

**Akademia Ekonomiczna w Krakowie**

**Wydział Zarządzania**

**Katedra Ekonometrii**

**Jerzy Marzec**

**Ekonometryczna analiza efektywności kosztów  
w bankach komercyjnych**

**Praca doktorska**

**napisana pod kierunkiem**

**Prof. AE dra hab. Jacka Osiewalskiego**

**Kraków, kwiecień 2000**

## Spis treści:

Wstęp.....	1
1. Mikroekonomiczne podstawy analizy efektywności kosztowej firm .....	7
1.1. Funkcja produkcji, kosztów i zysku a kategorie efektywności ekonomicznej.....	7
1.2. Modelowanie nieefektywności kosztowej.....	16
1.3. Charakterystyki procesu produkcji.....	23
2. Bank jako przedsiębiorstwo - ujęcie mikroekonomiczne.....	31
2.1. Model instytucji finansowych wg Sealeya i Lindleya.....	31
2.2. Czynniki produkcji i produkty banku w badaniach efektywności ekonomicznej.....	37
2.3. Funkcja kosztów w przypadku oddziałów banku komercyjnego.....	42
3. Elementy klasycznej estymacji modeli granicznych.....	50
3.1. Podstawowe techniki w przypadku danych przekrojowych.....	50
3.2. Model efektów stałych w przypadku danych przekrojowo-czasowych.....	56
4. Bayesowska analiza efektywności.....	60
4.1. Elementy podejścia bayesowskiego.....	60
4.2. Bayesowski stochastyczny model graniczny o wspólnym rozkładzie efektywności.....	62
4.3. Próbkowanie Gibbsa jako numeryczna metoda wyznaczania rozkładów a posteriori.....	67
4.4. Specyfikacja i testowanie egzogenicznych przyczyn nieefektywności.....	71
5. Empiryczna analiza efektywności kosztowej oddziałów.....	78
5.1. Charakterystyka danych źródłowych.....	78
5.2. Estymacja modelu granicznej funkcji kosztu z wykorzystaniem MNK.....	86
5.3. Estymacja bayesowskiego modelu losowych efektów indywidualnych (o zmiennym rozkładzie efektywność - VED).....	93
5.4. Estymacja modelu bayesowskiego modelu VED – przypadek dwóch produktów.....	118
Podsumowanie.....	135
Wykaz pozycji literaturowych.....	142
Załączniki.....	149

## Wstęp

Jednym z fundamentalnych zagadnień rozważanych na gruncie mikroekonomii jest problem maksymalizacji zysku firmy, rozważanego jako różnica między osiąganym przychodem a kosztem jego uzyskania. Warunkiem koniecznym maksymalizacji zysku jest minimalizacja kosztu produkcji. Przedsiębiorstwo osiąga zysk niższy od maksymalnego, jeżeli nie wytwarza produkcji po minimalnym koszcie, czyli gdy przy danych cenach czynników produkcji ponosi koszt całkowity wytworzenia określonego poziomu produkcji większy niż wynika on z mikroekonomicznej (granicznej) funkcji kosztu. Sytuację tę określamy mianem nieefektywności kosztowej bądź nieefektywności kosztów.

Przedsiębiorstwo ponosi koszty większe niż niezbędne (jest nieefektywne kosztowo; ang. *cost inefficient*) jeżeli 1<sup>o</sup> jest nieefektywne technicznie, tj. angażuje zbyt duże (w stosunku do wymagań technologicznych) nakłady czynników produkcji w celu osiągnięcia danej wielkości produkcji, lub 2<sup>o</sup> jest nieefektywne alokacyjnie, tj. proporcje nakładów czynników nie odpowiadają relacji ich cen rynkowych (dla zastosowanej kombinacji nakładów krańcowe stopy substytucji nie są równe odpowiednim ilorazom cen czynników). **Przez efektywność kosztową bądź efektywność kosztów (ang. *cost efficiency*) rozumie się iloraz minimalnego kosztu niezbędnego do wytworzenia danej wielkości produkcji (przy danych cenach czynników) i kosztu rzeczywiście poniesionego<sup>1</sup>.** Analogicznie, pod pojęciem efektywności technicznej rozumiemy iloraz wielkości produkcji rzeczywiście wytworzonej do maksymalnej, możliwej do uzyskania przy obserwowanych nakładach czynników produkcji.

Farrell w 1957 roku jako pierwszy teoretycznie rozważał możliwości empirycznej analizy efektywności produkcji. Podstawy obecnie stosowanej ekonometrycznej metodologii

---

<sup>1</sup> Powszechnie przez słowo „efektywność” rozumie się *pozytywny wynik, wydajność, skuteczność, sprawność*, a w polskiej terminologii ekonomicznej pojęcie *efektywność ekonomiczna* zwykle oznacza: *rezultat działalności gospodarczej, określony przez stosunek uzyskanego efektu do nakładu*. (źródło: *Słownik języka polskiego*, pod redakcją Mieczysława Szymczaka, PWN, Warszawa 1978). W niniejszej pracy autor bardziej skłania do tłumaczenia *cost efficiency* jako „efektywność kosztowa”, a nie „efektywność kosztów” bądź np. „wydajność kosztów”, bo (po pierwsze) analogicznie mówi się o efektywności ekonomicznej, technicznej, oraz (po drugie) termin efektywność ma - wg zaproponowanej definicji - inne znaczenie niż jego potoczne, słownikowe

badania efektywności (technicznej lub kosztowej) stworzyły w 1977 roku dwa zespoły badawcze: Aigner, Lovell i Schmidt oraz Meeusen i van den Broeck, którzy niezależnie od siebie zaproponowali tzw. stochastyczne modele graniczne (ang. *stochastic frontier models*). Zgodnie z tą metodologią pomiar efektywności (technicznej lub kosztowej) dokonywany jest zwykle za pomocą modelu jednorównaniowego składającego się z odpowiednio wyspecyfikowanej mikroekonomicznej funkcji produkcji lub kosztów (dla logarytmów tych zmiennych) oraz dwóch składników losowych, z których jeden (symetryczny względem zera) odzwierciedla efekt czynników przypadkowych i błędów pomiaru, zaś drugi (asymetryczny i stałego znaku) modeluje potencjalną nieefektywność. Dalszego rozwoju tej metodologii, prezentowanej głównie na łamach *Journal of Econometrics*, dokonali m. in. Stevenson [1980], Pitt i Lee [1981], Jondrow, Lovell, Materov i Schmidt [1982], Schmidt i Sickles [1984], Beckers i Hammond [1987], Greene [1980], [1990], van den Broeck, Koop, Osiewalski i Steel [1994] oraz Koop, Osiewalski i Steel [1994], [1997].

Jednym z obszarów zastosowania stochastycznych modeli granicznych jest zagadnienie efektywności kosztowej banków i innych instytucji finansowych, prezentowane od połowy lat osiemdziesiątych w czołowej specjalistycznej literaturze zachodniej, głównie na łamach *Journal of Banking and Finance* (JBF), *Journal of Productivity Analysis* (JPA) oraz *Journal of Money, Credit and Banking* (JMCB) (w 1993 roku całe numery JBF i JPA były poświęcone wyłącznie badaniu efektywności instytucji finansowych). Dość późny rozwój tych zastosowań (w stosunku do aplikacji dla przedsiębiorstw z innych branż) był spowodowany specyficznym charakterem działalności banków, który powoduje, że należy z dużą ostrożnością podchodzić do specyfikacji granicznej funkcji kosztu. Znaczący wkład w rozwój teoretyczny podstaw tych badań mają Sealey i Lindley, którzy w 1977 roku zaproponowali mikroekonomiczny model banków (i innych instytucji finansowych) oraz określili czynniki produkcji i produkty bankowe, umożliwiając tym samym budowę granicznej funkcji kosztu dla banków.

W polskiej literaturze naukowej, zarówno statystyczno-ekonometrycznej jak i bankowej, brak jest pozycji poświęconych aspektom metodologicznym i badaniom empirycznym z zakresu efektywności technicznej czy kosztowej banków. Warto dodać, że w ostatnich paru latach ukazały się popularyzatorskie artykuły przeglądowe Rogowskiego (m.in.. Bank nr 1, styczeń 1998, str. 24-27) na temat analizy efektywności oddziałów banku;

---

synonimy, tj. wydajność, sprawność, skuteczność, które mają ponadto ustalone i odmienne znaczenia na gruncie ekonomii, fizyki i techniki oraz prakseologii i nauk o zarządzaniu.

zob. też Rogowski [1998c], Kopczewski [1999]. Autor ten wskazał na ogromną potrzebę pomiaru i porównań efektywności ekonomicznej oddziałów, co może „stanowić cenne uzupełnienie tradycyjnych metod analiz wskaźnikowych stosowanych przez banki”. Uważa on, że zastosowanie metod ekonometrycznych (które podejmuje autor w tej pracy) „umożliwi między innymi obiektywizację kryteriów oceny oddziałów”.

Istniejące opracowania dotyczące analizy i oceny działalności banków (a także przedsiębiorstw należących do jednej branży - tzw. analizy branżowe) w większości są sporządzane w oparciu o analizę wskaźnikową. Prezentowane metody rzadko wykorzystują formalne wnioskowanie statystyczne i w małym stopniu nawiązują do zmatematyzowanej teorii mikroekonomii. Opracowania w języku polskim wskazują zatem na potrzebę opisu problemu oceny działalności banków komercyjnych językiem mikroekonomii z wykorzystaniem zaawansowanych metod statystyczno-ekonometrycznych, zgodnie z tendencjami panującymi w światowej ekonomii.

Głównymi motywami podjęcia tematu pracy są zatem: chęć przedstawienia metod ekonometrycznych (głównie bayesowskich) analizy efektywności kosztowej firm i potrzeba zainicjowania takich właśnie badań empirycznych w polskich bankach komercyjnych, w szczególności w celu zobiektywizowania ocen działalności oddziałów. W badaniach tej pracy wykorzystuje się rzeczywiste dane pochodzące z jednego z dużych polskich banków; warunkiem ich udostępnienia przez bank było jednak nie ujawnianie tych danych, ani źródła ich pochodzenia.

Zarówno pomiar efektywności w kontekście stochastycznego modelu granicznego, jak i stosowane w tym celu metody bayesowskie (które proponują i rozwijają Koop, Osiewalski i Steel [1994], [1997], [1999]) stanowią na polskim gruncie zupełnie nowe podejście do zagadnienia analizy i oceny działalności banków oraz ich oddziałów. Wydaje się, że metody konkurencyjne nie mają tak silnych podstaw ekonomicznych i statystycznych. Przeprowadzona w tej pracy analiza empiryczna umożliwiła pomiar efektywności kosztowej badanych oddziałów, sporządzenie ich rankingu i zbadanie przyczyn występowania nieefektywności. Rezultaty badań empirycznych nad efektywnością i pomiar charakterystyk technologii mogą być zatem wykorzystane w praktyce. Pomiar i analiza efektywności kosztowej może stać się narzędziem pomocniczym w zarządzaniu i polityce ekonomicznej banku, gdyż informuje o skali możliwych do uzyskania oszczędności (możliwego zwiększenia zysku) przy zachowaniu danej skali działalności banku. Efektywność banku stanowi istotny element jego konkurencyjności. Prezentowana metodologia ekonometryczna może być wykorzystywana w badaniach efektywności technicznej i kosztowej innych

jednorodnych grup przedsiębiorstw (należących do tej samej branży, np. wg Europejskiej Klasyfikacji Działalności).

Wśród banków europejskich została przeprowadzona w 1992 roku ankieta dotycząca problemów, które towarzyszą zjawisku obniżania się rentowności tamtejszych banków. Wśród czynników ekonomicznych, praktycy bankowi wymieniają przede wszystkim: istnienie wolnych mocy produkcyjnych, braki w technikach zarządzania, brak efektywnego systemu bankowej rachunkowości zarządczej, niska jakość personelu, niska kapitalizacja w stosunku do potrzeb (ekspansji), przerosty w zatrudnieniu oraz niedostosowane do rzeczywistości cenniki za usługi (zob. Groszek M., 1997, „Banki komercyjne wobec procesu Europejskiej Integracji Walutowej”, *Bank i Kredyt*, 6). Są to problemy aktualne również w odniesieniu do polskich banków komercyjnych. Podjęta w pracy metodologia pozwala ilościowo scharakteryzować proces produkcyjny, określić stopień wykorzystania nakładów finansowych i dostarczyć informacji istotnych dla potrzeb zarządzania bankiem jako przedsiębiorstwem. Wyniki empiryczne tej pracy mogłyby być wykorzystane m.in. przy optymalizacji poziomu kosztu i zatrudnienia banku oraz w celu określenia podstawowych obszarów, w których można poszukiwać oszczędności kosztowych.

Głównym celem pracy jest bayesowska analiza mikroekonomicznej efektywności kosztowej oddziałów polskiego banku komercyjnego za pomocą stochastycznych modeli granicznych. Na cele cząstkowe składają się:

1. określenie czynników produkcji i produktów oddziałów badanego banku komercyjnego,
2. konstrukcja stochastycznego modelu granicznego opisującego nieefektywność kosztową,
3. opis i zastosowanie metod bayesowskiej analizy modeli granicznych,
4. pomiar indywidualnej efektywności oddziałów banków na podstawie próby statystycznej,
5. określenie determinant zróżnicowania efektywności kosztowej oddziałów,
6. próba zbadania występowania korzyści pełnego zakresu produkcji bądź specjalizacji i ewentualnego pomiaru tych korzyści.

W celu realizacji powyższych celów przyjęto kilka założeń wyjściowych będących przedmiotem weryfikacji. Podstawową hipotezą jest przypuszczenie, że stochastyczny model graniczny jest dogodnym narzędziem pomiaru efektywności kosztowej, natomiast podejście bayesowskie jest uniwersalnym i elastycznym sposobem wnioskowania o charakterystykach funkcji kosztów i wskaźnikach efektywności, umożliwiającym pełne uwzględnienie założeń mikroekonomicznych i wstępnych informacji. Hipotezą empiryczną jest przypuszczenie, że można wyróżnić czynniki egzogeniczne, które wyjaśniają potencjalne różnice w poziomach

efektywności oddziałów. Wnioskowanie bayesowskie dostarcza wygodnych narzędzi do weryfikacji statystycznej tej hipotezy empirycznej.

Niniejsza praca ma charakter metodologiczno-empiryczny i składa się ze wstępu, pięciu rozdziałów, zakończenia, spisu literatury oraz załączników. W rozdziale pierwszym zostały przedstawione mikroekonomiczne podstawy analizy efektywności kosztowej firm. W części 1.1 przypomniano pojęcie funkcji produkcji, kosztu oraz zdefiniowano termin efektywności kosztowej. Następnie w podrozdziale 1.2 wprowadzono najprostszy jednorównaniowy stochastyczny model granicznej krótkookresowej funkcji kosztu zmiennego oraz zaproponowano i omówiono funkcję translogarytmiczną jako podstawową postać analityczną wykorzystywaną w niniejszej pracy. W części 1.3 podano podstawowe charakterystyki wieloproduktowego procesu produkcji, tj. elastyczności, różne definicje miar efektu skali produkcji oraz miar korzyści pełnego zakresu bądź specjalizacji.

W rozdziale drugim przedstawiono mikroekonomiczny model instytucji finansowych zaproponowany przez Sealeya i Lindleya. Część 2.2 stanowi przegląd literatury światowej dotyczącej analizy efektywności kosztowej banków komercyjnych ze zwróceniem szczególnej uwagi na problem definicji czynników produkcji i produktów w przypadku tych przedsiębiorstw. W części 2.3 autor podjął się określenia czynników produkcji i produktów w przypadku badanych oddziałów banku komercyjnego oraz specyfikacji krótkookresowej funkcji kosztu zmiennego wykorzystanej w części empirycznej.

W rozdziale trzecim przedstawiono i poddano krytyce podstawowe klasyczne techniki estymacji modeli granicznych i pomiaru efektywności, przede wszystkim skorygowaną metodę najmniejszych kwadratów (MNK) i metodę największej wiarygodności (MNW) w przypadku danych przekrojowych oraz model efektów stałych i estymator wewnątrzgrupowy w przypadku danych przekrojowo-czasowych.

Rozdział czwarty (a zwłaszcza część 4.2 i 4.4) stanowi przede wszystkim omówienie bayesowskiego stochastycznego modelu granicznego o wspólnym rozkładzie efektywności dla losowych efektów indywidualnych (CED) oraz jego uogólnienia, tj. modelu o zmiennym rozkładzie efektywności (VED). Wstępna część 4.1 przedstawia zarys bayesowskich reguł wnioskowania statystycznego. Prezentowane jest również (w części 4.3) losowanie Gibbsa jako numeryczna metoda wyznaczania rozkładów a posteriori dla nieznanymi parametrów granicznej funkcji kosztów i dla wskaźników efektywności.

Rozdział piąty prezentuje empiryczną analizę efektywności kosztowej oddziałów. Część 5.1 poświęcona jest prezentacji danych. Następnie w części 5.2 przedstawione są

rezultaty estymacji granicznej funkcji kosztu i pomiaru wskaźników efektywności dla zmodyfikowanej funkcji Cobba i Douglasa, uzyskane na podstawie danych przekrojowych za pomocą skorygowanej MNK. Części 5.3 i 5.4 stanowią opis wyników empirycznych uzyskanych dla bayesowskiego modelu losowych efektów indywidualnych (o zmiennym rozkładzie efektywności - VED) w przypadku jednego zaagregowanego produktu oraz dwóch produktów; podrozdziały te zawierają m.in. interpretację charakterystyk procesu produkcji. Przedstawiono także wyniki pomiaru efektywności oraz testowania istotności egzogenicznych przyczyn zróżnicowania efektywności oddziałów. Pracę zamyka podsumowanie oraz załączniki zawierające uzupełniające, cząstkowe wyniki empiryczne.

Niniejsze opracowanie jest efektem współpracy naukowej autora z prof. AE dr hab. Jackiem Osiewalskim rozpoczętej w ramach grantu KBN nr 1-H02B-015-11, a kontynuowanej w ramach grantu KBN nr 1-H02B-022-18. Podstawy metodologiczne i wstępne wyniki tych badań zostały opublikowane w czasopiśmie ogólnopolskich bądź były referowane na konferencjach naukowych; zob. Marzec i Osiewalski [1996-97], Marzec [1998a, 1998b], [1999], Osiewalski, Marzec i Pipień [1999], Osiewalski i Marzec [1998a, 1998b, 1998c].

Pragnę złożyć wyrazy serdecznego podziękowania Profesorowi Jackowi Osiewalskiemu za nieocenioną pomoc i zaangażowanie, sugestie, dyskusje naukowe i wnikliwe uwagi merytoryczne w tworzeniu niniejszej pracy.



# 1. Mikroekonomiczne podstawy analizy efektywności kosztowej firm

## 1.1. Funkcja produkcji, kosztów i zysku a kategorie efektywności ekonomicznej.

Podstawą do stworzenia modelu opisującego działalność firmy jest określenie jej celów oraz warunków ograniczających tę działalność, a narzucanych przez otoczenie. Mikroekonomiczny model przedsiębiorstwa jako układ założeń przedstawia zachowanie się firmy, która prowadzi działalność przy pewnych ograniczeniach, będących efektem istnienia rynku konsumentów i producentów oraz działania przyrody. Przedsiębiorstwo dąży do realizacji określonych celów, z których najważniejszym i docelowym jest maksymalizacja zysku<sup>2</sup>, rozumianego jako różnica między przychodem a kosztem jego uzyskania. Dążąc do tego celu przedsiębiorstwo określa m.in. rozmiary (skalę) produkcji i decyduje o jej strukturze napotykając na ograniczenia ze strony przyrody i technologii, która wyznacza sposoby wytwarzania produkcji z dostępnych nakładów czynników produkcji.

Niech wektor  $Q=(Q_1, \dots, Q_G) \in R^G_+$  oznacza wektor  $G$  produktów otrzymanych w wyniku zaangażowania w procesie produkcyjnym  $H$  czynników produkcji, których nakłady reprezentuje wektor  $x=(x_1, \dots, x_H) \in R^H_+$ . Ograniczenia technologiczne są opisywane przez kombinacje nakładów i wyników (produktów), czyli zbiór  $Z=\{(x, Q)\}$  taki, że z wektora nakładów  $x$  można otrzymać wektor produktów  $Q$ . Zbiór ten obejmuje technologicznie wykonywalne sposoby produkcji, i jest nazywany zbiorem możliwości produkcyjnych lub zbiorem produkcyjnym (ang. *production possibility set* lub *technology set*), opisującym możliwe wybory technologiczne stojące przed firmą. Przedmiotem szczególnego zainteresowania firmy jest taki element zbioru produkcyjnego (plan)  $\{x, Q\}$ , który umożliwia uzyskanie największej możliwej produkcji z danej ilości nakładów czynników. Taki plan produkcji określa się jako technicznie (technologicznie) efektywny.

---

<sup>2</sup> Maksymalizacja zysku nie stanowi nadrzędnego celu działania szczególnego rodzaju przedsiębiorstw tzw. *non-profit* (np. przedsiębiorstw użyteczności publicznej), dla których głównym celem jest minimalizacja kosztów wytworzenia określonej wielkości produkcji.

W analizie procesu produkcji (budując model matematyczny) dla uproszczenia często przyjmuje się, że firma wytwarza jeden produkt albo produkcję wyraża się w postaci agregatu mierzonego w jednostkach pieniężnych (przy cenach stałych). W tym przypadku funkcja opisująca ten zbiór  $\{x, Q\}$  nazywana jest mikroekonomiczną funkcją produkcji  $f(x_1, \dots, x_H)$  lub graniczną funkcją produkcji (ang. *frontier production function*), co formalnie zapisuje się  $f(x_1, \dots, x_H) = \max\{Q: (x_1, \dots, x_H) \in V(Q)\}$ , gdzie  $V(Q) = \{x \in R^H_+ : \{x, Q\} \in Z\}$ , jest zbiorem nakładów czynników, z których można uzyskać produkcję  $Q$ . Funkcja produkcji wyraża maksymalną produkcję  $Q$  możliwą do uzyskania z danych nakładów czynników  $x_1, \dots, x_H$ . W przypadku rozważania  $G$  produktów firmy ( $G \geq 1$ ) analogonem funkcji produkcji jest funkcja transformacji  $T: R^{G+H}_+ \rightarrow R$  taka, że  $T(Z^*) = 0$ , gdzie  $Z^*$  oznacza (technologicznie) efektywny plan produkcyjny. Przez efektywny plan produkcyjny rozumie się taką kombinację produktów  $Q_1, \dots, Q_G$  i nakładów czynników  $x_1, \dots, x_H$ , że nie istnieje inna, która umożliwia wytworzenie większych wielkości produktów przy tych samych nakładach czynników lub wytworzenie tych samych wielkości produktów przy mniejszych nakładach czynników (Varian [1992]). Z własności zbiorów produkcyjnych  $\{x, Q\}$  wynika, że funkcja produkcji jest  $1^0$  niemalejąca ze względu na czynniki produkcji,  $2^0$  wklęsła<sup>3</sup>. Z historycznego punktu widzenia najczęściej stosowanymi w badaniach empirycznych funkcjami produkcji (a także funkcjami kosztu) są funkcje jednorodne stopnia  $\nu > 0$  (ang. *homogeneous function*), Szerszą klasę stanowią funkcje jednokładne<sup>4</sup> (ang. *homothetic function*), których szczególnym przypadkiem są funkcje jednorodne. W pracach teoretycznych z zakresu teorii firm od początku lat siedemdziesiątych, a w pracach empirycznych od początku lat osiemdziesiątych wprowadza się i upowszechnia ogólniejsze formy funkcyjne np. translogarytmiczna funkcja produkcji lub kosztu. Wykorzystanie tej ostatniej w analizie efektywności kosztowej będzie przedmiotem szczególnego zainteresowania w niniejszej pracy.

Przed firmą zorientowaną na zysk jawią się dwa problemy decyzyjne. Dotyczą one wyboru metod wytwarzania i planów produkcji. Pierwszy z nich dotyczy wyboru takich nakładów czynników ekonomicznie niezbędnych do uzyskania produkcji na poziomie  $Q$ , które gwarantują minimalny koszt jej wytworzenia<sup>5</sup>. Opisuje się go za pomocą modelu

<sup>3</sup>  $f: R^n \rightarrow R$  jest funkcją wklęsłą, gdy  $f(tx_1 + (1-t)x_2) \geq tf(x_1) + (1-t)f(x_2)$  dla każdego  $x_1, x_2$  i  $0 \leq t \leq 1$ ; Varian [1992].

<sup>4</sup>  $f$  jest funkcją jednokładną wtedy i tylko wtedy, gdy  $f(x) = g(h(x))$ , gdzie  $h(\cdot)$  jest funkcją jednokładną stopnia 1, a  $g(\cdot)$  jest monotonicznym przekształceniem  $g: R \rightarrow R$ ; Varian [1992].

<sup>5</sup> Jak już wspomniano, minimalizacja kosztu jest postulatem uniwersalnym, którym winny się kierować wszystkie jednostki produkcyjne (w tym *non-profit*).

minimalizacji kosztu. Natomiast model maksymalizacji zysku ujmuje problem wyboru takiej wielkości produkcji (i jej struktury, jeśli  $G > 1$ ), która pozwoli na uzyskanie maksymalnego przychodu po najniższym koszcie, a w efekcie maksymalnego zysku. Minimalizacja kosztów jest warunkiem koniecznym dla maksymalizacji zysku.

Budując obydwie modele zakłada się, że firma napotyka na ograniczenia rynkowe - ceny używanych czynników i wyrobów gotowych są egzogeniczne, określone przez rynek, na którym żadna z firm nie ma pozycji dominującej - pozycji monopolisty lub zbliżonej. W tym przypadku problem minimalizacji kosztów, gdy rozważa się jeden produkt (lub produkcję wyraża się w postaci agregatu wyrażonego w jednostkach pieniężnych) i  $H$  czynników produkcji o ustalonych cenach  $w_1, \dots, w_H$ , polega na znalezieniu nakładów czynników  $x_1, \dots, x_H$ , które minimalizują koszty całkowite ( $TC$ ) wytworzenia danego poziomu produkcji  $Q = f(x_1, \dots, x_H)$ , gdzie  $f(x_1, \dots, x_H)$  jest mikroekonomiczną funkcją produkcji,  $TC = (x_1 \cdot w_1 + \dots + x_H \cdot w_H)$  a  $Q$  to zadana (ustalona) wielkość produkcji. Rozwiązanie problemu, czyli minimalny koszt całkowity niezbędny do osiągnięcia pożądanego poziomu produkcji  $Q$ ,  $c(w_1, \dots, w_H, Q)$ , jest funkcją cen czynników i wielkości produkcji  $Q$ .

Nakłady czynników produkcji  $x_1^*(w_1, \dots, w_H, Q)$ , ...,  $x_H^*(w_1, \dots, w_H, Q)$  zapewniające firmie minimalny koszt całkowity, również zależą od cen czynników i poziomu produkcji; określane jako warunkowe funkcje popytu na czynniki. Przedstawiony wyżej model dotyczy sytuacji, gdy firma minimalizująca koszt mogła zmieniać nakłady wszystkich czynników produkcji. Otrzymana funkcja kosztu  $c(w_1, \dots, w_H, Q)$  nazywana jest wówczas długookresową funkcją kosztu. Odzwierciedla ona minimalny koszt całkowity wytworzenia danego poziomu produkcji  $Q$  przy dostosowaniu wszystkich czynników produkcji. Problem maksymalizacji zysku firmy, gdy ma ona swobodę wyboru poziomu wszystkich ponoszonych nakładów (tj. w długim okresie), polega na wyborze takiej skali produkcji, która gwarantuje maksymalny zysk - największą różnicę między przychodem a kosztem jego uzyskania. Zapisuje się go następująco:

$$\max_Q [p \cdot Q - c(w_1, \dots, w_H, Q)],$$

gdzie:  $p$  - cena produktu (w przypadku produkcji jednoasortymentowej) lub uśredniona cena wielu produktów, a  $Q$  jest agregatem produkcji wyrażonym w jednostkach pieniężnych.

O ile analiza długookresowa przedstawiona zostanie bardzo zwięźle i jedynie dla przypadku jednoproduktowego ( $G=1$ ), to dalsze rozważania (ważniejsze dla pracy) dotyczą analizy krótkookresowej w przypadku produkcji wieloasortymentowej.

Jeżeli w badanym okresie zaangażowanie chociaż jednego z czynników produkcji, np.  $H$ -tego, nie podlega zmianie, wtedy minimalny koszt wytworzenia  $G$  produktów (o wielkości  $Q_1, \dots, Q_G$ ) przy zmianie jedynie czynników podlegających optymalizacji (czynników zmiennych:  $x_1, \dots, x_{H-1}$ ) jest rozwiązaniem zagadnienia:

$$\min_{x_1, \dots, x_{H-1}} [x_1 \cdot w_1 + \dots + x_{H-1} \cdot w_{H-1} + \tilde{x}_H \cdot w_H] \text{ przy warunku } T(Q_1, \dots, Q_G, x_1, \dots, x_{H-1}, \tilde{x}_H) = 0,$$

$$\text{czyli } \tilde{x}_H \cdot w_H + \min_{x_1, \dots, x_{H-1}} [x_1 \cdot w_1 + \dots + x_{H-1} \cdot w_{H-1}] \text{ przy warunku } T(Q_1, \dots, Q_G, x_1, \dots, x_{H-1}, \tilde{x}_H) = 0,$$

gdzie:  $\tilde{x}_H$  - zaangażowanie czynnika stałego, nie podlegającego optymalizacji. Koszt całkowity zaangażowania optymalnych nakładów czynników  $x_1^*, \dots, x_{H-1}^*$  zależy od cen wszystkich czynników  $w_1, \dots, w_H$ , wielkości produktów  $Q_1, \dots, Q_G$  i dodatkowo nakładów czynnika stałego  $\tilde{x}_H$ , i nosi nazwę krótkookresowej funkcji kosztu całkowitego (ang. *short-run total cost*). Na koszt całkowity,  $c^s(w_1, \dots, w_H, Q_1, \dots, Q_G, \tilde{x}_H)$ , składa się koszt zmienny  $VC = c^s_v(w_1, \dots, w_{H-1}, Q_1, \dots, Q_G, \tilde{x}_H)$  równy  $\sum_{i=1}^{H-1} (w_i \cdot x_i^*(w_1, \dots, w_{H-1}, Q_1, \dots, Q_G, \tilde{x}_H))$  i koszt stały  $c^s_f(\tilde{x}_H, w_H) = \tilde{x}_H \cdot w_H$ . Zatem krótkookresowa funkcja kosztu zmiennego (ang. *short-run variable cost*) -  $c^s_v(\cdot)$  - jest funkcją jedynie cen czynników zmiennych, wielkości produktów oraz nakładów czynników stałych (nie występuje w niej cena czynnika stałego). Problem maksymalizacji zysku w krótkim okresie może być zapisany następująco:

$$\begin{aligned} \max_{Q_1, \dots, Q_G} \left[ \sum_{i=1}^G p_i \cdot Q_i - c^s(w_1, \dots, w_H, Q_1, \dots, Q_G, \tilde{x}_H) \right] = \\ \max_{Q_1, \dots, Q_G} \left[ \sum_{i=1}^G p_i \cdot Q_i - c^s_v(w_1, \dots, w_{H-1}, Q_1, \dots, Q_G, \tilde{x}_H) \right] - w_H \cdot \tilde{x}_H. \end{aligned}$$

Powyższy zapis podkreśla fakt, że warunkiem koniecznym maksymalizacji zysku jest minimalizacja kosztu, por. Varian [1992].

Dotychczasowe rozważania można zilustrować przykładem krótkookresowej funkcji kosztu otrzymanej z funkcji produkcji typu Cobba-Douglasa. Dla uproszczenia przyjęto, że wielkość produkcji wyraża się poprzez  $Q$  ( $G=1$ ), zaangażowanie zmiennych czynników produkcji wynosi odpowiednio  $x_1, x_2$ , a stałego  $x_3$  ( $H=3$ ), gdzie ostatni z czynników nie podlega optymalizacji w procesie minimalizacji kosztu. Problem ten zapisuje się:

$$\begin{cases} \min_{x_1, x_2} (w_1 x_1 + w_2 x_2 + w_3 \tilde{x}_3) \\ Q = a x_1^b x_2^c \tilde{x}_3^d \end{cases}, \text{ przy czym } a, b, c, d > 0 \text{ i } w_1, w_2, w_3 > 0.$$

Wykorzystując metodę mnożników Lagrange'a w celu wyznaczenia ekstremum warunkowego otrzymano rozwiązanie - warunkowe funkcje popytu na czynniki produkcji - postaci:

$$x_1^*(w_1, w_2, Q, \bar{x}_3) = a^{-1/(b+c)} \left(\frac{w_2}{w_1}\right)^{c/(b+c)} \left(\frac{b}{c}\right)^{c/(b+c)} Q^{1/(b+c)} \bar{x}_3^{-d/(b+c)},$$

$$x_2^*(w_1, w_2, Q, \bar{x}_3) = a^{-1/(b+c)} \left(\frac{w_1}{w_2}\right)^{b/(b+c)} \left(\frac{c}{b}\right)^{b/(b+c)} Q^{1/(b+c)} \bar{x}_3^{-d/(b+c)}.$$

Zatem na minimalny koszt całkowity  $c^s(w_1, w_2, Q, \bar{x}_3)$  składa się koszt stały (niezależny od poziomu produkcji)  $c^s_f(\bar{x}_3, w_3) = \bar{x}_3 \cdot w_3$  oraz koszt zmienny  $c^s_v(w_1, w_2, Q, \bar{x}_3)$  równy  $\beta_0 w_1^{\beta_1} w_2^{\beta_2} Q^{\beta_3} \bar{x}_3^{\beta_4}$ , gdzie:  $\beta_0 = a^{-1/(b+c)} \left( \left(\frac{b}{c}\right)^{c/(b+c)} + \left(\frac{c}{b}\right)^{b/(b+c)} \right)$ ,  $\beta_1 = b/(b+c)$ ,  $\beta_2 = c/(b+c)$ ,  $\beta_3 = 1/(b+c)$ ,  $\beta_4 = -d/(b+c)$ .

W przypadku funkcje kosztu zmiennego można mówić o ogólnym efekcie skali produkcji i efekcie skali produkcji względem czynników zmiennych. Pierwszy z nich - ogólny efekt skali produkcji - określa wpływ proporcjonalnych zmian nakładów wszystkich czynników na wielkość produkcji, a jego miarą jest współczynnik efektu skali, który dla przypadku omawianej funkcji kosztu typu Cobba-Douglasa wynosi  $(1-\beta_4)/\beta_3$ , czyli  $b+c+d$ . Natomiast efekt skali produkcji względem czynników zmiennych informuje o ile procent wzrośnie wielkość produkcji na skutek proporcjonalnego wzrostu nakładów jedynie czynników zmiennych; w omawianym przypadku współczynnik efektu skali wynosi  $1/\beta_3$ , czyli  $b+c$ . Zatem w omawianym przypadku współczynnik ogólnego efektu może być większy albo mniejszy od współczynnika efektu skali względem czynników zmiennych w zależności od znaku parametru  $\beta_4$ , a dokładniej od sytuacji substytucji czynników produkcji bądź ich komplementarności. Wyprowadzenie wzoru na współczynnik ogólnego efektu skali produkcji w przypadku translogarytmicznej funkcji kosztu zmiennego nie jest łatwe, zatem w części dotyczącej empirycznej analizy produkcji będzie wykorzystywany współczynnik efektu skali względem tylko czynników zmiennych.

Otrzymana w tym przypadku krótkookresowa funkcja kosztu zmiennego jest funkcją potęgową, a warunki regularności ekonomicznej nakładają następujące restrykcje na parametry  $\beta_i > 0$  dla  $i=1, \dots, 3$  i  $\beta_4 < 0$ . Ujemny znak parametru przy poziomie zaangażowania czynnika stałego ( $\beta_4$ ) wynika ze zjawiska substytucji czynników produkcji dla przyjętej funkcji produkcji (Cobba-Douglasa). Warto zauważyć, iż z własności funkcji produkcji (niemalejąca ze względu na czynniki produkcji) wynika możliwość występowania zarówno komplementarności, gdy jest ona stała względem jednego z czynników, jak i substytucji

między czynnikami (w przypadku silnej monotoniczności). Prowadzi to do następującego stwierdzenia: jeżeli funkcja produkcji, z której otrzymano krótkookresową funkcję kosztu, dopuszcza substytucję czynników, to elastyczność względem zaangażowania czynnika stałego dla funkcji kosztu zmiennego powinna być ujemna. W przypadku funkcji produkcji dopuszczającej komplementarność czynników, np. Leontiewa, co prawda funkcja warunkowego popytu na zmienny czynnik produkcji nie zależy od nakładów czynników stałych, więc nie pojawią się one jako zmienne w równaniu krótkookresowej funkcji kosztu zmiennego, lecz dziedzina tej funkcji jest ograniczona przez nakłady czynników stałych. Zatem w przypadku funkcji Leontiewa, rozwiązanie problemu:

$$\begin{cases} \min_{x_1, x_2} (w_1 x_1 + w_2 x_2 + w_3 \tilde{x}_3) \\ Q = \min\{ax_1, bx_2, c\tilde{x}_3\} \end{cases} \text{ przy czym } a, b, c > 0 \text{ i } w_1, w_2, w_3 > 0,$$

prowadzi do funkcji kosztu zmiennego postaci:  $c^s_v(w_1, w_2, Q) = Q(w_1/a + w_2/b)$  dla  $Q \in (0, c\tilde{x}_3]$ . Zatem specyfikacja zmiennych (argumentów) krótkookresowej funkcji kosztu ma podstawy w mikroekonomicznej teorii firm i zależy m.in. od tego czy dopuszcza się komplementarność czynników, czy nie.

Przedmiotem dalszych rozważań będzie właśnie krótkookresowa funkcja kosztu zmiennego. Własności tej funkcji implikują restrykcje na parametry, które powinno się ewentualnie uwzględnić przy ich estymacji. Zakłada się, że jest ona:

- niemalejąca ze względu na ceny czynników zmiennych ( $w_1, \dots, w_{H-1}$ ) i produkcję ( $Q_1, \dots, Q_G$ ),
- jednorodna (homogeniczna) stopnia jeden ze względu na ceny czynników zmiennych,
- nierosnąca ze względu na nakłady czynników stałych, jeśli zakłada się substytucyjność.

Warunki regularności ekonomicznej wymagają, aby elastyczności funkcji kosztów względem produktów oraz cen czynników zmiennych (podlegających optymalizacji w procesie minimalizacji kosztów) były nieujemne i aby te ostatnie sumowały się do jedności. Własność homogeniczności można explicite narzucić dokonując prostej operacji wydzielenia kosztu i każdej z cen przez jedną z nich. Narzucenie globalnych warunków (na całą dziedzinę funkcji kosztu) ma praktyczne znaczenie, ponieważ zmniejsza liczbę parametrów funkcji. Narzucenie pozostałych restrykcji jest trudne i nie zawsze możliwe w przypadku klasycznych metod estymacji, w przeciwieństwie do podejścia bayesowskiego, które poprzez rozkład a priori pozwala uwzględnić nawet nieliniowe i nierównościowe restrykcje na parametry. W praktyce, stosując podejście bayesowskie, uwzględnia się warunki monotoniczności funkcji kosztu jedynie lokalnie, tj. dla wielkości zmiennych obserwowanych w próbie.

Mikroekonomiczna (tzw. graniczna) funkcja kosztów, otrzymana jako rozwiązanie problemu minimalizacji kosztów, ma szerokie zastosowanie w analizie empirycznej kosztów firmy. Jednym z obszarów jej wykorzystania jest analiza efektywności kosztowej firmy, a w szczególności instytucji finansowych (w tym banków komercyjnych lub ich oddziałów). W badaniach efektywności kosztowej przedmiotem badania jest grupa przedsiębiorstw należących do jednej branży i posiadających dostęp do tej samej myśli technicznej (technologii). Zakłada się, że firmy te są dla siebie konkurentami zarówno na rynku czynników produkcji, jak i produktów (oznacza to, że pozyskują czynniki produkcji i sprzedają swoje produkty po cenach kształtowanych przez rynek i żadna z nich nie ma pozycji monopolistycznej lub zbliżonej). Obserwuje się, że firmy - ze względu na swój potencjał - angażują różne ilości tych samych czynników, które w procesie produkcji są w odmienny sposób zużywane powodując, że każda z firm posiada inny poziom i strukturę produkcji. Zakłada się, że wszystkie firmy prowadząc działalność nastawioną na osiągnięcie zysku dążą do jego maksymalizacji, a stopień wykonania tego celu jest swoistym kryterium oceny przedsiębiorstwa. Można zatem postawić pytanie, która z firm robi to najlepiej?

Odpowiedź wydaje się być prosta: firma maksymalizuje zysk, gdy wytwarza produkcję po najniższym koszcie i jednocześnie uzyskuje maksymalny przychód z jej sprzedaży (skończone rozwiązanie istnieje w przypadku procesu produkcji charakteryzującego się malejącym efektem skali). W przypadku firmy w pełni maksymalizującej zysk powinny być spełnione trzy warunki: 1<sup>o</sup> dostępna technologia produkcji wykorzystywana jest w sposób najlepszy z możliwych (liczony w pewien sposób stosunek efektu produkcyjnego do poniesionych nakładów jest maksymalny), 2<sup>o</sup> struktura nakładów czynników odpowiada relacji ich cen rynkowych, 3<sup>o</sup> skala i struktura produkcji odpowiada cenom produktów (warunek na strukturę ma znaczenie tylko w przypadku produkcji wieloasortymentowej). Istnieje zatem potrzeba dokonania porównań i oceny firm ze względu na przyjęte kryterium. Analiza efektywności technicznej pozwoli na zbadanie, w jakim stopniu spełniony jest pierwszy warunek, natomiast analiza efektywności kosztowej bada spełnienie jednocześnie dwóch pierwszych warunków. Badanie czy przedsiębiorstwo maksymalizuje zysk, czyli czy spełnione są wszystkie trzy warunki, możliwe jest dopiero w oparciu o funkcję zysku, jednakże analiza tego problemu wychodzi poza ramy niniejszej pracy. Pierwsze próby badań empirycznych ogólnej efektywności ekonomicznej na podstawie funkcji zysku, podjęli dla banków: Berger, Hancock, Humphrey [1993] oraz Akhainen, Swamy, Taubman, Singamsetti [1997].

Analiza efektywności kosztowej umożliwia stwierdzenie, czy przy danych cenach czynników produkcji (ewentualnie także danych nakładach czynników stałych) firma ponosi minimalny koszt (całkowity lub zmienny) wytworzenia określonego poziomu produkcji lub wektora produktów. Innymi słowy, bada się czy dana wielkość produkcji została osiągnięta przy odpowiedniej skali i strukturze nakładów tych czynników. Jeżeli ten warunek jest spełniony, to firma wytwarzając optymalną wielkość produkcji (gwarantującą maksymalny zysk) osiąga ją przy odpowiednich nakładach czynników produkcji (po najniższym koszcie). Jeżeli przedsiębiorstwo ponosi większy koszt niż wynika on z mikroekonomicznej (granicznej) funkcji kosztu, to spowodowane jest to nieefektywnością kosztową - alokacyjną lub techniczną.

Termin „efektywność techniczna” wprowadzony przez Farella [1957] może być interpretowany na dwa różne sposoby. Przedsiębiorstwo charakteryzuje się efektywnością techniczną zorientowaną na efekty (ang. *output - oriented technical efficiency*), gdy z danych nakładów czynników produkcji uzyskuje produkcję niższą niż wynika ona z mikroekonomicznej funkcji produkcji, czyli  $Q/a = f(x_1, \dots, x_H)$ , gdzie  $0 < a \leq 1$  to parametr mierzący stopień odchylenia *in minus* obserwowanego poziomu produkcji  $Q$  od funkcji granicznej (por. np. Atkinson i Cornwell [1994]). Przedsiębiorstwo charakteryzuje się nieefektywnością techniczną zorientowaną na nakłady (ang. *input - oriented technical efficiency*), gdy określoną wielkość produkcji  $Q$  mogłoby uzyskać redukując proporcjonalnie zaangażowanie wszystkich czynników produkcji (obniżając koszt poniesionych nakładów), czyli  $Q = f(b \cdot x_1, \dots, b \cdot x_H)$ , gdzie  $0 < b \leq 1$  jest parametrem możliwych redukcji nakładów. W przypadku, gdy funkcja produkcji jest jednorodna stopnia  $\nu$  ( $\nu > 0$ ) zachodzi związek  $a = b^\nu$ , a zatem w przypadku szczególnym liniowej jednorodności oba sposoby definiowania nieefektywności technicznej są równoważne (por. Färe i Lovell [1978]). Jeżeli funkcja produkcji nie jest jednorodna rozróżnienie obu przypadków ma istotne znaczenie.

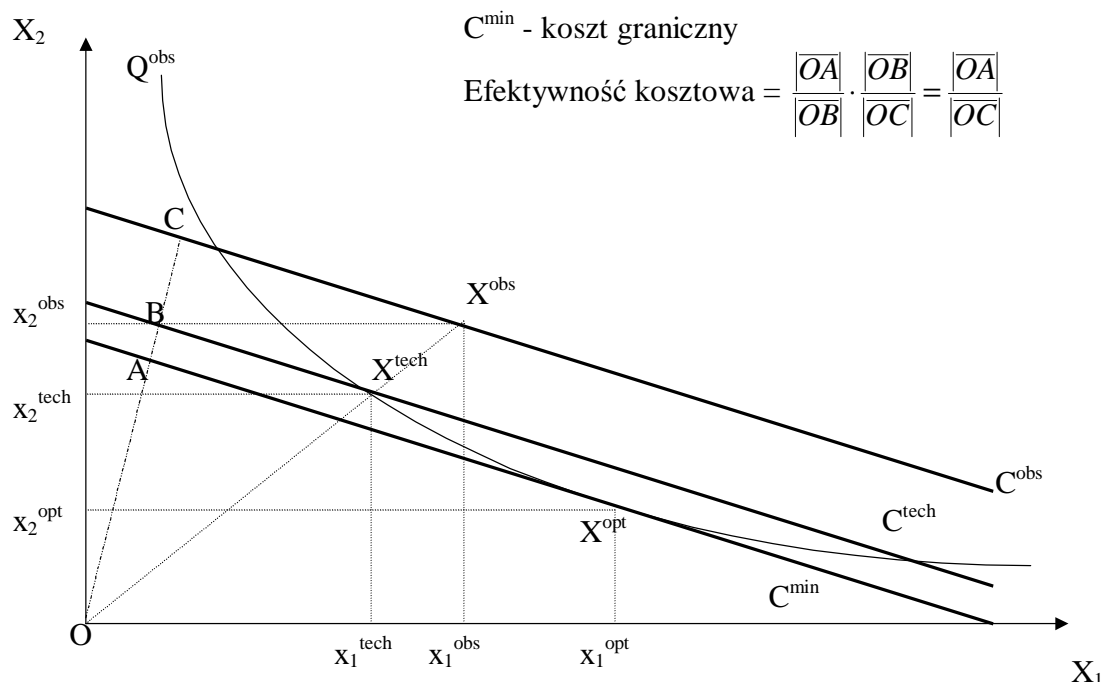
Nieefektywność alokacyjna polega na tym, że czynniki produkcji są zaangażowane w nieodpowiednich proporcjach. Ma to miejsce, gdy choć jedna z krańcowych stóp substytucji między czynnikami nie odpowiada ilorazowi ich cen rynkowych (równość ta wynika bezpośrednio z warunku koniecznego na istnienie ekstremum funkcji – zerowanie się pierwszych pochodnych).

Można pokazać, że odchylenie rzeczywiście poniesionego kosztu od granicznej funkcji kosztu jest rezultatem błędu alokacji lub nieefektywności technicznej (zob. Marzec i Osiewalski [1996-97]). Rysunek 1 przedstawia graficzną dekompozycję efektywności



kosztowej na techniczną (zorientowaną na nakłady) i alokacyjną (poczynione założenia: jeden produkt, dwa czynniki produkcji, ich ceny są na ustalonym poziomie). Przez  $X^{opt} = (x_1^{opt}, x_2^{opt})$  oznaczono optymalne nakłady obu czynników produkcji, które przy danych cenach czynników gwarantują uzyskanie produkcji na poziomie  $Q^{obs}$  po najniższym koszcie  $C^{min}$ . Natomiast  $X^{obs} = (x_1^{obs}, x_2^{obs})$  oznacza obserwowane nakłady, przy których uzyskano produkcję  $Q^{obs}$ , ponosząc koszt  $C^{obs}$ . Niech izokwanta  $\{(x_1, x_2): f(x_1, x_2) = Q^{obs}, f - \text{funkcja produkcji}\}$  jest zbiorem takich kombinacji nakładów, które są technicznie niezbędne do uzyskania produkcji  $Q^{obs}$ . Widoczna na rysunku nieefektywność techniczna polega na tym, że  $X^{obs}$  leży powyżej izokwenty  $Q^{obs}$ . Można więc tę samą produkcję  $Q^{obs}$  uzyskać przez proporcjonalną (tj. nie zmieniającą struktury) redukcję nakładów do punktu  $X^{tech}$  i obniżenie kosztu do  $C^{tech}$ . Różnica między  $C^{obs}$  i  $C^{tech}$  jest kosztem nieefektywności technicznej. Nieefektywność alokacyjna polega na tym, że struktura nakładów  $X^{obs}$  i  $X^{tech}$  nie odpowiada relacji cen czynników produkcji; produkcję  $Q^{obs}$  można uzyskać taniej (po koszcie  $C^{min} < C^{tech}$ ) przez zastosowanie optymalnej kombinacji  $X^{opt}$ . Podsumowując, miarą efektywności alokacyjnej jest więc stosunek długości odcinków  $OA/OB$ , a technicznej  $OB/OC$ . Efektywność kosztowa to  $OA/OC = (OA/OB) \cdot (OB/OC)$ , czyli iloczyn obu mierników cząstkowych; por. Kopp i Diewert [1982]. Na przykład: efektywność techniczna rzędu 0.8 i efektywność alokacyjna rzędu 0.9 prowadzi do efektywności kosztowej 0.72 (koszt niezbędny stanowi 72% kosztu poniesionego).

Rysunek 1. Nieefektywność kosztowa i jej dekompozycja (założenie: jeden produkt, dwa czynniki produkcji, a ich ceny są na ustalonym poziomie).



W dalszej części zostanie przedstawiony sposób modelowania nieefektywności kosztowej za pomocą stochastycznych modeli granicznych, które stanowią alternatywne podejście w stosunku do historycznie wcześniejszych metod opartych o techniki programowania matematycznego.

## 1.2. Modelowanie nieefektywności kosztowej.

Empiryczne badania efektywności ekonomicznej firm prowadzi się w oparciu o stochastyczne modele ekonometryczne, bądź o modele deterministyczne. Wykorzystanie tych ostatnich modeli, wykorzystujących technikę programowania liniowego<sup>6</sup>, w odniesieniu do instytucji finansowych (banków komercyjnych), prezentowane jest m.in. w pracach następujących autorów: Berg, Forsund, Hjalmarsson i Souminen [1993], Ferrier i Lovell [1990], Grabowski, Ragan i Rezvanian [1993], Hassan, Grabowski, Pasurka i Ragan [1990], Rangan, Grabowski, Aly i Pasurka [1988], Sherman i Gold [1985] oraz Rogowski [1998b,

<sup>6</sup> podstawową metodą jest tzw. DEA (*Data Envelopment Analysis*).

1998c] i Kopczewski [1999]. W obu podejściach model formułuje się tak, aby w sposób uproszczony, lecz właściwy ujmował podstawowe założenia związane z mikroekonomicznym opisem procesu produkcyjnego firmy. Wykorzystuje się w nich dane pochodzące z firm należących do jednej branży - więc stosujących tę samą technologię lub przynajmniej mających swobodny do niej dostęp - przy założeniu, że firmy mogą nabywać czynniki produkcji po egzogenicznie danych cenach. Założenie to pozwala przyjąć dla każdej z firm identyczną postać mikroekonomicznej funkcji produkcji, a w konsekwencji także funkcji kosztów. W przypadku danych przekrojowo-czasowych, jeżeli istnieją przesłanki merytoryczne, to dla prostoty modelu zakłada się, że technologia produkcji firm nie ulega istotnej zmianie w badanym okresie czasu.

Koszt graniczny przy swobodnym dostępie wszystkich firm do tej samej technologii reprezentuje minimalny koszt ponoszony przy danych cenach czynników i danym (zrealizowanym) poziomie produkcji. Odchylenie *in plus* od granicznej funkcji kosztów interpretowane może być jako błąd pomiaru, który ze swej istoty może być dowolnego znaku (jest zmienną symetryczną względem zera) lub nieefektywność, która jest zmienną nieujemną.

Według metodologii ekonometrycznej, problem efektywności (technicznej lub kosztowej) formułuje się zwykle za pomocą modelu jednorównaniowego składającego się z odpowiednio wyspecyfikowanej mikroekonomicznej funkcji produkcji lub kosztów (dla logarytmów tych zmiennych) oraz dwóch składników losowych, z których jeden (symetryczny względem zera) odzwierciedla efekt czynników przypadkowych i błędów pomiaru, zaś drugi (asymetryczny i stałego znaku) modeluje potencjalną nieefektywność. Występowanie złożonego składnika losowego definiuje modele, które w literaturze określa się jako tzw. stochastyczne modele graniczne (ang. *stochastic frontier models*). Zostały one zaproponowane przez dwa niezależne zespoły badawcze: Aigner, Lovell i Schmidt [1977] oraz Meeusen i van den Broeck [1977]. Dalszego rozwoju tej metodologii, prezentowanej głównie na łamach *Journal of Econometrics*, dokonali m. in. Stevenson [1980], Pitt i Lee [1981], Jondrow, Lovell, Materov i Schmidt [1982], Schmidt i Sickles [1984], Beckers i Hammond [1987], Greene [1980] i [1990], van den Broeck, Koop, Osiewalski i Steel [1994], Koop, Osiewalski i Steel [1994], [1997] oraz Fernández, Osiewalski i Steel [1997]. Wykorzystanie niektórych spośród tych modeli w badaniu efektywności kosztowej banków prezentowane jest w czołowej specjalistycznej literaturze, głównie na łamach *Journal of Banking and Finance*, *Journal of Money, Credit and Banking* oraz *Journal of Productivity*

*Analysis* (zob. Bauer, Hancock [1993], Cebenoyan, Cooperman, Register i Hudgins [1993], Ferrier i Lovell [1990], Kaparakis, Miller i Noulas [1994], Kraft, Tirtiroglu [1998] oraz Mester [1993]).

Natomiast w polskiej literaturze naukowej, zarówno statystyczno-ekonometrycznej jak i bankowej, niewiele pozycji poświęconych jest metodologicznym aspektom i badaniom empirycznym z zakresu efektywności technicznej czy kosztowej banków (i firm w ogóle). Najczęściej do porównywania banków i oceny ich działalności stosuje się tradycyjne metody taksonomiczne lub analizy wskaźnikowe (np. Grabczan [1995], Tarczyński [1998], Capiga, Kurzak [1998]). Dopiero ostatnio na łamach czasopism *Bank i Kredyt* oraz *Bank* ukazały się popularyzatorskie artykuły przeglądowe G. Rogowskiego na temat analizy efektywności oddziałów banku (Rogowski [1996], [1998a], [1998b], [1998c]). Wskazał on na ogromną potrzebę pomiaru i porównań efektywności ekonomicznej oddziałów, co może „stanowić cenne uzupełnienie tradycyjnych metod analiz wskaźnikowych stosowanych przez banki”. Uważa on, że zastosowanie metod ekonometrycznych „umożliwi między innymi obiektywizację kryteriów oceny oddziałów” (Rogowski [1998a], s. 27).

W przypadku danych przekrojowo-czasowych najprostszym jednorównaniowym stochastycznym modelem granicznej krótkookresowej funkcji kosztu zmiennego firmy może być sformułowany następująco (dla logarytmu kosztu):

$$y_{it} = h(x_{it}, \beta) + v_{it} + z_{it}, \quad (1)$$

gdzie:

$y_{it}$  - obserwowany logarytm kosztu zmiennego ( $VC_{it}$ )  $i$ -tej firmy w okresie  $t$  ( $i=1, \dots, N$ ;  $t=1, \dots, T$ ),

$x_{it}$  - wektor-wiersz zmiennych egzogenicznych (będących funkcjami wielkości produktów, cen czynników zmiennych i nakładów czynników stałych),

$\beta$  - wektor  $K+1$  nieznanych parametrów,

$h(x_{it}, \beta)$  - odpowiednio wyspecyfikowana graniczna krótkookresowa funkcja kosztu zmiennego, liniowa względem parametrów.

W równaniu (1) zakłada się o  $z_{it}$  i  $v_{it}$ , że  $1^0$  są zmiennymi losowymi niezależnymi od siebie, zarówno po firmach jak i po czasie ( $z_{it} \perp v_{kt}$ ),  $2^0$  są niezależne od regresorów, przy czym:

- $z_{it}$  to zmienna o wartościach wyłącznie nieujemnych ( $z_{i,t} \geq 0$ ), która odzwierciedla nieefektywność  $i$ -tej firmy w okresie  $t$  ( $i=1, \dots, N$ ;  $t=1, \dots, T$ ).

- $v_{it}$  to składnik czysto losowy o rozkładzie symetrycznym, wyrażający wpływ czynników przypadkowych bądź błędów w pomiarze kosztów; o wszystkich zmiennych  $v_{it}$  ( $i=1, \dots, N$ ;  $t=1, \dots, T$ ) zakłada się, że posiadają niezależne ( $v_{it} \perp v_{kt}$ ), identyczne rozkłady o zerowej wartości oczekiwanej, np. rozkłady normalne, tj.  $v_{it} \sim iin(0, \sigma_v^2)$ .

Składnik  $z_{it}$  reprezentuje wzrost kosztu spowodowany nieefektywnością techniczną (zbyt duża skala nakładów) lub alokacyjną (proporcje nakładów niezgodne są z relacją cen rynkowych czynników).

Główny problem w ekonometrycznej, parametrycznej analizie efektywności kosztowej firm sprowadza się do estymacji nieznanymi parametrów stochastycznej granicznej funkcji kosztów, a następnie do pomiaru wskaźników nieefektywności (bądź efektywności). Problemy w przypadku modelu (1) polegają na znalezieniu takiej metody estymacji, która uwzględni informacje a priori o nieujemnym składniku modelującym nieefektywność oraz pozwoli na estymację (lub predykcję) składników  $z_{it}$ . Rodzaj posiadanych danych - przekrojowo-czasowe (panelowe) czy tylko przekrojowe - i specyfikacja zmiennej reprezentującej nieefektywności są zatem podstawowymi czynnikami, które winny decydować o wyborze typu modelu i sposobu jego estymacji (por. np. Fernández, Osiewalski i Steel [1997]). O  $z_{it}$  zakłada się m.in., że posiadają niezależne rozkłady ( $z_{it} \perp z_{kt}$ ), np. wykładnicze indeksowane przez parametr  $\lambda_{it}$  ( $z_{it} \sim \text{EXP}(\lambda_{it})$ ). Rozważa się także przypadek identycznych rozkładów dla  $z_{it}$  po obu indeksach ( $z_{it} \sim \text{EXP}(\lambda)$ ), lub tylko po czasie ( $z_{it} \sim \text{EXP}(\lambda_i)$ ). Częstym założeniem (przyjmowanym zwłaszcza przy małym  $T$ ) jest nie tylko stałość technologii, ale i stałość efektywności firmy w czasie  $z_{it} = z_i$  dla  $t=1, \dots, T$ . Umożliwia to bardziej precyzyjny szacunek przeciętnej efektywności każdej firmy, gdyż istnieje dla niej  $T$  obserwacji, a nie tylko jedna. W tym przypadku zakłada się niezależność efektywności po firmach, czyli ( $z_i \perp z_k$ ) i przyjmuje się np.  $z_i \sim \text{EXP}(\lambda)$  lub  $z_i \sim \text{EXP}(\lambda_i)$ . Modele bardziej rozbudowane pozwalają na uwzględnienie i testowanie systematycznych różnic w efektywności, spowodowanych przez czynniki zewnętrzne. Zakłada się w nich m.in., że składniki reprezentujące nieefektywność nie mają identycznych rozkładów dla różnych firm, tzn. mają ten sam typ rozkładu, lecz o różnych parametrach będących funkcją pewnych zmiennych egzogenicznych (por. Kumbhakar, Ghosh i McGuckin [1991] oraz Koop, Osiewalski i Steel [199]).

Najważniejszą cechą odróżniającą tzw. modele deterministyczne od modeli stochastycznych jest brak składnika czysto losowego  $v_{it}$ . Modele deterministyczne zakładają istnienie jednego składnika  $z_{it}$ , więc dopuszczają tylko odchylenia *in plus* od granicznej

funkcji kosztu. To jest jedna z głównych przyczyn ostrej krytyki tego podejścia, nie uwzględniającego jakichkolwiek zaburzeń losowych oraz błędów pomiaru, które potencjalnie w różnym kierunku i z różną siłą mogą wpływać na poziom nieefektywności badanych firm. W modelu deterministycznym jedna nietypowa obserwacja może mieć znaczący wpływ na otrzymane wyniki - pomiar efektywności. Jeżeli natomiast w równaniu (1) zachowa się składnik czysto losowy  $v_{it}$  a pominię się składnik  $z_{it}$ , to przedstawia ono model przeciętnej funkcji kosztów (*average practice model*), formalnie odpowiadający założeniu pełnej efektywności. Należy jednak zwrócić uwagę, że tak wyspecyfikowane równanie (1) nie pozwala na ujęcie dekompozycji efektywności kosztowej na efektywność kosztową i alokacyjną. Próbę dekompozycji źródeł nieefektywności podjęli się Atkinson i Cornwell [1994a], [1994b] oraz Kumbhakar [1997] w tzw. modelu *shadow cost function*, będącym układem równań: odpowiednio zmodyfikowanej funkcji kosztów i równań warunkowego popytu na czynniki. Estymacja takiego modelu w przypadku translogarytmicznej funkcji kosztu (z uwzględnieniem nieliniowych restrykcji) z uwagi na stopień komplikacji może stanowić obszar dalszych pogłębionych badań.

Obecnie w badaniach empirycznych z tego zakresu najczęściej używaną formą funkcyjną jest translogarytmiczna funkcja kosztów (ang. *translog cost function*) określana również terminem *giętka forma funkcyjna*. Mimo, że do opisu procesu produkcji została ona użyta po raz pierwszy w 1971 roku przez Christensena, Jorgensona i Lau (zob. także Christensen, Jorgensona i Lau [1973]), to w polskiej literaturze ekonometrycznej praktycznie nie jest stosowana. Otrzymuje się ją - dla logarytmu kosztów ( $\ln VC$ ) - poprzez aproksymację drugiego rzędu dowolnej funkcji przynajmniej trzykrotnie różniczkowalnej (drugie pochodne ciągłe) w otoczeniu pewnego (arbitralnie dobranego) punktu. Lepszą aproksymację nieznanej funkcji kosztu (aczkolwiek nieznaczną przy dodatkowych założeniach o różniczkowalności) można uzyskać rozwijając w szereg Taylora, ucinając go po odpowiednich pochodnych rzędu wyższego niż drugi. Funkcja translogarytmiczna kosztu jest autonomiczna w stosunku do funkcji produkcji i jak dotąd nie pokazano z jakiej postaci funkcji produkcji otrzymuje się ją w wyniku rozwiązania problemu minimalizacji kosztu (o funkcji produkcji wiadomo, że nie jest funkcją jednorodną). Ponieważ translogarytmiczna funkcja kosztu nie jest funkcją jednorodną, więc jest wrażliwa na skalowanie danych. Zatem konieczne jest zwrócenie uwagi na fakt, że w równaniu (1) na potencjalną nieefektywność kosztową składa się nieefektywność techniczna zdefiniowana jako zorientowana na nakłady, a nie zorientowana na efekty. Rozróżnienie dwóch interpretacji nieefektywności technicznej jest w tym przypadku istotne.

Translogarytmiczna funkcja kosztów w przypadku krótkookresowym ma następująca postać:

$$\begin{aligned}
 \ln VC_{it} = & \beta_0^* + \sum_{g=1}^G \beta_g^{*(1)} \cdot \ln Q_{it,g} + \sum_{p=1}^P \beta_p^{*(2)} \cdot \ln \tilde{X}_{it,p} + \sum_{h=1}^H \beta_h^{*(3)} \cdot \ln w_{it,h} + \\
 & + \sum_{g=1}^G \sum_{h=1}^H \beta_{g,h}^{(4)} \cdot \ln Q_{it,g} \cdot \ln w_{it,h} + \sum_{p=1}^P \sum_{h=1}^H \beta_{p,h}^{(5)} \cdot \ln \tilde{X}_{it,p} \cdot \ln w_{it,h} + \\
 & + \sum_{p=1}^P \sum_{g=1}^G \beta_{p,g}^{(6)} \cdot \ln \tilde{X}_{it,p} \cdot \ln Q_{it,g} + \sum_{g=1}^G \sum_{j \geq g}^G \beta_{g,j}^{(7)} \cdot \ln Q_{it,g} \cdot \ln Q_{it,j} + \\
 & + \sum_{p=1}^P \sum_{j \geq p}^P \beta_{p,j}^{(8)} \cdot \ln \tilde{X}_{it,p} \cdot \ln \tilde{X}_{it,j} + \sum_{h=1}^H \sum_{j \geq h}^H \beta_{h,j}^{(9)} \cdot \ln w_{it,h} \cdot \ln w_{it,j} + v_{it} + z_{it}
 \end{aligned} \tag{2}$$

gdzie o  $v_{it}$  i  $z_{it}$  czyni się założenia identyczne jak w równaniu (1) i wykorzystuje się twierdzenie Younga<sup>7</sup> celu oszczędnej parametryzacji. Ponadto:

$VC_{it}$  - obserwowany koszt zmienny  $i$ -tej firmy w okresie  $t$  ( $i=1, \dots, N$ ;  $t=1, \dots, T$ ),

$Q_{it,g}$  - produkcja  $g$ -tego produktu ( $g=1, \dots, G$ ),

$w_{it,h}$  - cena  $h$ -tego czynnika zmiennego produkcji ( $h=1, \dots, H$ ),

$\tilde{X}_{it,p}$  - zaangażowanie  $p$ -tego czynnika stałego ( $p=1, \dots, P$ ).

Wykorzystywany często model kosztów Cobba-Douglasa stanowi formalnie szczególny przypadek funkcji translogarytmicznej (przy odpowiednich restrykcjach na parametry, których istotność można testować). W równaniu (2) występuje duża liczba parametrów, lecz można dodatkowo ograniczyć liczbę swobodnych parametrów, jeżeli uwzględnimy własność jednorodności funkcji kosztów stopnia jeden ze względu na ceny zmiennych czynników produkcji. Powoduje ona następujące restrykcje na parametry równania (2):

$$\begin{aligned}
 \sum_{h=1}^H \beta_h^{*(3)} &= 1 \\
 \sum_{h=1}^H \beta_{g,h}^{(4)} + \sum_{h=1}^H \beta_{p,h}^{(5)} &= 0 \quad \text{dla } p=1, \dots, P; g=1, \dots, G \\
 \beta_{h,h}^{(9)} + \frac{1}{2} \sum_{j>h}^H \beta_{h,j}^{(9)} + \frac{1}{2} \sum_{j=1}^{h-1} \beta_{j,h}^{(9)} &= 0 \quad \text{dla } h=1, \dots, H
 \end{aligned} \tag{3}$$

Warunek jednorodności funkcji kosztu względem cen będzie automatycznie spełniony, gdy koszt zmienny ( $VC_{it}$ ) i ceny  $w_{it,h}$  ( $h=1, \dots, H$ ) wydzieli się przez jedną z nich (np.  $w_{it,H}$ ):

<sup>7</sup> wynika z niego w szczególności, że jeżeli funkcja jest 2-krotnie różniczkowalna w pewnym punkcie i pochodne 2-go rzędu są ciągłe, to macierz pochodnych 2-go rzędu w tym punkcie jest macierzą symetryczną.

$$\begin{aligned}
\ln\left(\frac{VC_{it}}{w_{it,H}}\right) &= \beta_0^* + \sum_{g=1}^G \beta_g^{*(1)} \cdot \ln Q_{it,g} + \sum_{p=1}^P \beta_p^{*(2)} \cdot \ln \tilde{X}_{it,p} + \sum_{h=1}^H \beta_h^{*(3)} \cdot \ln\left(\frac{w_{it,h}}{w_{it,H}}\right) \\
&+ \sum_{g=1}^G \sum_{h=1}^H \beta_{g,h}^{(4)} \cdot \ln Q_{it,g} \cdot \ln\left(\frac{w_{it,h}}{w_{it,H}}\right) + \sum_{p=1}^P \sum_{h=1}^H \beta_{p,h}^{(5)} \cdot \ln \tilde{X}_{it,p} \cdot \ln\left(\frac{w_{it,h}}{w_{it,H}}\right) \\
&+ \sum_{p=1}^P \sum_{g=1}^G \beta_{p,g}^{(6)} \cdot \ln \tilde{X}_{it,p} \cdot \ln Q_{i,g} + \sum_{g=1}^G \sum_{j \geq g}^G \beta_{g,j}^{(7)} \cdot \ln Q_{it,g} \cdot \ln Q_{it,j} \\
&+ \sum_{p=1}^P \sum_{j \geq p}^P \beta_{p,j}^{(8)} \cdot \ln \tilde{X}_{it,p} \cdot \ln \tilde{X}_{it,j} + \sum_{h=1}^H \sum_{j \geq h}^H \beta_{h,j}^{(9)} \cdot \ln\left(\frac{w_{it,h}}{w_{it,H}}\right) \cdot \ln\left(\frac{w_{it,j}}{w_{it,H}}\right) + v_{it} + z_{it},
\end{aligned} \tag{4}$$

Po dokonaniu prostych operacji matematycznych model (2) przyjmie postać równoważną:

$$\begin{aligned}
\ln VC_{it} &= \beta_0 + \sum_{g=1}^G \beta_g^{(1)} \cdot \ln Q_{it,g} + \sum_{p=1}^P \beta_p^{(2)} \cdot \ln \tilde{X}_{it,p} + \sum_{h=1}^{H-1} \beta_h^{(3)} \cdot \ln w_{it,h} \\
&+ \sum_{g=1}^G \sum_{h=1}^{H-1} \beta_{g,h}^{(4)} \cdot \ln Q_{it,g} \cdot \ln w_{it,h} + \sum_{p=1}^P \sum_{h=1}^{H-1} \beta_{p,h}^{(5)} \cdot \ln \tilde{X}_{it,p} \cdot \ln w_{it,h} \\
&+ \sum_{p=1}^P \sum_{g=1}^G \beta_{p,g}^{(6)} \cdot \ln \tilde{X}_{it,p} \cdot \ln Q_{i,g} + \sum_{g=1}^G \sum_{j \geq g}^G \beta_{g,j}^{(7)} \cdot \ln Q_{it,g} \cdot \ln Q_{it,j} \\
&+ \sum_{p=1}^P \sum_{j \geq p}^P \beta_{p,j}^{(8)} \cdot \ln \tilde{X}_{it,p} \cdot \ln \tilde{X}_{it,j} + \sum_{h=1}^{H-1} \sum_{j \geq h}^{H-1} \beta_{h,j}^{(9)} \cdot \ln w_{it,h} \cdot \ln w_{it,j} + v_{it} + z_{it},
\end{aligned} \tag{5}$$

gdzie nowe parametry  $\beta_0$ ,  $\beta_g^{(1)}$ ,  $\beta_p^{(2)}$ ,  $\beta_h^{(3)}$  są funkcjami parametrów modelu (4) i ceny  $w_{it,H}$ , w szczególności:

$$\beta_0 = \beta_0^* - \ln w_{it,H} \cdot \left( 1 + \sum_{h=1}^{H-1} \beta_h^{*(3)} - \ln w_{it,H} \cdot \sum_{h=1}^{H-1} \sum_{j \geq h}^{H-1} \beta_{h,j}^{(9)} \right).$$

W efekcie liczba estymowanych parametrów funkcji kosztu w równaniu (4) - z wyjątkiem wyrazu wolnego - wynosi  $K=(G+H+P-1)(G+H+P+2)/2$  (np. w przypadku, gdy  $G=H=3$  i  $P=1$ , to  $K=27$ ).

W dalszej części zostaną opisane podstawowe charakterystyki służące do opisu procesu produkcji i kosztu, m.in. poprzez pomiar efektu skali i korzyści produkcji w przypadku translogarytmicznej, krótkookresowej funkcji kosztu zmiennego.



### 1.3. Charakterystyki procesu produkcji.

Oszacowana mikroekonomiczna funkcja kosztów jest cennym źródłem informacji o badanym procesie produkcji i koszcie produkcji, zwłaszcza, gdy rozważa się wiele produktów ( $G > 1$ ). W tym przypadku funkcja transformacji jest bardzo rzadko wykorzystywana w badaniach empirycznych, ponieważ występują problemy z jej modelowaniem i estymacją, chociaż ostatnio próby jej estymacji dokonali Fernández, Koop i Steel [2000]. Zatem w praktyce w przypadku wielu produktów wykorzystuje się funkcję kosztu, co ma teoretyczne uzasadnienie, ponieważ w myśl zasady dualizmu funkcja kosztu sumaryzuje wszystkie ekonomicznie istotne aspekty technologii (procesu produkcji). Istnieje jednoznaczna relacja wiążąca obie funkcje; zob. Shephard [1981].

Oryginalne parametry funkcji translogarytmicznej nie mają ekonomicznej interpretacji, lecz koszt krańcowy, elastyczności względem cen czynników zmiennych i nakładów czynników stałych oraz współczynnik efektu skali produkcji niosą informacje o kształtowaniu się kosztu i własnościach procesu produkcji. Z uwagi na fakt, że funkcja translogarytmiczna jest określona dla logarytmów zmiennych, elastyczności liczy się łatwo jako pochodne logarytmu kosztu względem logarytmu zmiennej. W efekcie elastyczność kosztu zmiennego  $VC(W, Q, \tilde{X})$  względem  $h$ -tej ceny ( $h=1, \dots, H$ ) dla translogarytmicznej funkcji danej równaniem (5) wyraża się wzorem:

$$\begin{aligned} \eta(VC/w_h)_{it} = \frac{\partial \ln VC(w_{it}, Q_{it}, \tilde{X}_{it})}{\partial \ln w_{it,h}} = & \beta_h^{(3)} + \sum_{g=1}^G \beta_{g,h}^{(4)} \cdot \ln Q_{it,g} + \sum_{p=1}^P \beta_{p,h}^{(5)} \cdot \ln \tilde{X}_{it,p} + \\ & + 2\beta_{h,h}^{(9)} \cdot \ln w_{it,h} + \sum_{j \neq h}^{H-1} \beta_{h,j}^{(9)} \cdot \ln w_{it,j}. \end{aligned} \quad (6)$$

Natomiast elastyczność kosztu względem ceny  $H$ -tego czynnika produkcji wyznacza się korzystając z własności jednorodności funkcji kosztów względem cen, a mianowicie:

$$\eta(VC/w_H)_{it} = \frac{\partial \ln VC(w_{it}, Q_{it}, \tilde{X}_{it})}{\partial \ln w_{it,H}} = 1 - \sum_{h=1}^{H-1} \eta(VC/w_{it,h}). \quad (7)$$

Elastyczność względem zaangażowania  $p$ -tego czynnika stałego ( $p=1, \dots, P$ ) wynosi:

$$\eta(VC / \tilde{X}_p)_{it} = \frac{\partial \ln VC(w_{it}, Q_{it}, \tilde{X}_{it})}{\partial \ln \tilde{X}_{it,p}} = \beta_p^{(2)} + \sum_{h=1}^{H-1} \beta_{p,h}^{(5)} \cdot \ln w_{it,h} + \sum_{g=1}^G \beta_{p,g}^{(6)} \cdot \ln Q_{it,g} + 2\beta_{p,p}^{(8)} \cdot \ln \tilde{X}_{it,p} + \sum_{j \neq p}^G \beta_{p,j}^{(8)} \cdot \ln \tilde{X}_{it,p}. \quad (8)$$

Elastyczność względem  $g$ -tego produktu ( $g=1, \dots, G$ ) wynosi:

$$\eta(VC / Q_g)_{it} = \frac{\partial \ln VC(w_{it}, Q_{it}, \tilde{X}_{it})}{\partial \ln Q_{it,g}} = \beta_g^{(1)} + \sum_{h=1}^{H-1} \beta_{g,h}^{(4)} \cdot \ln w_{it,h} + \sum_{p=1}^P \beta_{p,g}^{(6)} \cdot \ln \tilde{X}_{it,p} + 2\beta_{g,g}^{(7)} \cdot \ln Q_{it,g} + \sum_{j \neq g}^G \beta_{g,j}^{(7)} \cdot \ln Q_{it,j}. \quad (9)$$

Należy zauważyć, że w przypadku translogarytmicznej funkcji kosztu technologia charakteryzuje się zmiennymi elastycznościami, które zależą od cen czynników produkcji, wielkości produktów i nakładów czynników stałych. W przypadku danych przekrojowo-czasowych liczba wszystkich elastyczności wynosi  $N \cdot T \cdot (H+P+G)$ , lecz w praktyce wyznacza się je dla przeciętnej firmy, np. charakteryzującej się średnimi z próby (po  $i$  oraz  $t$ ) wielkościami produktów, cenami czynników zmiennych i nakładami czynników stałych. Zatem w dalszej części rozdziału dla prostoty zapisu oba indeksy zostaną pominięte. Elastyczność względem jednej zmiennej przy ustalonych wartościach pozostałych zmiennych ma postać funkcji logarytmicznej, której kształt zależy od znaku i wartości parametru występującego przy kwadracie logarytmu tej zmiennej w funkcji kosztu.

W przypadku rozważania wielu produktów analiza kosztu jest dogodnym narzędziem poznania charakteru procesu produkcji poprzez określenie ekonomii skali i korzyści zakresu produkcji. W przypadku jednego produktu ( $G=1$ ) ekonomia skali lub efekt skali produkcji (ang. *scale economies*) mierzony jest poprzez współczynnik efekt skali będący odwrotnością elastyczności kosztu względem produkcji (ang. *returns-to-scale; RTS*), czyli

$$RTS = \left( \frac{\partial \ln VC(Q, W, \tilde{X})}{\partial \ln Q} \right)^{-1} = \left( \frac{\partial VC(Q, W, \tilde{X}) / \partial Q}{VC(Q, W, \tilde{X}) / Q} \right)^{-1} \approx \left( \frac{\Delta VC / VC}{\Delta Q / Q} \right)^{-1}. \quad (10)$$

W przypadku produkcji wieloproduktowej pomiar efektu skali produkcji jest bardziej skomplikowany. Panzar i Willig [1977] zaproponowali zagregowany współczynnik efektu skali (ang. *overall or aggregate scale economies*; por. także Youn Kim [1986], Molyneux,

Altunbas i Gardener [1996]), który w przypadku funkcji krótkookresowej można zdefiniować następująco<sup>8</sup>:

$$SL(Q) = VC(Q, W, \tilde{X}) / \sum_{g=1}^G Q_g \frac{\partial VC(Q, W, \tilde{X})}{\partial Q_g} = \left( \sum_{g=1}^G \eta(VC / Q_g) \right)^{-1}. \quad (11)$$

Miara ta przedstawia zmiany skali produkcji spowodowane wzrostem kosztu zmiennego produkcji w wyniku proporcjonalnego wzrostu wszystkich nakładów czynników zmiennych produkcji. Malejący (rosnący) efekt skali względem wyłącznie czynników zmiennych występuje, gdy  $SL < 1$  ( $SL > 1$ ), czyli zmiana kosztu zmiennego o 1% spowodowana proporcjonalnym wzrostem nakładów zmiennych czynników produkcji powoduje wzrost produkcji każdego z produktów o mniej (więcej) niż 1%, wówczas przeciętny koszt zmienny jest rosnącą (malejącą) funkcją produkcji. Efekt stały ( $SL=1$ ) występuje, gdy wzrost kosztu zmiennego powoduje proporcjonalny wzrost produkcji (przeciętny koszt zmienny jest stały).

W oparciu o koncepcje wskaźnika  $SL$  skonstruowano współczynnik efektu skali określonego produktu (ang. *product-specific scale economic*; por. Baumol, Panzar i Willig [1982]):

$$SL_g(Q) = \left( \frac{\partial IVC(Q, W, \tilde{X})}{\partial Q_g} / \frac{IVC_g}{Q_g} \right)^{-1} = \frac{IVC_g(Q, W, \tilde{X})}{Q_g \cdot \frac{\partial VC(Q, W, \tilde{X})}{\partial Q_g}} = \frac{IVC_g / VC}{\eta(VC / Q_g)}, \quad (12)$$

gdzie  $IVC_g = VC(Q_1, \dots, Q_G, W, \tilde{X}) - VC(Q_1, \dots, Q_{g-1}, 0, Q_{g+1}, \dots, Q_G, W, \tilde{X})$  jest kosztem zmiennym wytworzenia jednego produktu na poziomie  $Q_g$ . Informuje on o ile procent wzrośnie produkcja  $g$ -tego produktu na wskutek jednoprocentowego wzrostu kosztu zmiennego. Koncepcję wskaźnika  $SL_g$  można rozszerzyć na grupę produktów (por. Youn Kim [1986]).

Inna propozycja pomiaru efektu skali produkcji jest oparta o tzw. radialną krzywą kosztu przeciętnego (*Ray Average Cost*; por. Baumol, Panzar i Willig [1982] oraz Kim 1986, Zardokoohi i Kolari [1994], Molyneux, Altunbas i Gardener [1996]), która – dla kosztu zmiennego - jest zdefiniowana jako:  $RAC(k) = VC(kQ^0, W, \tilde{X}) / k$ . Przez  $Q^0 = [Q_1^0, \dots, Q_G^0]$  oznaczono odpowiednio dobrany wektor produktów, zawierający np. średnie wartości

---

<sup>8</sup> Można zauważyć, że  $\Delta VC / VC \approx \sum_{g=1}^G (\partial \ln VC / \partial \ln Q_g \cdot \Delta Q_g / Q_g)$ , zatem jeżeli tylko

$$\Delta Q_g / Q_g = \Delta Q / Q \text{ dla każdego } g=1, \dots, G, \text{ to } \frac{\Delta Q / Q}{\Delta VC / VC} \approx \left( \sum_{g=1}^G \partial \ln VC / \ln Q_g \right)^{-1}.$$

poszczególnych produktów dla wszystkich banków w próbie lub dla grup banków wyodrębnionych np. ze względu na wielkość aktywów (por. Gilligan, Smirlock i Marshall [1984]).  $RAC(k)$  przedstawia zatem koszt produkcji jako funkcję  $k$  – współczynnika zmiany skali produkcji - przy ustalonej strukturze produkcji. Do określenia typu efektu skali produkcji proponuje się współczynnik radialnego efektu skali  $S(k)$  zdefiniowany jako:

$$S(k) = (1 + \eta(RAC/k))^{-1}, \quad (13)$$

gdzie  $\eta(RAC/k)$  to elastyczność radialnej krzywej przeciętnego kosztu zmiennego –  $RAC(k)$  - względem parametru  $k$ ; zob. Baumol, Panzar i Willig [1982]. Współczynnik radialnego efektu skali względem czynników zmiennych informuje, w przypadku banku wytwarzającego  $k$ -krotnie więcej każdego z produktów niż przeciętnie w grupie banków, o ile procent zmienia się wielkości każdego z produktów pod wpływem jednoprocentowego wzrostu kosztu zmiennego. Dla translogarytmicznej funkcji kosztu zmiennego ten współczynnik wynosi:

$$S(k) = \left( \frac{\partial \ln VC(k \cdot Q^0, W, X)}{\partial \ln k} \right)^{-1} = \quad (14)$$

$$\left( \sum_{g=1}^G \beta_g^{(1)} + \sum_{h=1}^{H-1} \left( \ln w_{it,h} \cdot \sum_{g=1}^G \beta_{g,h}^{(4)} \right) + \sum_{p=1}^{P-1} \left( \ln X_{it,p} \cdot \sum_{g=1}^G \beta_{p,g}^{(6)} \right) + \sum_{g=1}^G \left( \ln Q_g^0 \cdot \left( 2 \cdot \beta_{g,g}^{(7)} + \sum_{j \neq g}^G \beta_{g,j}^{(7)} \right) \right) + 2 \cdot \ln k \cdot \sum_{g=1}^G \left( \beta_{g,g}^{(7)} + \sum_{j \neq g}^G \beta_{g,j}^{(7)} \right) \right)^{-1}$$

W przypadku szczególnym, gdy  $k=1$  (lub dla jednorodnych funkcji kosztu, np. Cobba-Douglasa), miernik  $S(k)$  sprowadza się do zagregowanego współczynnika efektu skali danego wzorem (11). Uzasadnione jest stosowanie tej miary skali produkcji przede wszystkim dla niejednorodnych funkcji kosztu, aczkolwiek można mieć uwagi krytyczne pod kątem definicji tej miary korzyści skali produkcji, co potwierdziły badania empiryczne prezentowane w dalszej części pracy.

Badania empiryczne prowadzone zarówno w Stanach Zjednoczonych jak i w Europie Zachodniej potwierdzają tezę, że krzywa kosztów przeciętnych przedsiębiorstw sektora finansowego jest krzywą ukształtną (ma kształt litery U). W dobie konsolidacji banków czyni się zatem próby wyznaczenia optymalnej wielkości banku (ze względu na wielkość produkcji), tzn. takiej, aby przeciętne koszty produkcji były najmniejsze (np. Berger, Hanweck i Humphrey [1987], Ferrier i Lovell [1990], Berger i Humphrey [1991], Noulas, Subhash i Miller [1990], McAllister i McManus [1993]). W tym celu analizuje się działalność banków pod względem zakresu świadczonych usług i kosztów z nimi związanych. Fuzje banków wyspecjalizowanych jak dotąd w świadczeniu tylko określonych usług bankowych

(np. hipotecznych, inwestycyjnych), mają spowodować m.in. znaczące obniżenie się kosztów działalności powstających, silnych kapitałowo struktur holdingowych w postaci dużych banków uniwersalnych (komercyjnych). Badania tego typu prowadzi się w oparciu o analizę zakresu produkcji, której podstawy teoretyczne zostały sformułowane na początku lat osiemdziesiątych (zob. Panzar, Willig [1981] i Baumol, Panzar, Willig [1982]).

Analiza zakresu produkcji wielu produktów (ang. *economies of scope*) pozwala na zbadanie czy istnieją możliwości obniżenia kosztów działalności banku poprzez odstąpienie od dotychczasowej specjalizacji banku i poszerzenie zakresu świadczonych usług (produktów). Efekt (korzyść) pełnego zakresu produkcji jest związany z subaddytywnością kosztów i ma miejsce, gdy koszt produkcji wektora produktów  $Q_1, \dots, Q_G$  w przypadku jednego banku jest niższy od sumy kosztu produkcji tychże produktów wytworzonych przez wyspecjalizowane banki. Zatem technologia produkcji wykazuje korzyści pełnego zakresu produkcji, jeżeli  $VC(Q_1, \dots, Q_G) < VC(Q_1, 0, \dots, 0) + VC(0, Q_2, 0, \dots, 0) + \dots + VC(0, \dots, 0, Q_G)$ .

Przyczynami występowania korzyści zakresu produkcji są m.in. redukcja kosztów obsługi klienta wskutek łącznej sprzedaży produktów, lepsze wykorzystanie majątku trwałego (koszty stałe przypadają na większą liczbę produktów), obniżka kosztów reklamy produktów oraz obniżenie kosztów ponoszonego ryzyka, które ulega zmniejszeniu wskutek większej dywersyfikacji aktywów bankowych.

Jeżeli nie występują korzyści pełnego zakresu produkcji, to zazwyczaj ma miejsce efekt specjalizacji; ponadto formalnie może występować sytuacja jednoczesnego braku korzyści zakresu produkcji i braku korzyści specjalizacji. Baumol, Panzar i Willig [1982] wykazali, że dla wieloproduktowej funkcji kosztu warunkiem wystarczającym (lecz nie koniecznym) jest tzw. komplementarność funkcji kosztu między produktami. Powoduje ona, że koszt krańcowy względem ustalonego produktu (np. o numerze  $g$ ) jest malejącą funkcją wszystkich pozostałych produktów. Zwiększenie o małą jednostkę produkcji każdego z pozostałych produktów ( $Q_1, \dots, Q_{g-1}, Q_{g+1}, \dots, Q_G$ ) spowoduje spadek kosztu krańcowego względem tego produktu, zatem wytworzenie dodatkowej jednostki produktu w sytuacji łącznej produkcji nastąpi po koszcie niższym niż w przypadku, gdyby zwiększono produkcję tylko jednego produktu, a nie wszystkich jednocześnie. Formalnie warunek wystarczający na istnienie korzyści zakresu produkcji wymaga, aby:

$$\frac{\partial^2 VC(Q, W, \bar{X})}{\partial Q_g \partial Q_j} < 0 \quad \text{dla } g \neq j; g, j = 1, \dots, G, \quad (15)$$

Logiczną konsekwencją powyższego warunku jest stwierdzenie, że jeżeli występuje efekt specjalizacji, to  $\partial^2 VC(Q, W, \tilde{X}) / \partial Q_g \partial Q_j > 0$  dla przynajmniej jednej pary  $(g, j)$ ,  $g \neq j$ . W przypadku translogarytmicznej funkcji kosztu zmiennego warunek (15) oznacza, że

$$\frac{VC}{Q_j} \cdot \left[ \frac{\partial}{\partial Q_g} \left( \frac{\partial \ln VC}{\partial \ln Q_j} \right) + \frac{1}{Q_g} \cdot \frac{\partial \ln VC}{\partial \ln Q_g} \cdot \frac{\partial \ln VC}{\partial \ln Q_j} \right] < 0, \text{ czyli:} \quad (16)$$

$$\beta_{g,j}^{(7)} + \eta(VC/Q_{it,g}) \cdot \eta(VC/Q_{it,j}) < 0 \quad \text{dla } g \neq j. \quad (17)$$

Dla najprostszej funkcji kosztu, tj. potęgowej (Cobba i Douglasa) warunek (15) jest zawsze nie spełniony, zatem w badaniach korzyści pełnego zakresu produkcji należy wykorzystywać bardziej ogólne postacie funkcji.

W badaniach empirycznych zwykle testuje się warunek wystarczający na istnienie korzyści zakresu produkcji. Z uwagi na skomplikowaną postać warunku (17) dla translogarytmicznej funkcji kosztu stosuje się test przybliżony. Murray i White [1983], Lawrence [1989], Dietsch [1993] oraz Muldur i Sassenou [1993] przyjmują  $w_{it,h} = Q_{it,g} = \tilde{X}_{it,p} = 1$  i sprawdzają warunek:

$$\beta_{g,j}^{(7)} + \beta_g^{(2)} \cdot \beta_j^{(2)} < 0 \quad \text{dla } g \neq j. \quad (18)$$

Baumol, Panzar i Willig [1982] zaproponowali naturalną miarę korzyści z tytułu obniżki kosztu w wyniku łącznej produkcji, zwaną współczynnikiem korzyści zakresu produkcji (SC):

$$SC = \frac{\sum_{g=1}^G VC_g(Q_g, W, \tilde{X}) - VC(Q, W, \tilde{X})}{VC(Q, W, \tilde{X})}, \quad (19)$$

gdzie  $VC_g(Q_g, W, \tilde{X}) = VC(0, \dots, 0, Q_g, 0, \dots, 0, W, \tilde{X})$  jest kosztem zmiennym wyprodukowania  $g$ -tego produktu na poziomie  $Q_g$ . Współczynnik  $SC > 0$  mierzy procentową oszczędność kosztu (gdy  $SC > 0$ ) lub nadwyżkę kosztu (gdy  $SC < 0$ ) spowodowaną łączną produkcją wszystkich produktów zamiast specjalizacji. Miarę korzyści zakresu produkcji można uogólnić także na przypadek grupy produktów (por. Youn Kim [1986]).

Konstrukcja miernika  $SC$  jest krytykowana zwłaszcza w zastosowaniu do translogarytmicznej funkcji kosztu (por. Youn Kim [1986]), ponieważ wyznaczając koszt wytworzenia jednego produktu  $VC_g(Q_g, W, \tilde{X})$  przyjmuje się zerową wielkość produkcji

pozostałych, czego zwykle nie obserwuje się w praktyce. A zatem obliczenie  $SC$  wymagać może ekstrapolacji oszacowanej funkcji kosztu daleko poza obszar empirycznych wartości zmiennych. Ponadto w przypadku najczęściej wykorzystywanych funkcji kosztu, np. translogarytmicznej, które zwykle nie są określone dla zerowych wartości zmiennych, miernik  $SC$  jest nie zdefiniowany, więc jego zastosowanie w badaniach empirycznych jest wielce ograniczone. W praktyce - z uwagi na niezdefiniowaną wielkość logarytmu zera - przyjmuje się dla zmiennych reprezentujących produkty arbitralnie małe wartości, np. Youn Kim [1986] przyjmuje 0.1, natomiast Kim [1986] podstawia 0.001. Zaletą miernika  $SC$  jest to, że przyjmując wartości mniejsze (większe) od zera wskazuje na występowanie korzyści specjalizacji (pełnego zakresu) zwłaszcza, że warunek wystarczający na istnienie korzyści pełnego zakresu produkcji w badaniach empirycznych często jest nie spełniony (jest zbyt mocny).

Zardkoohi i Kolari [1994] zaproponowali modyfikację miary efektu (korzyści) zakresu produkcji w przypadku dwóch produktów (przy ustalonych cenach i ewentualnie czynnikach stałych):

$$SCP = \frac{\Delta C_1 + \Delta C_2 - \Delta C_{1,2}}{\Delta C_{1,2}}, \text{ gdzie:} \quad (20)$$

$$\Delta C_1 = C(Q_1^{min} + \Delta Q_1, Q_2^{min}) - C(Q_1^{min}, Q_2^{min})$$

$$\Delta C_2 = C(Q_1^{min}, Q_2^{min} + \Delta Q_2) - C(Q_1^{min}, Q_2^{min})$$

$$\Delta C_{1,2} = C(Q_1^{min} + \Delta Q_1, Q_2^{min} + \Delta Q_2) - C(Q_1^{min}, Q_2^{min})$$

Wielkości  $Q_1^{min}$  i  $Q_2^{min}$  oznaczają minimalne wielkości produktów zaobserwowane w próbie,  $\Delta Q_1$  i  $\Delta Q_2$  to różnice między średnią a minimalną wielkością produktów,  $\Delta C_{1,2}$  stanowi zatem różnicę między kosztem wytworzenia obu wielkości produktów na poziomie przeciętnym i minimalnym, natomiast  $\Delta C_1$  określa dodatkowy koszt jaki należy ponieść, aby zwiększyć produkcję pierwszego produktu z poziomu minimalnego do poziomu średniego w próbie przy ustalonej wielkości drugiego produktu na poziomie minimalnym. Miara ta informuje o możliwym stopniu redukcji kosztu ( $SCP > 0$ ) wskutek jednoczesnego zwiększenia łącznej produkcji obu produktów z poziomu minimalnego (obserwowanego w próbie) do poziomu średniego. Jedną z wad tego miernika jest to, że liczy się go dla hipotetycznego banku, oraz fakt, że rozważany wzrost produkcji z poziomu minimalnego do poziomu średniego w praktyce jest niemożliwy. Inną modyfikacją miernika korzyści zakresu

produkcji, także nie pozbawioną tych wad, jest propozycja Hughes i Mester [1993] i Mester [1993] tzw. *within-sample economies of scope*.

W badaniach empirycznych prezentowanych w dalszej części zostanie wykorzystana miara *SCP* liczona dla każdego oddziału przyjmując wartości pozostałych zmiennych, tj. cen czynników zmiennych i nakładów czynnika stałego, na poziomie obserwowanym. Należy zwrócić uwagę na arbitralność przyjętego miernika *SCP*, zatem uzyskane w ten sposób wyniki trzeba traktować jako wstępne. Istnieje potrzeba wprowadzenia innej miary potencjalnych korzyści zakresu produkcji banku, która jednocześnie uwzględni ograniczenia wynikające z przyjętej postaci translogarytmicznej funkcji kosztu i rzeczywisty, obserwowany poziom zmiennych, w tym wielkości produktów w przypadku analizowanego banku.

Z punktu widzenia mikroekonomicznej teorii opisu działalności banków komercyjnych oraz konkretnych badań empirycznych efektywności kosztów w bankach konieczne jest określenie czynników produkcji, produktów oraz pomiar tych wielkości i kosztu całkowitego. Specyficzne cechy procesu produkcji w bankach powodują, że pojawia się przy tym szereg trudnych pytań. Dotyczą one np. tego, czy depozyty - lokaty przyjmowane od różnych klientów banku - mają być traktowane jako produkty, czy jako czynniki produkcji. Czy produktem banku są szeroko pojęte usługi bankowe? W jaki sposób mierzyć produkcję - czy poprzez liczbę rachunków, czy liczbę transakcji bankowych, a może przez ich wartość? Analiza charakteru procesu produkcyjnego w kontekście kryteriów podejmowania decyzji ekonomicznych instytucji finansowych pozwoli na określenie czynników produkcji banku i jego produktów. W następnej części pracy zostanie przedstawiona koncepcja produktów, czynników produkcji i kosztu instytucji finansowych w świetle mikroekonomicznej teorii firm, zaproponowana przez Sealeya i Lindleya w 1977 na łamach *The Journal of Finance* w pracy „*Inputs, outputs, and a theory of production and cost at depository financial institutions*”.

W celu wykorzystania stochastycznego modelu granicznej funkcji kosztu do badania i pomiaru efektywności kosztowej oraz analizy procesu produkcji w bankach komercyjnych, konieczne jest określenie zmiennych występujących w równaniu krótkookresowej funkcji kosztu zmiennego danej wzorem (2). Analiza działalności banków w świetle mikroekonomicznej teorii firm umożliwi ich poprawną specyfikację.



## **2. Bank jako przedsiębiorstwo - ujęcie mikroekonomiczne.**

### **2.1. Model instytucji finansowych wg Sealeya i Lindleya.**

W literaturze bankowej dominują dwa różne podejścia do problemu ujęcia działalności (procesu produkcji) banku, które mają wpływ na sposób określenia czynników i produktów banku. Pierwsze podejście, tzw. teoria portfela, zakłada, że bank jest szczególnym przedsiębiorstwem, którego celem jest utrzymywanie równowagi między aktywami (kredytami) i pasywami (depozytami) uwzględniając płynność poszczególnym składników bilansu i dywersyfikację ryzyka. W tym podejściu depozyty i kredyty są traktowane jednakowo - jako produkty, a działalność banku jest rozważana z punktu widzenia działalności całego sektora finansowego. Banki (komercyjne) w ujęciu makroekonomicznym traktuje się jako instytucje, które dokonują transformacji informacji, wielkości sumy pieniędzy, ryzyka i terminu przy przyjmowaniu wkładów i udzielaniu kredytów oraz pełnią rolę pośrednika, płatnika, agenta, gwaranta i instrumentu realizacji polityki gospodarczej rządu (Jaworski [1996]; str. 9). Szczególną rolę w tym podejściu przywiązuje się do roli banku centralnego, który oddziałuje bezpośrednio na cały sektor bankowy poprzez operacje rynku otwartego, politykę redyskontową i procentową. Analizuje się bank z uwzględnieniem m.in. takich czynników zewnętrznych jak konieczność odprowadzania przez bank rezerw obowiązkowych od przyjętych środków (od wszystkich podmiotów gospodarczych i osób fizycznych z wyłączeniem innych banków) czy konieczność posiadania zdolności kredytowej (płynności). Spojrzenie na bank z punktu widzenia całej gospodarki nie umożliwia jednak dokonania analizy sprawności, jakości działania i stopnia realizacji jego celów. Ponadto podejście to całkowicie pomija rolę ograniczeń w procesie podejmowania decyzji instytucji finansowej, które związane są z technologią i kosztami własnymi produkcji.

Alternatywne podejście, tzw. *real approach*, rozważa działalność banku w świetle mikroekonomicznej teorii firm (produkcyjnych) i uwzględnia rolę czynników pominiętych w podejściu portfelowym. Podejście to traktuje elementy otoczenia banku, tj. działalność banku

centralnego i obowiązujące uregulowania prawno-organizacyjne, jako zewnętrzne warunki działania, które regulują działalność tych przedsiębiorstw.

Ujęcie działalności instytucji finansowych w świetle mikroekonomicznej teorii firm jest możliwe, jeżeli poprawnie określi się kluczowe elementy tej teorii, tzn. dokona się klasyfikacji czynników produkcji i produktów finansowych, rozważając kryteria podejmowania przez banki decyzji ekonomicznych. W początkowych pracach z zakresu mikroekonomicznego modelowania działalności banków jako produkty najczęściej wymieniane były: aktywa całkowite, aktywa generujące dochód (*earning assets*), depozyty całkowite, depozyty na żądanie, liczba rachunków depozytowych i kredytowych, przychody operacyjne brutto i inne kombinacje tych mierników (zob. Sealey i Lindley [1977]). Różnorodność wymienionych wyżej produktów wskazuje, że określenie dla nich jednolitych miar nie było proste.

Podstawy obecnie stosowanej mikroekonomicznej klasyfikacji czynników i produktów instytucji finansowych (w tym banków komercyjnych), wykorzystywanych m.in. w badaniach efektywności ekonomicznej, stworzyli Sealey i Lindley. Przeprowadzona przez nich analiza obydwu aspektów procesu produkcyjnego instytucji finansowych: technicznego i ekonomicznego w kontekście kryteriów podejmowania decyzji przyczyniła się do stworzenia modelu depozytowo-kredytowego opisującego działalność banku, a w konsekwencji do klasyfikacji czynników i produktów. W swym modelu przyjęli oni, że **głównym celem działalności instytucji finansowych jest maksymalizacja zysku**, a dywersyfikacja ryzyka i ograniczenia organizacyjno-prawne stanowią wspólne dla każdego z podmiotów warunki ograniczające zakres podejmowanych decyzji. Sealey i Lindley [1977] rozważali działalność instytucji finansowych **z punktu widzenia celów tych firm, a nie całej gospodarki**. Dla łatwiejszego zrozumienia źródeł dokonanej przez nich klasyfikacji pomocne jest przedstawienie podstawowych terminów z teorii produkcji i odniesienie ich do instytucji finansowych, ze zwróceniem szczególnej uwagi na rolę depozytów w procesie produkcji tych podmiotów.

Proces produkcyjny w sensie technicznym to (wg Frischa) proces transformacji kierowany przez człowieka; transformacja ta powoduje, że pewne dobra i/lub usługi wchodzi do procesu, w którym tracą swoją tożsamość, tzn. kończą istnienie w oryginalnej formie, podczas gdy tworzone są inne dobra lub usługi (Frisch [1965]). W przypadku instytucji finansowych proces transformacji polega na pozyskaniu pieniędzy od podmiotów mających nadwyżkę finansową i pożyczanie tych środków tym podmiotom gospodarczym, które mają

ich deficyt (niedobór). Szczególną rolę spełniają banki komercyjne, tzn. instytucje finansowe specjalizujące się w pozyskiwaniu depozytów i udzielaniu kredytów. Produktem w sensie technicznym jest dla nich wachlarz usług finansowych świadczonych deponentom i kredytobiorcom. Usługi te mogą być kategoryzowane jako:

- usługi pośrednictwa (*intermediation*); np. przyjmowanie lokat, udzielanie kredytów, gwarancji i poręczeń,
- administrowanie płatnościami klientów,
- inne usługi w postaci doradztwa finansowego i zarządzania finansami klientów instytucji finansowych.

Rozwój mikroekonomicznej teorii instytucji finansowych przyniósł zmianę sposobu opisu (dotychczas tylko technicznego) procesu produkcji. Kluczowym elementem nowej teorii było przyjęcie założenia, że proces produkcji ma dwa wymiary: jeden techniczny (proces transformacji), a drugi ekonomiczny. Przez proces produkcji w sensie ekonomicznym rozumie się próbę stworzenia produktu, który posiada wyższą wartość niż oryginalne czynniki produkcji (Frisch [1965]).

Przyjęcie rozszerzonej definicji procesu produkcji powoduje, iż za produkty instytucji finansowych rozumiane w sensie ekonomicznym można uważać tylko te „produkty techniczne”, które mają wyższą wartość niż zaangażowane nakłady czynników produkcji (gdzie ich wartość wyrażona jest w cenach rynkowych). Pozwala to na właściwe ujęcie maksymalizacji zysku jako podstawowego kryterium działania: instytucja finansowa opierając się na własnych kosztach i przychodach - mierzonych w sposób jawny przez rynek - podejmuje decyzje o strukturze i skali produkcji. W tym kontekście usługi świadczone deponentom przez instytucje finansowe należy skojarzyć z nabywaniem ekonomicznych czynników produkcji. Koszt tych usług w postaci np. udostępniania skrytek bankowych, obrotu czekowego i kart płatniczych stanowi częściową zapłatę za dysponowanie środkami deponentów, choć nie rodzi bezpośrednich lub znaczących przychodów. Instytucje finansowe co prawda pobierają opłaty za te usługi (np. za wydanie karty płatniczej lub prowadzenie rachunku oszczędnościowo-rozliczeniowego), jednak zwykle nie pokrywają one rzeczywiście poniesionych kosztów związanych z tymi czynnościami. W sytuacji silnej konkurencji banki niekiedy rezygnują z tych opłat celem pozyskania nowych i utrzymania dotychczasowych klientów. Koszty ponoszone w wyniku dostarczanych usług stanowią element całkowitych kosztów pozyskania środków pieniężnych. W konsekwencji koszt, który ponosi instytucja finansowa w celu pozyskania środków od deponentów składa się z niejawnych (ukrytych)

kosztów zaangażowania kapitału, pracy ludzkiej i kosztów materiałowych związanych z produkcją usług świadczonych deponentom oraz jawnych kosztów odsetkowych od zdeponowanych przez nich środków.

Z przeprowadzonej przez Sealeya i Lindleya analizy procesu produkcyjnego instytucji finansowych wynika, iż z punktu widzenia tych firm jest on procesem wieloetapowym wymagającym półproduktów (*intermediate outputs*). W procesie tym środki pozyskane bezpośrednio od deponentów lub pośrednio na rynku międzybankowym (a przeznaczone na kredyty) i usługi świadczone przez instytucje finansowe przy użyciu kapitału, pracy i materiałowych czynników produkcji są użyte do wytworzenia aktywów generujących przychód (ang. *earning assets*). Proces ten jest analogiczny do tego w wielowydziałowym przedsiębiorstwie produkcyjnym, w którym jeden wydział wytwarza półprodukty, bezpośrednio użyte jako czynnik produkcji w innym wydziale. Ostatecznie półprodukty powalają na uzyskanie finalnego produktu ekonomicznego: aktywów generujących przychód, tj. kredytów, udziałów w innych podmiotach i pozostałych papierów wartościowych (np. bony, weksle itp.); zob. Rysunek 2.

Właściwe - z punktu widzenia maksymalizacji zysku jako kryterium podejmowania decyzji instytucji finansowych - pojęcie produktu sprowadza się do usług świadczonych ich dłużnikom. Depozyty są traktowane jako finansowy czynnik produkcji, kapitał jako czynnik fizyczny a pracownicy jako czynnik ludzki (praca). Natomiast uzyskiwane produkty bankowe wyrażają się przez aktywa generujące przychód (*earning assets*), które proponuje się mierzyć poprzez wartość różnych typów tych aktywów wyrażoną w jednostkach pieniężnych, które są w tym przypadku jednostkami naturalnymi.

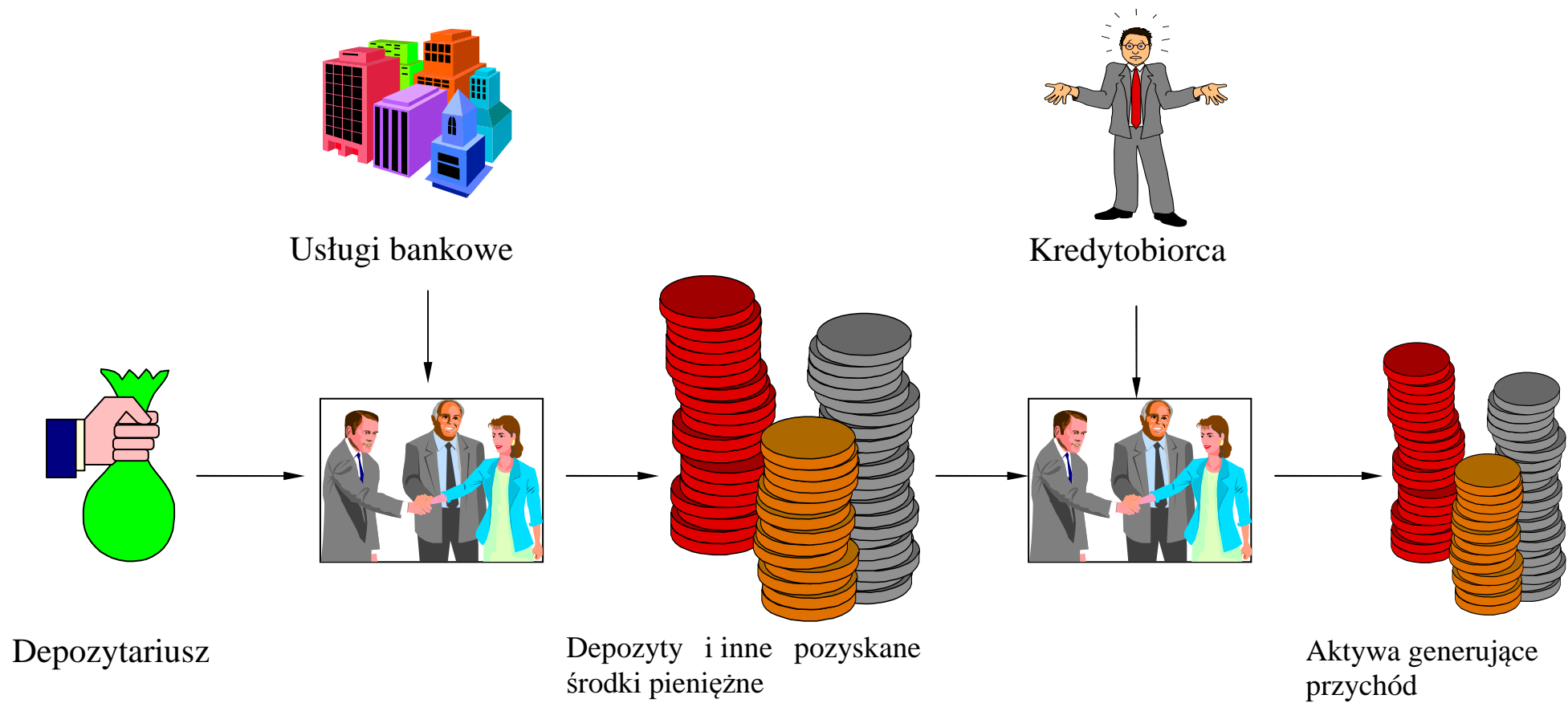
Proponowane podejście różni się od poprzednich, po pierwsze - podejściem do roli depozytów, po drugie - sposobem pomiaru produktów, wyrażając ich wielkość w jednostkach pieniężnych. Inni badacze zajmujący się tym problemem, np. Pesek i Towey, swoje propozycje klasyfikacji czynników i produktów opierali na założeniu, że różnica między podmiotami niefinansowymi a finansowymi polega na tym, że produkty instytucji finansowych nie przynoszą bezpośrednich przychodów, lecz reprezentują realny koszt dla firmy. W konsekwencji traktowali oni m.in. depozyty na żądanie wyrażone wartościowo jako miarę produkcji banków komercyjnych (Pesek [1970] i Towey [1974]). Sealey i Lindley zakwestionowali sensowność tego założenia i wyjaśnili, że skupiając się na makroekonomicznych aspektach systemu bankowego można uzasadnić, iż operacje banków komercyjnych związane z koncentracją depozytów na żądanie są ważnym aspektem ich

aktywności. Uważają oni jednak, że nie można stosować tego rozumowania jako usprawiedliwienia w traktowaniu depozytów jako mikroekonomicznych produktów instytucji finansowych. Błąd tego podejścia można łatwo zrozumieć, znajdując analogie w innych zagadnieniach ekonomii. Na przykład, ekonomiści pracy są zainteresowani określeniem równowagi na rynku pracy - podaży i popytu na pracę ludzką, lecz nie rozważają oni zatrudnionych w przedsiębiorstwach jako produktu tych firm. Usługi świadczone deponentom przez instytucje finansowe (zwłaszcza przez banki komercyjne) przynoszą tym instytucjom korzyści w postaci dodatkowego przychodu (np. w postaci prowizji od operacji gotówkowych), i dlatego niektórzy ekonomiści uważają, że to one są ich produktami. Oczywiście, w technicznym sensie banki świadczą te usługi, ale w każdorazowej analizie otwartą kwestią pozostaje czy są one też produktami w sensie ekonomicznym.

Sealey i Lindley zwrócili uwagę na fakt, że mikroekonomiczna teoria zachowania się firm jest modelem, w którym bank zainteresowany jest przede wszystkim oceną jego aktywności przez rynek i maksymalizacją swego zysku, a nie korzyściami otrzymywanymi przez społeczeństwo. Tylko w tym kontekście należy prowadzić ekonomiczną klasyfikację produktów bankowych. Uregulowania organizacyjne i prawne sektora finansowego, odzwierciedlające jego makroekonomiczną rolę, należy traktować w mikroekonomicznej analizie produkcji i kosztów jako warunki zewnętrzne regulujące ich działalność (jako specyficzne i wspólne dla danej branży warunki ograniczające zakres podejmowanych decyzji).

W przypadku instytucji finansowych mikroekonomicznymi produktami są zatem różne kategorie aktywów generujących przychód. Pomiaru produkcji dokonuje się przez wyrażenie różnych typów tych aktywów w jednostkach pieniężnych. Czynnikiem produkcji są: praca, kapitał i pozyskane przez bank środki pieniężne (w tym depozyty), których zaangażowanie mierzy się odpowiednio liczbą zatrudnionych, wartością składników majątku trwałego i wartością pozyskanych środków. Proces produkcji wg podejścia *real approach* polega na „... transformacji depozytów w kredyty, analogicznie jak w przemyśle motoryzacyjnym ze stali i innych surowców wytwarza się samochody” (zob. Dietsch [1993]).

Rysunek 2. Proces produkcji w instytucjach finansowych (schemat).



## 2.2. Czynniki produkcji i produkty banku w badaniach efektywności ekonomicznej.

Badania empiryczne z zakresu analizy produkcji i kosztu banków prowadzi się często wg omówionego w poprzednim podrozdziale teorii banku zwanej *real approach*. Mimo trwającej już od 20 lat dyskusji, stosowane są alternatywne sposoby określenia produktów i czynników produkcji. Pierwsze, historycznie wcześniejsze podejście „produkcyjne” (*production approach*) opiera się na definicji procesu produkcyjnego w sensie technicznym. Drugi sposób określenia czynników i produktów oparty na definicji procesu produkcyjnego w sensie ekonomicznym zaproponowali Sealey i Lindley [1977]. We współczesnej literaturze podejście to najczęściej określa się jako „intermediacyjne” (*intermediation approach*; Berger i Humphrey [1991], English, Grosskopf, Hayes i Yaiswarng [1993], Mester [1993] używają określenia *assets approach*). O obu podejściach: „produkcyjnym” i „intermediacyjnym” w kontekście empirycznych badań efektywności ekonomicznej piszą m.in: Berger, Hanweck i Humphrey [1987], Dietsch [1993], Ferrier i Lovell [1990], Grabowski, Ragan i Rezvanian [1993], Humphrey [1985], Kaparakis, Miller i Noulas [1994], Muldur i Sassenou [1993], Noulas, Subhash i Miller [1990]. Dalszy rozwój badań spowodował pojawianie się nowych koncepcji ekonomicznej definicji czynników i produktów banku: *user cost approach* i *value-added approach*, omówionych i poddanych krytyce m.in. w pracy Bergera i Humphreya [1982].

Podejście „produkcyjne” traktuje banki jako firmy, które używając kapitału i pracy świadczą różne kategorie usług. Z punktu widzenia przyjętej w pracy teorii banku są to produkty w sensie technicznym (a więc zarówno produkty finalne, jak i półprodukty np. depozyty). Muldur i Sassenou [1993] dzielą te usługi na dwie grupy: pierwsze - pozyskują środki pieniężne (depozyty na żądanie, depozyty terminowe, wkłady oszczędnościowe, itd.), drugie - umożliwiają ich sprzedaż (jako kredyty, inwestycje, itp.). Obydwie kategorie usług traktowane są jako różne produkty, więc jako jednostkę miary do określenia wielkości ich produkcji stosuje się albo liczbę bankowych rachunków albo liczbę transakcji (operacji) bankowych. Produkt jest traktowany jako strumień wyrażony przez ilość produktu wytwarzanego w jednostce czasu. W analizie kosztu rozważa się wtedy jedynie koszty operacyjne, będące pochodną liczby czynnych rachunków bankowych i operacji z nimi związanych, które wystąpiły w badanym okresie. Humphrey [1985] natomiast traktuje banki

jako podmioty, które „produkują” depozyty na żądanie, lokaty i wkłady oszczędnościowe, kredyty komercyjne, kredyty hipoteczne i ratalne przy użyciu kapitału fizycznego, ludzkiego i materiałów. Miarą produkcji jest liczba rachunków depozytowych i kredytowych, a na koszty całkowite składają się koszty produkcji wspomnianych pięciu produktów.

Pomiar świadczenia różnych typów usług bankowych dokonuje się także przez wyrażenie jej przez liczbę rachunków bankowych (zob. Berger, Hanweck i Humphrey [1987], Ferrier i Lovell [1990], Humphrey [1985] i [1993], Lawrence [1989], Sherman i Gold [1985]). Inni jak Berg, Forsund, Hjalmarsson i Souminen [1993], Dietsch [1993], Kim [1986], McAllister i McManus [1993], Rangan, Grabowski, Aly i Pasurka [1988] wyrażają produkt bankowy wartościowo, za pomocą wartości rachunków depozytowych i kredytowych.

Podejście produkcyjne jest krytykowane m.in. ze względu na ignorowanie roli depozytów i innych pozyskanych przez bank środków jako czynników niezbędnych do finansowania kredytów i innych składników aktywów. Shaffer i Dawid [1986] przytaczają następujące argumenty za tym, aby depozyty traktowane były jako czynniki produkcji:

- po pierwsze - depozyty generują koszt i nie rodzą przychodu, zanim nie zostaną przeznaczone na kredyty i inne składniki aktywów;
- po drugie - depozyty są substytutami innych źródeł środków pieniężnych wykorzystywanych w produkcji kredytów i innych składników aktywów (należności banku).

Proponowany sposób pomiaru produktu za pomocą liczby rachunków ignoruje także związki między wielkością rachunku a kosztem jego obsługi i utrzymania go przez bank. Wielkość środków na rachunkach i liczba transakcji mają wpływ na całkowite koszty wytworzenia tak rozumianego produktu. Tymczasem w podejściu produkcyjnym, rozważając funkcję kosztów, uwzględnia się tylko koszty operacyjne i pomija koszty odsetkowe od depozytów, co ma istotny wpływ m.in. na oszacowanie efektu skali.

Również Kolari i Zardkoohi [1987] poddają krytyce sposób mierzenia produktu za pomocą liczby rachunków, a nie wartościowo w jednostkach pieniężnych. Uważają, że użycie liczby rachunków nie może być poprawnym miernikiem produktu, ponieważ wszystkie rachunki nie powodują powstawania tych samych kosztów. Na przykład, aktywny rachunek depozytowy na żądanie może być bardziej kosztowny niż terminowy. Podobnie rachunek, na którym znajdują się duże środki, może pociągać za sobą mniejsze koszty niż rachunek o małych osadzie środków, lecz z dużą liczbą transakcji. Wyrażenie wielkości produktów w jednostkach pieniężnych nie ignoruje natomiast wieloproduktowej natury banków



komercyjnych, lecz pozwala na pomiar produkcji, uwzględniając szeroki wachlarz usług bankowych, gdy rozważa się różne formy aktywności banków.

W podejściu „intermediacyjnym” traktuje się bank jako pośrednika między deponentami a kredytobiorcami, który dokonuje transformacji depozytów w kredyty, czerpiąc z tego korzyści (kierując się kryterium maksymalizacji zysku). Zobowiązania banku, stanowiąc surowiec do finansowania potencjalnych źródeł przychodu banku (kredytów, lokat w papiery wartościowe, akredytyw i innych pożyczonych środków), mają charakter czynników produkcji. I analogicznie - aktywa banku mają własności produktów, ponieważ ich odpowiednie użycie pozwala na uzyskanie bezpośrednich przychodów. Kredyty oraz inne należności (aktywa generujące przychód) są traktowane jako produkty bankowe, zaś depozyty i inne zobowiązania (kapitał finansowy), są – obok oprócz kapitału fizycznego i ludzkiego - czynnikami produkcji. Wartość rachunków kredytowych jest odpowiednią miarą produktów bankowych, a koszty operacyjne plus koszty odsetkowe pozyskanych środków składają się na koszty całkowite. Podejście „intermediacyjne” jest stosowane w następujących pracach z zakresu efektywności banków komercyjnych lub innych instytucji finansowych: Akhainen, Swamy, Taubman i Singamsetti [1997], Berger i Humphrey [1991], Cebenoyan, Cooperman, Register i Hudgins [1993], English, Grosskopf, Hayes i Yaiswarng [1993], Grabowski, Ragan i Rezvanian [1993], Hassan, Grabowski, Pasurka i Ragan [1990], Hughes i Mester [1993], Humphrey [1985], Kaparakis, Miller i Noulas [1994], Mester [1987], [1993], Muldur i Sassenou [1993], Noulas, Subhash, Miller [1990], Zardokoohi i Kolari [1994].

Krytyka tego podejścia dotyczy m.in. roli depozytów na żądanie jako czynnika produkcji, który powoduje powstawanie kosztów operacyjnych mimo braku kosztów odsetkowych (dotyczy to przeważnie rachunków podmiotów gospodarczych i osób fizycznych prowadzących działalność gospodarczą). Operacje wypłat z rachunków bieżących i wpłat generują natomiast przychody w postaci prowizji od wartości przelewów oraz zryczałtowanych opłat za prowadzenie rachunku (jednorazowo także za jego otwarcie). Banki próbują ustalić prowizje tak, aby pokryły rzeczywiście poniesione koszty. Takie spojrzenie na depozyty na żądanie powoduje, że m.in. w pracach Berger i Humphrey [1991], Grabowski, Ragan i Rezvanian [1993], Hassan, Grabowski, Pasurka i Ragan [1990] są one traktowane jako produkty bankowe. Oczywiście, środki na rachunkach bieżących wykazują duże fluktuacje i klient banku może nimi dysponować w każdej chwili, więc banki nie mogą przeznaczyć ich na kredyty bądź inne należności. Średni dzienny stan środków na rachunkach bieżących zależy przede wszystkim od wpłat i wypłat dokonywanych na rachunku klienta

banku. Zgodnie „ze złotą regułą bankową”, mającą na celu zabezpieczenie płynności banków komercyjnych, wartość depozytów na żądanie powinna odpowiadać ilości środków płatniczych (gotówce) będących w bezpośrednim jego posiadaniu. Gotówka w kasie banku, jako element aktywów, nie generuje przychodu tak jak to czynią np. kredyty. Skoro depozyty na żądanie są najczęściej nie oprocentowane (przynajmniej w sensie realnym), to w krótkim okresie czasu banki nie mają wpływu na ich wielkość; niektórzy autorzy traktują je więc jako czynnik stały produkcji (m.in. w pracy Berger, Hancock i Humphrey [1993]). W innej pracy depozyty w kwocie mniejszej niż 100 tys. USD, są traktowane jako czynnik stały tak jak kapitał fizyczny (zakłada się, że wielkość osadu środków na rachunku nie zależy bezpośrednio od oferowanej przez bank ceny za nie, lecz od ilości i wartości transakcji przeprowadzanych przez klienta banku), a depozyty większe od tej kwoty są traktowane jako czynnik zmienny (Akhaiven, Swamy, Taubman i Singamsetti [1997]).

W badaniach empirycznych z zakresu analizy efektywności kosztowej sektora bankowego, na etapie budowy mikroekonomicznej funkcji kosztów należy określić (zdefiniować) czynniki produkcji, ich ceny oraz produkty. Badacze wykorzystujący podejście „intermediacyjne” jako czynniki produkcji przyjmują: kapitał fizyczny (majątek trwały), kapitał ludzki (zatrudnieni w banku), kapitał finansowy (depozyty i inne pożyczone pieniądze); zob. Akhaiven, Swamy, Taubman i Singamsetti [1997], Berger [1993], Berger i Humphrey [1991], Cebenoyan, Cooperman, Register i Hudgins [1993], Dietsch [1993], English, Grosskopf, Hayes i Yaiswarng [1993], Grabowski, Ragan i Rezvanian [1993], Hughes i Mester [1993], Humphrey [1993], Kaparakis, Miller i Noulas [1994], Mester [1993], Muldur i Sassenou [1993]. W podejściu „produkcyjnym” czynnikami produkcji są tylko kapitał fizyczny i ludzki.

Cenę czynnika produkcji wyznacza się jako iloraz kosztu związanego z danym czynnikiem i wielkości jego zaangażowania. Zgodnie z tą koncepcją, cenę kapitału ludzkiego wyraża się przez stosunek wynagrodzenia pracowników banku do liczby zatrudnionych (w przeliczeniu na pełne etaty). Ceną kapitału finansowego jest iloraz kosztów odsetkowych od depozytów (i innych pożyczonych pieniędzy) i wielkości tych depozytów, czyli średnie oprocentowanie tych środków. Cenę kapitału fizycznego najlepiej można wyrazić przez średni koszt wynajmu powierzchni biurowej na rynku lokalnym, na którym działa bank (zob. Berger, Hanweck i Humphrey [1987], Humphrey [1993]). W praktyce na wskutek trudności z uzyskaniem tej wielkości, za cenę przyjmuje się iloraz wydatków związanych z majątkiem trwałym i jego wartości (zob. Berger [1993], Cebenoyan, Cooperman, Register i Hudgins

[1993], Grabowski, Ragan i Rezvanian [1993], Humphrey [1985], Kaparakis, Miller, Noulas [1994], Muldur i Sassenou [1993]). Jako zmienna zastępująca cenę kapitału fizycznego przyjmuje się niekiedy iloraz wspomnianych wydatków i wielkości depozytów (zob. Dietsch [1993], Ferrier, Lovell [1990]). Koszt całkowity (zmienna  $TC$ ) jest sumą kosztów związanych z zaangażowaniem każdego z czynników produkcji: kapitału fizycznego, ludzkiego i finansowego.

Sposób wyrażenia produkcji banku zależy od stosowanego podejścia, lecz badacze zwykle agregują ją tak, aby liczba produktów nie przekroczyła pięciu. Stosujący podejście „intermediacyjne” Mester [1993] oraz English, Grosskopf, Hayes, Yaiswarng [1993] rozważają trzy produkty:

- kredyty hipoteczne (*real estate mortgage loans*);
- kredyty konsumpcyjne i kredyty dla przedsiębiorstw (*consumer and commercial loans*);
- gwarancje, poręczenia i inne należności od podmiotów finansowych (*securities and other investments, assets in trading accounts*).

Noulas, Subhash i Miller [1990] oraz Kaparakis, Miller i Noulas [1994] traktują kredyty konsumpcyjne dla indywidualnych gospodarstw domowych i osób fizycznych nie prowadzących działalności gospodarczej jako czwarty produkt. Akhainen, Swamy, Taubman i Singamsetti [1997] agregują produkcję banku do dwóch produktów: kredytów ratalnych (*installment loans*) oraz pozostałych zawierających kredyty dla przedsiębiorstw, kredyty hipoteczne, faktoring i kredyt z tytułu leasingu (*industrial loans*).

Przyjmujący alternatywne podejście „produkcyjne” Ferrier i Lovell [1990] oraz Berger, Hanweck i Humphrey [1987] zaproponowali pomiar produkcji banku przez liczbę depozytów na żądanie, lokat terminowych, liczbę udzielonych kredytów hipotecznych, ratalnych i kredytów dla przedsiębiorstw. Berger i Humphrey [1991] oraz McAllister i McManus [1993] traktują natomiast depozyty na żądanie (*demand deposits*), depozyty terminowe i wkłady oszczędnościowe (*time and savings deposits*) jak produkty bankowe. Podobny podział stosują Rangan, Grabowski, Aly i Pasurka [1988], lecz nie wyodrębniają oni kredytów ratalnych, a osobną kategorią produktów są dla nich kredyty konsumpcyjne. Kredyty hipoteczne, kredyty kasowe, kredyty z tytułu leasingu i faktoring dla przedsiębiorstw, kredyt konsumpcyjny, gwarancje, poręczenia i inne należności od podmiotów finansowych oraz depozyty na żądanie stanowią produkty dla takich autorów, jak Grabowski, Ragan i Rezvanian [1993] i Hassan, Grabowski, Pasurka i Ragan [1990].

Przy próbie dokonania analizy kosztu i badania efektywności w przypadku polskiego sektora finansowego zasadniczymi problemami (oprócz problemów metodologicznych) wydają się: określenie czynników i produktów (w zgodzie zarówno z naszymi realiami jak i metodologią wypracowaną na Zachodzie) oraz dostęp do porównywalnych danych o polskich bankach. Podejście Sealeya i Lindleya wydaje się poprawnym podejściem w sensie mikroekonomicznej teorii firm. Głównymi kryteriami decydującymi o typach produktów kredytowych mogą być: czas trwania i cel umowy finansowej (kredytowej), typ zabezpieczenia, osoba wierzyciela, sposób oprocentowania lub sposób wypłaty i spłaty kredytu. Te czynniki będą podstawą do określenia kategorii produktów w warunkach polskiej bankowości, zwłaszcza że wachlarz usług bankowych, z których korzystają klienci polskich banków, różni się od oferty produktowej np. banków amerykańskich, najczęściej badanych. Za finansowe czynniki produkcji przyjmuje się różne rodzaje depozytów i innych środków pozyskanych przez bank. Oczywiście postać mikroekonomicznej funkcji kosztu determinuje liczbę czynników i produktów.

Zaproponowany model depozytowo-kredytowy banku stanowi uproszczoną konstrukcję, która umożliwia sformułowanie metodologii pomiaru i badania efektywności kosztowej banków. W szczególności model ten można wykorzystać w określeniu produktów i czynników produkcji w przypadku analizy ekonomicznej oddziałów jednego banku komercyjnego.

### **2.3. Funkcja kosztów w przypadku oddziałów banku komercyjnego.**

Przy budowie funkcji kosztu służącej do badania i pomiaru efektywności oddziałów jednego z polskich banków komercyjnych zasadniczymi problemami wydaje się być określenie - zgodnie z przedstawioną metodologią - mikroekonomicznych produktów oddziałów banku oraz czynników produkcji. Oczywiście aktualny jest problem dostępu do odpowiednich danych (z uwagi na istniejący układ sprawozdawczości banku). Fakt, że bada się oddziały jednego banku ma wpływ na określenie mikroekonomicznych produktów i czynników produkcji, lecz równocześnie gwarantuje porównywalność pozyskanych danych. Zgodnie z dyskusją przeprowadzoną w poprzednich częściach pracy, wykorzystując podejście intermediacyjne za czynniki produkcji przyjęto:

- pracowników oddziału (zmienną oznaczono jako  $L$ );
- nieruchomości (a także inwestycje w nieruchomości obce) –  $K$ ;
- inne środki trwałe (m.in. sprzęt komputerowy) oraz wartości niematerialne i prawne (m.in. oprogramowanie) -  $M$ ;
- depozyty (złotowe i walutowe) i środki pieniężne pozyskane przez oddział z centrali i innych oddziałów -  $D$ .

Rozróżnienie dwóch kategorii kapitału fizycznego,  $K$  i  $M$ , jest spowodowane przyjęciem założenia, że nieruchomości muszą być traktowane w analizie krótkookresowej jako czynnik stały (nie podlegający optymalizacji) w rozważanym tutaj horyzoncie czasowym. Natomiast charakter czynnika  $M$  powoduje, że można go traktować jako czynnik zmienny, a zatem w funkcji kosztów pojawia się jego cena. Jeżeli centrala banku na drodze przetargu wybiera dostawców oprogramowania oraz sprzętu (m.in. komputerów, drukarek, akcesoriów komputerowych) dla oddziałów, to można przyjąć, że cena tego czynnika ( $w_M$ ) jest jednakowa dla każdego oddziału.

Oddziały banku mają siedziby w miastach różnej wielkości (np. poniżej 50 tys. lub powyżej 500 tys. mieszkańców), w regionach o zróżnicowanym charakterze (przemysłowym, rolniczym, turystycznym), zakres świadczonych usług zależy więc od charakteru lokalnego rynku i potrzeb klientów. Podstawowy podział wyróżnia: oddziały depozytowe, tj. posiadające nadwyżkę przyjętych lokat (depozytów) od ludności i przedsiębiorstw ponad wartość udzielonych kredytów, oraz oddziały kredytowe, w których wartość udzielonych kredytów przewyższa wartość depozytów. Samodzielne banki, mające nadwyżkę udzielonych kredytów w stosunku do przyjętych lokat, mogą ją finansować poprzez pożyczki w innych bankach, mających nadmiar pozyskanych środków, lub poprzez zakup środków na rynku międzybankowym, lecz takie operacje prowadzi wyłącznie centrala banku. W przypadku oddziałów banku przepływ tych środków następuje bądź w ramach wzajemnych rozliczeń międzyoddziałowych, bądź rozliczeń między centralą i oddziałem. Podział oddziałów na depozytowe i kredytowe powoduje m.in., że wynik z działalności operacyjnej oddziałów kredytowych koryguje się *in minus* o różnicę między wartością udzielonych kredytów a wartością posiadanych depozytów, przeliczoną wg ceny pieniądza ustalonej przez centralę (na podstawie stawki WIBOR). Uzyskuje się w ten sposób tzw. skorygowany wynik z działalności operacyjnej oddziałów, a nadwyżka kredytów nad depozytami stanowi może być traktowana jako finansowy czynnik produkcji o cenie ustalonej, która jest wyższa od przeciętnego oprocentowania depozytów avista i terminowych od osób fizycznych i

prawnych. Natomiast w przypadku oddziałów depozytowych dokonuje się korekty wyniku z działalności operacyjnej *in plus* o różnicę między wielkością depozytów nad kredytami pomnożoną przez cenę umowną, ustalaną przez centralę na poziomie niższym niż minimalne oprocentowanie kredytów dla osób fizycznych i prawnych. Zatem w przypadku rozważanej krótkookresowej funkcji kosztu zmiennego przyjęto, iż owa różnica stanowi dla oddziałów kredytowych dodatkowy zmienny czynnik produkcji o cenie jednakowej dla wszystkich oddziałów. Jednakże z uwagi na przyjętą translogarytmiczną postać funkcji kosztu, będzie on traktowany łącznie wraz z depozytami jako jeden czynnik produkcji<sup>9</sup>.

Zaangażowanie czynników produkcji proponuje się mierzyć odpowiednio przez: liczbę zatrudnionych w przeliczeniu na pełne etaty, powierzchnię pomieszczeń biurowych własnych i najmowanych przez oddział, wartość brutto składników niematerialnych i prawnych oraz wartość depozytów wraz z innymi pozyskanymi środkami pieniężnymi. Produkcję oddziałów bankowych ( $Q$ ) proponuje się wyrazić przez łączną wartość różnego rodzaju kredytów złotych i walutowych (m.in. kredyty w rachunku bieżącym, obrotowy, ratalny, dyskontowy, inwestycyjny), pożyczek gotówkowych dla ludności, faktoringu, zrealizowanych gwarancji i poręczeń oraz - w przypadku oddziałów depozytowych - nadwyżki między wartością przyjętych lokat a wartością udzielonych kredytów; nadwyżka ta może być traktowana jak dodatkowy typ produktu, odsprzedawanego centrali i oddziałom kredytowym (po cenie ustalonej przez centralę i niższej niż rynkowa).

Całkowite koszty zmienne ( $VC$ ) stanowią sumę kosztów związanych z zaangażowaniem każdego z czynników zmiennych, czyli wynagrodzeń wraz z narzutami, kosztów rzeczowych (m.in. zużycie tonera i papieru do drukarek, dyskietek, materiałów biurowych) oraz kosztów kapitału finansowego (odsetek od depozytów, a także prowizji płaconych od zaciągniętych przez dany oddział kredytów i pożyczek). W oddziałach kredytowych w skład kosztów całkowitych wchodzi koszt pozyskania z centrali lub innych oddziałów środków będących różnicą między wartością udzielonych kredytów a wartością przyjętych lokat, tj. wielkość tej różnicy razy cena (wynosząca kilka punktów procentowych mniej niż WIBOR, jednak znacznie więcej niż rynkowa cena depozytów.) Należy pamiętać, że omawiana różnica traktowana jest jako czynnik produkcji. Konstrukcję cen czynników produkcji, przedstawioną poniżej, oparto w głównym zarysie na światowej literaturze

---

<sup>9</sup> W omawianym przypadku nadwyżka kredytów nad depozytami nie może być oddzielnym czynnikiem produkcji, ponieważ dla oddziałów depozytowych nadwyżka ta jest ujemna (tzn. jest dodatkowym produktem), a logarytm jest funkcją określoną dla wartości wyłącznie większych od zera.

przedmiotu, por. Muldur i Sassenou [1993]. Ceną pracy ludzkiej jest średnie wynagrodzenie pracownika oddziału wraz z narzutami ( $w_L$ ). Zakłada się, że wynagrodzenie to odzwierciedla rynkowy koszt zatrudnienia wykwalifikowanego pracownika sektora bankowego w geograficznym otoczeniu oddziału. Ceną czynnika finansowego ( $w_D$ ) - depozytów pozyskanych bezpośrednio od klientów lub innych środków pieniężnych pozyskanych pośrednio poprzez centralę - jest średnia ważona rzeczywistej ceny przyjmowanych lokat bankowych (a więc cena obowiązująca na lokalnym rynku) i omówionej wcześniej wewnętrznej ceny rozliczeniowej środków przeznaczonych na finansowanie ewentualnej nadwyżki kredytów na depozytami. Wagami wykorzystywanymi do konstrukcji  $w_D$  są udziały depozytów i środków z centrali w ich sumie. W szczególności w przypadku oddziałów depozytowych udział środków z centrali wynosi zero, a zatem  $w_D$  pokrywa się dla nich z ceną depozytów.

Zatem przyjmując założenie, że nieruchomości w analizie krótkookresowej traktowane są jako czynnik stały, stochastyczny model graniczny krótkookresowej, translogarytmicznej funkcji kosztu zmiennego w przypadku danych przekrojowo-czasowych przyjmuje postać:

$$\begin{aligned} \ln VC_{it} = & \beta_0^* + \beta_1^* \ln w_{it,D} + \beta_2^* \ln w_{it,L} + \beta_3^* \ln Q_{it} + \beta_4^* \ln K_{it} + \beta_5 \ln w_{it,D} \ln w_{it,L} \\ & + \beta_6 \ln w_{it,D} \ln Q_{it} + \beta_7 \ln w_{it,D} \ln K_{it} + \beta_8 \ln w_{it,L} \ln Q_{it} + \beta_9 \ln w_{it,L} \ln K_{it} \\ & + \beta_{10} \ln Q_{it} \ln K_{it} + \beta_{11} (\ln w_{it,D})^2 + \beta_{12} (\ln w_{it,L})^2 + \beta_{13} (\ln Q_{it})^2 + \beta_{14} (\ln K_{it})^2 \\ & + \beta_{15}^* \ln w_{it,M} + \beta_{16} \ln w_{it,M} \ln w_{it,D} + \beta_{17} \ln w_{it,M} \ln w_{it,L} + \beta_{18} \ln w_{it,M} \ln Q_{it} \\ & + \beta_{19} \ln w_{it,M} \ln K_{it} + \beta_{20} (\ln w_{it,M})^2 + v_{it} + z_{it}, \end{aligned} \quad (21)$$

gdzie:

$VC_{it}$  = obserwowany poziom kosztu zmiennego (koszty kapitału ludzkiego - wynagrodzeń, koszty kapitału finansowego w postaci odsetek od depozytów i innych pozyskanych środków oraz koszty zaangażowania kapitału fizycznego z wyjątkiem nieruchomości)  $i$ -tego oddziału banku w okresie  $t$  ( $i=1, \dots, N$ ;  $t=1, \dots, T$ )

$w_{it,D}$  = cena czynnika finansowego (depozytów i innych pozyskanych środków, w tym finansujących nadwyżkę kredytów nad depozytami) = (koszty odsetkowe)/(średnia wartość tych środków),

$w_{it,L}$  = cena pracy = (wynagrodzenia wraz z narzutami)/(średnia liczba zatrudnionych w przeliczeniu na pełne etaty),

$w_{it,M}$  = cena zmiennego kapitału fizycznego (sprzętu komputerowego, wartości niematerialnych i prawnych (m.in. oprogramowania), jednakowa dla każdego z oddziałów, czyli  $w_{it,M} = w_{t,M}$ .

$Q_{it}$  = zaobserwowana wielkość produkcji wyrażona poprzez agregat wartości różnego rodzaju kredytów + nadwyżka depozytów nad kredytami w dyspozycji centrali (jeżeli jest dodatnia),

$K_{it}$  = zaangażowanie czynnika stałego (kapitału fizycznego) mierzonego poprzez powierzchnię w metrach kwadratowych pomieszczeń biurowych (budynków) własnych i najmowanych, a eksploatowanych przez oddział. Ze względu na rozważany horyzont czasowy cztery kwartały jednego roku przyjmuje się, że  $K_{it}$  jest stałe po czasie dla  $i$ -tego oddziału, czyli  $K_{it} = K_i$ .

Cena jednego wyróżnionego zmiennego czynnika fizycznego ( $w_{it,M}$ ) jest stała dla każdego oddziału, lecz na tym etapie specyfikacji modelu granicznego dopuszcza się jej zmienność po czasie. Zatem podstawowy warunek ekonomicznej regularności, tj. homogeniczności względem cen czynników zmiennych będzie spełniony, gdy każdą z cen i koszt wydzieli się przez jedną z nich, np.  $w_{it,M}$ . Wielkość ta z uwagi na dość złożoną jej konstrukcję nie jest obserwowana, zatem traktuje się ją jako nieznaną stałą. Po odpowiednich przekształceniach równanie (21) przyjmie formę:

$$\begin{aligned} \ln VC_{it} = & \beta_{0,t} + \beta_{1,t} \ln w_{it,D} + \beta_{2,t} \ln w_{it,L} + \beta_{3,t} \ln Q_{it} + \beta_{4,t} \ln K_{it} + \\ & \beta_5 \ln w_{it,D} \ln w_{it,L} + \beta_6 \ln w_{it,D} \ln Q_{it} + \beta_7 \ln w_{it,D} \ln K_{it} + \\ & \beta_8 \ln w_{it,L} \ln Q_{it} + \beta_9 \ln w_{it,L} \ln K_{it} + \beta_{10} \ln Q_{it} \ln K_{it} + \beta_{11} (\ln w_{it,D})^2 + \\ & \beta_{12} (\ln w_{it,L})^2 + \beta_{13} (\ln Q_{it})^2 + \beta_{14} (\ln K_{it})^2 + v_{it} + z_{it}, \end{aligned} \quad (22)$$

gdzie:

$$\begin{aligned} \beta_{0,t} &= \beta_0^* - \ln w_{t,M} (1 + \beta_1^* + \beta_2^* - \ln w_{t,M} (\beta_5 + \beta_{11} + \beta_{12})) \\ \beta_{1,t} &= \beta_1^* - \ln w_{t,M} (\beta_5 + 2\beta_{11}) \\ \beta_{2,t} &= \beta_2^* - \ln w_{t,M} (\beta_5 + 2\beta_{12}) \\ \beta_{3,t} &= \beta_3^* - \ln w_{t,M} (\beta_6 + \beta_8) \\ \beta_{4,t} &= \beta_4^* - \ln w_{t,M} (\beta_7 + \beta_9) \end{aligned} \quad (23)$$

W modelu danym wzorem (22) zmienność w czasie nieobserwowanej ceny  $w_M$  można opisać używając najprostszej parametryzacji poprzez model trendu liniowego (dla logarytmu tej ceny), co odpowiada przyjęcia założenia, że cena charakteryzuje się stałym tempem wzrostu, np. w wyniku inflacji. Przyjmując



$$\ln w_{t,M} = \alpha_0 + \alpha_1 t \quad (24)$$

i podstawiając to wyrażenie do równania (22), po odpowiednich przekształceniach otrzymuje się równanie postaci:

$$\begin{aligned} \ln VC_{it} = & \beta_0^{**} + \beta_1^{**} \ln w_{it,D} + \beta_2^{**} \ln w_{it,L} + \beta_3^{**} \ln Q_{it} + \beta_4^{**} \ln K_{it} + \beta_5 \ln w_{it,D} \ln w_{it,L} + \\ & \beta_6 \ln w_{it,D} \ln Q_{it} + \beta_7 \ln w_{it,D} \ln K_{it} + \beta_8 \ln w_{it,L} \ln Q_{it} + \beta_9 \ln w_{it,L} \ln K_{it} + \\ & \beta_{10} \ln Q_{it} \ln K_{it} + \beta_{11} (\ln w_{it,D})^2 + \beta_{12} (\ln w_{it,L})^2 + \beta_{13} (\ln Q_{it})^2 + \beta_{14} (\ln K_{it})^2 + \\ & \beta_{21} \ln w_{it,D} \cdot t + \beta_{22} \ln w_{it,L} \cdot t + \beta_{23} \ln Q_{it} \cdot t + \beta_{24} \ln K_{it} \cdot t + \\ & \beta_{25} t + \beta_{26} t^2 + v_{it} + z_{it}. \end{aligned} \quad (25)$$

gdzie:

$$\begin{aligned} \beta_0^{**} &= \beta_0^* - \alpha_0 (1 + \beta_1^* + \beta_2^* - \alpha_0 (\beta_5 + \beta_{11} + \beta_{12})) \\ \beta_1^{**} &= \beta_1^* - \alpha_0 (\beta_5 + 2\beta_{11}) \\ \beta_2^{**} &= \beta_2^* - \alpha_0 (\beta_5 + 2\beta_{12}) \\ \beta_3^{**} &= \beta_3^* - \alpha_0 (\beta_6 + \beta_8) \\ \beta_4^{**} &= \beta_4^* - \alpha_0 (\beta_7 + \beta_9) \\ \beta_{21} &= -\alpha_1 (\beta_5 + 2\beta_{11}) \\ \beta_{22} &= -\alpha_1 (\beta_5 + 2\beta_{12}) \\ \beta_{23} &= -\alpha_1 (\beta_6 + \beta_8) \\ \beta_{24} &= -\alpha_1 (\beta_7 + \beta_9) \\ \beta_{25} &= -\alpha_1 (1 + \beta_1^* + \beta_2^* - 2\alpha_0 (\beta_5 + \beta_{11} + \beta_{12})) \\ \beta_{26} &= (\alpha_1)^2 (\beta_5 + \beta_{11} + \beta_{12}) \end{aligned} \quad (26)$$

Otrzymane równanie (25) jest nieliniowe ze względu na parametry  $\alpha$  i  $\beta$ . Równanie to - wraz z warunkami na parametry danymi przez (26) - można traktować jako postać zredukowana modelu z restrykcjami na parametry, otrzymanego z modelu składającego się z równań danych wzorem (22) i (24). Liczba parametrów w tym równaniu jest o 4 większa od łącznej liczby parametrów postaci strukturalnej. Pojawia się zatem problem wyboru odpowiedniej metody estymacji modelu (25), zwłaszcza, że proponowane w dalszej części pracy techniki estymacji bayesowskiej przy użyciu algorytmu Gibbsa wymagają, aby funkcja kosztu  $h(x_{it}, \beta)$  była liniowa względem parametrów. Problemy z estymacją modelu (25) skłaniają do przyjęcia założenia, że cena  $w_M$  jest stała nie tylko dla każdego oddziału, ale także po czasie, czyli  $w_{itM} = w_M = \exp(\alpha_0)$  dla każdego  $i$  i  $t$ . W rzeczywistości przetargi na zakup sprzętu komputerowego są ogłaszane przez centralę mniej więcej raz do roku, a ich dostawa do oddziału jest organizowana na zasadzie wymiany zużytego sprzętu na nowy. W

części empirycznej utrzymanie tego założenie w mocy wydaje się być uzasadnione, ponieważ estymacji modelu oparta zostanie o dane pochodzące z krótkiego szeregu czasowego (z czterech kwartałów). W efekcie przyjmując, że  $K_{it}=K_i$  i  $w_{it,M} = w_M = \exp(\alpha_0)$  model będący przedmiotem dalszego zainteresowania przybiera postać:

$$\begin{aligned} \ln VC_{it} = & \beta_0 + \beta_1 \ln w_{it,D} + \beta_2 \ln w_{it,L} + \beta_3 \ln Q_{it} + \beta_4 \ln K_i + \beta_5 \ln w_{it,D} \ln w_{it,L} \\ & + \beta_6 \ln w_{it,D} \ln Q_{it} + \beta_7 \ln w_{it,D} \ln K_i + \beta_8 \ln w_{it,L} \ln Q_{it} + \beta_9 \ln w_{it,L} \ln K_i \\ & + \beta_{10} \ln Q_{it} \ln K_i + \beta_{11} (\ln w_{it,D})^2 + \beta_{12} (\ln w_{it,L})^2 + \beta_{13} (\ln Q_{it})^2 + \beta_{14} (\ln K_i)^2 \\ & + v_{it} + z_{it}, \end{aligned} \quad (27)$$

gdzie

$$\begin{aligned} \beta_0 &= \beta_0^* - \ln w_M (1 + \beta_1^* + \beta_2^* - \ln w_M (\beta_5 + \beta_{11} + \beta_{12})) \\ \beta_1 &= \beta_1^* - \ln w_M (\beta_5 + 2\beta_{11}) \\ \beta_2 &= \beta_2^* - \ln w_M (\beta_5 + 2\beta_{12}) \\ \beta_3 &= \beta_3^* - \ln w_M (\beta_6 + \beta_8) \\ \beta_4 &= \beta_4^* - \ln w_M (\beta_7 + \beta_9). \end{aligned}$$

Specyfikacja (27), zakładająca jeden zaagregowany produkt oddziału, będzie punktem wyjścia większości dalszych analiz empirycznych.

Analiza potencjalnych korzyści zakresu produkcji wymaga rozważania przynajmniej dwóch produktów. W tym celu zbudowano model, w którym rozróżniono dwa rodzaje kredytów, tj. kredyty udzielone klientom detalicznych (osobom fizycznym z wyłączeniem prowadzących działalność gospodarczą; zmienna  $Q_2$ ) i kredyty udzielone klientom komercyjnym (podmiotom finansowym, budżetowym i niefinansowym z wyłączeniem osób fizycznych nie prowadzących działalności gospodarczej;  $Q_1$ ). Produkt  $Q_1$  obejmuje również nadwyżkę depozytów nad kredytami (w przypadku oddziałów specjalizujących się pozyskiwaniu środków), która następnie jest przekazywana do centrali bądź innych oddziałów. Postać końcowa tego dwuproduktowego modelu przyjmie postać:

$$\begin{aligned} \ln VC_{it} = & \delta_0 + \delta_1 \ln w_{it,D} + \delta_2 \ln w_{it,L} + \delta_3 \ln Q_{it,1} + \delta_4 \ln K_i + \delta_5 \ln w_{it,D} \ln w_{it,L} + \\ & \delta_6 \ln w_{it,D} \ln Q_{it,1} + \delta_7 \ln w_{it,D} \ln K_i + \delta_8 \ln w_{it,L} \ln Q_{it,1} + \delta_9 \ln w_{it,L} \ln K_i + \\ & \delta_{10} \ln Q_{it,1} \ln K_i + \delta_{11} (\ln w_{it,D})^2 + \delta_{12} (\ln w_{it,L})^2 + \delta_{13} (\ln Q_{it,1})^2 + \delta_{14} (\ln K_i)^2 + \\ & \delta_{15} \ln Q_{it,2} + \delta_{16} \ln w_{it,D} \ln Q_{it,2} + \delta_{17} \ln w_{it,L} \ln Q_{it,2} + \delta_{18} \ln Q_{it,1} \ln Q_{it,2} + \\ & \delta_{19} \ln Q_{it,2} \ln K_i + \delta_{20} (\ln Q_{it,2})^2 + v_{it} + z_{it}, \end{aligned} \quad (28)$$

Uwzględnienie jednej dodatkowej zmiennej spowodowało wzrost liczby parametrów granicznej translogarytmicznej funkcji kosztu o sześć, więc w sumie jest ich 21, a więc dużo w

stosunku do 58 obserwacji po firmach. W przypadku ewentualnej estymacji powyższego równania dla danych przekrojowych mała liczba stopni swobody miałaby zapewne istotny wpływ na precyzję szacunku parametrów; estymacja równania (28) dokonana będzie jedynie na podstawie danych przekrojowo-czasowych. Analiza procesu produkcji i kosztu w przypadku wielu produktów stanowi przedmiot dalszych, pogłębionych badań, a w części empirycznej zostaną zaprezentowane wyniki tylko dla dwóch wyżej wymienionych produktów.

W dalszej części pracy zostaną przedstawione metodologiczne aspekty estymacji granicznej funkcji kosztu, modelowania i pomiaru efektywności kosztowej przy użyciu podejścia klasycznego i bayesowskiego, zarówno w przypadku danych przekrojowych jak i przekrojowo-czasowych.

### 3. Elementy klasycznej estymacji modeli granicznych.

#### 3.1. Podstawowe techniki w przypadku danych przekrojowych.

Najprostsze metody estymacji jednorównaniowego stochastycznego modelu granicznego funkcji kosztu oparte są na metodzie najmniejszych kwadratów (MNK). W przypadku danych przekrojowych ( $T=1$ ) model dany wzorem (1), gdy graniczna funkcji kosztu  $h(x_i, \beta)$  jest liniowa względem  $K+1$  parametrów, przybiera postać:

$$y_i = \beta_0 + x_i \beta_{(-0)} + v_i + z_i, \quad (i=1, \dots, N), \quad (29)$$

gdzie  $\beta_{(-0)} = [\beta_1, \dots, \beta_K]^T$  to wektor  $K$  nieznanych parametrów charakteryzujących technologię (z wyłączeniem wyrazu wolnego  $\beta_0$ ), składniki losowe  $v_i$  posiadają niezależne identyczne rozkłady normalne o średniej zero i skończonej wariancji (co oznacza się  $v_i \sim iin(0, \sigma_v^2)$ ). Ponadto, identycznie jako w modelu (1),  $z_i$  i  $v_i$  są zmiennymi losowymi niezależnymi od siebie ( $z_i \perp v_i$ ) i od regresorów. Z uwagi na fakt, że w modelu (29) występuje  $N$  wielkości nieobserwowalnych  $z_i$  modelujących potencjalną nieefektywność, a całe wnioskowanie o nich może być oparte na podstawie tej samej liczby obserwacji ( $N$ ), wprowadza się oszczędną parametryzację modelu zakładając, że składniki  $z_i$  posiadają niezależne między sobą rozkłady ( $z_i \perp z_j$ ) indeksowane przez wspólne parametry. W tym celu wystarczy np. przyjąć, że niezależne nieujemne zmienne losowe  $z_i$  posiadają identyczną wartość oczekiwaną  $E(z_i) = \mu$ . Należy zauważyć, że suma zmiennych losowych  $\varepsilon_i = v_i + z_i$  ma rozkład, który nie jest rozkładem normalnym, ani nawet rozkładem symetrycznym. Co najważniejsze, wartość oczekiwana złożonego składnika losowego  $\varepsilon_i$  jest niezerowa:  $E(\varepsilon_i) = E(v_i) + E(z_i) = E(z_i) = \mu > 0$ . Jednakże przyjmując oznaczenia  $\beta_\mu = \beta_0 + \mu$  oraz  $\varepsilon_{\mu,i} = \varepsilon_i - \mu$ , otrzymuje się równoważny modelowi (29) zapis:

$$y_i = \underbrace{\beta_0 + \mu}_{\beta_\mu} + x_i \beta_{(-0)} + \underbrace{v_i + z_i - \mu}_{\varepsilon_{\mu,i}}, \quad (30)$$

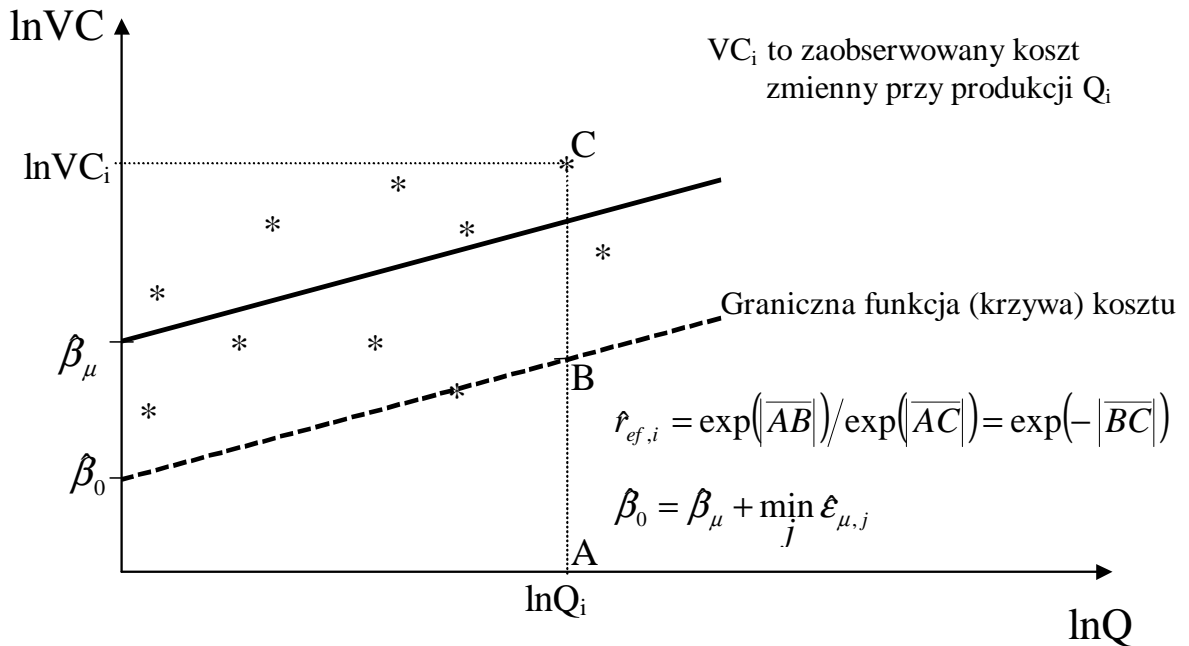
w którym wyrazem wolnym jest  $\beta_\mu$  zaś zakłóceniem losowym  $\varepsilon_{\mu,i}$ , a wektor  $x_i$  o wymiarze  $1 \times K$  zawiera obserwacje  $K$  zmiennych występujących w mikroekonomicznej funkcji kosztu. Ponieważ  $E(\varepsilon_{\mu,i}) = E(\varepsilon_i) - \mu = E(z_i) - \mu = 0$ , więc - przy standardowych założeniach o zmiennych objaśniających (przy założeniu braku skorelowania między  $z_i$  i składowymi wektora  $x_i$ ) - estymator MNK wektora zawierającego  $\beta_\mu$  i wszystkie współczynniki  $\beta_k$  ( $k=1, \dots, K$ ) funkcji kosztu danej wzorem (30) jest estymatorem (przynajmniej) zgodnym, jeżeli tylko  $N \rightarrow \infty$  (por. Greene [1993]). Uzasadnione jest więc wykorzystanie MNK do estymacji równania (30), jednakże w ten sposób uzyska się tzw. przeciętną funkcję kosztu. Funkcję graniczną otrzyma się w wyniku korekty oceny wyrazu wolnego ( $\beta_\mu$ ), gdyż w sposób zgodny szacuje się  $\beta_\mu = \beta_0 + \mu$ , a nie  $\beta_0$ . Zatem zgodność estymatora MNK wszystkich współczynników funkcji kosztów poza wyrazem wolnym może być podstawą różnych wariantów estymacji granicznej funkcji kosztów i wskaźników efektywności kosztowej, jeżeli tylko uwzględni się korektę wyrazu wolnego.

Wariant najprostszy nosi nawet skorygowanej MNK (ang. *corrected ordinary least squares; COLS*) i został zaproponowany przez Winstena [1957], choć to odkrycie przypisuje się zwykle Gabrielsenowi [1975]. Sprowadza się on do estymacji MNK równania (30) i wyznaczenia reszt  $\varepsilon_{\mu,i} = y_i - \hat{y}_i$ . Zakłada się, że jedna z badanych firm jest w pełni efektywna, więc mierzy się nieefektywność innych w relacji do tej wyróżnionej firmy. Zatem firmę, dla której reszta  $\varepsilon_{\mu,i}$  jest minimalna przyjmuje się jako w pełni efektywną (jako wzorzec efektywności). Minimalna reszta  $\varepsilon_{\mu,i}$  (która musi być ujemna traktowana) jest jako oszacowanie błędu systematycznego  $-\mu$ . Graficzną ilustrację tej metody przedstawia Rysunek 3 (przyjęte założenia: jeden produkt Q; ceny czynników produkcji oraz nakłady czynników stałych są ustalone dla wszystkich firm). Oszacowana przy pomocy MNK na podstawie punktów empirycznych krzywa kosztów, tzw. przeciętna krzywa kosztów (linia gruba) po przesunięciu wzdłuż osi rzędnych o wektor  $(0; \min_j(\varepsilon_{\mu,j}))$  tworzy graniczną funkcję kosztów (linia przerywana) o wyrazie wolnym równym  $\beta_0 = \beta_\mu + \min_j(\varepsilon_{\mu,j})$ . Liniowość krzywej kosztów wskazuje, że rozważa się technologię typu Cobba-Douglasa. Miarą względnej efektywności kosztowej będzie iloraz minimalnego kosztu niezbędnego (wynikającego z granicznej funkcji kosztów) do kosztu rzeczywiście poniesionego przy tym samym poziomie produkcji (por. Greene [1993]). Z uwagi na rozważanie logarytmu funkcji kosztu, efektywność  $i$ -tej firmy na skali (0,1] mierzy się jako:

$$\hat{r}_{ef,i} = \frac{\exp(\hat{y}_i + \min_j(\hat{\varepsilon}_{\mu,j}))}{\exp(y_i)} = \frac{\exp(\hat{y}_i + \min_j(\hat{\varepsilon}_{\mu,j}))}{\exp(\hat{y}_i + \hat{\varepsilon}_{\mu,i})} = \exp\left(\min_j(\hat{\varepsilon}_{\mu,j}) - \hat{\varepsilon}_{\mu,i}\right). \quad (31)$$

Przedstawiona metoda estymacji wskaźników efektywności jest wrażliwa na obserwacje nietypowe i intuicyjnie powinna dawać tym lepsze wyniki im mniejsze są symetryczne zakłócenia losowe (błąd pomiaru); pierwotnie była ona zaproponowana dla tzw. deterministycznych funkcji granicznych gdzie  $v_i=0$ ; por. Greene [1980]. Metoda ta nie wykorzystuje w sposób jawny informacji o złożonym składniku  $\varepsilon_i = v_i + z_i$ , a zwłaszcza o rozkładzie jego składowej  $z_i$  i nie pozwala na badanie egzogenicznych przyczyn potencjalnej nieefektywności. Nie są znane własności stochastyczne oszacowań efektywności uzyskanych tym uproszczonym sposobem, można więc traktować je jako podejście wstępne. Główną zaletą tej metody jest prostota obliczeń – do oszacowania stosuje się wyłącznie MNK, zaś oceny efektywności kosztowej uzyskuje się z reszt MNK.

Rysunek 3. Wyznaczanie wskaźnika efektywności na podstawie danych przekrojowych - skorygowana MNK.



W innym wariantcie estymacja granicznej funkcji polega na korekcie wyrazu wolnego o ocenę wartości oczekiwanej złożonego składnika losowego  $\varepsilon_i = v_i + z_i$ , po uprzednim zastosowaniu MNK do równania (30). Sprowadza się ona zatem do oszacowania parametru  $\mu$ , o którą koryguje się  $\beta_\mu$  w celu znalezienia  $\beta_0$ . Wymaga to przyjęcia pewnych założeń co do

rozkładu nieefektywności kosztowej  $z_i$ . Najczęściej przyjmuje się, że zmienna  $z_i$  ma rozkład wykładniczy o wartości oczekiwanej  $\mu$  ( $z_i \sim \text{EXP}(\mu^{-1})$ ) i wariancji  $\mu^2$  albo półnormalny (tj. normalny w wartości oczekiwanej zero, ucięty w punkcie zero), tzn.  $z_i \sim |N(0, \sigma_z^2)|$ . Przyjmuje się też inne rozkłady określone na dodatniej półosi, np. normalny ucięty w zerze, gamma, Weibulla, lognormalny czy loglogistyczny (por. Greene[1990], [1993], Stevenson [1980], Lovell[1993]).

W celu znalezienie oszacowania wartości oczekiwanej  $E(z_i)$  oraz pozostałych parametrów rozkładu złożonego składnika losowego  $\varepsilon_{\mu,i}$  stosuje się np. Metodę Momentów, czyli porównuje się odpowiednią liczbę niezerowych momentów teoretycznych składników losowych  $\varepsilon_{\mu,i}$  z momentami empirycznymi reszt MNK równania (30). Następnie dokonuje się następującej korekty wyrazu wolnego tego równania:

$$\hat{\beta}_0 = \hat{\beta}_\mu - E(z_i) = \hat{\beta}_\mu - \hat{\mu}.$$

Omówiona technika nosi nazwę Zmodyfikowanej Metody Najmniejszych Kwadratów i została zaproponowana przez Richmonda [1974].

Powyższe rozważania można zilustrować przykładem dla przypadku rozkładu wykładniczego dla  $z_i$ . Niech  $z_i \sim \text{EXP}(\mu^{-1})$ , a  $v_i \sim N(0, \sigma_v^2)$ , co prowadzi do następujących wzorów na estymatory nieznanymi parametrów  $\mu$  oraz  $\sigma_v$  (por. Osiewalski, Wróbel-Rotter [1999]):

$$\hat{\mu} = \sqrt[3]{\frac{1}{2N} \sum_{i=1}^N (\varepsilon_{\mu,i})^3} \quad \text{oraz} \quad \hat{\sigma}_v^2 = \frac{1}{N} \sum_{i=1}^N (\varepsilon_{\mu,i})^3 - (\hat{\mu})^2. \quad (32)$$

Estymatory zdefiniowane powyżej istnieją, jeżeli tylko trzeci moment reszt MNK  $\frac{1}{N} \sum_{i=1}^N (\varepsilon_{\mu,i})^3$  jest mniejszy od zera oraz  $\hat{\sigma}_v^2 > 0$ . Warunki te nie zawsze są w praktyce spełnione, co dyskryminuje tę metodę. Zatem jest ona wrażliwa na dane, w efekcie nie zawsze estymatory uzyskanie w ten sposób istnieją. Przyczyną tego faktu może być to, że porównuje się momenty teoretyczne zmiennej  $\varepsilon_{\mu,i}$  z momentami empirycznymi, a nie teoretycznymi reszt MNK  $\varepsilon_{\mu,i}$ . Ponadto nie ma pewności, że po przesunięciu przeciętnej funkcji kosztu wszystkie odchylenia od niej będą nieujemne, co powoduje, że stosowanie tej metody wymaga innego miernika samej nieefektywności niż ten dany wzorem (31). Jondrow, Lovell, Materov i Schmidt [1982] zaproponowali uniwersalną formułę pozwalającą na wyznaczenie oceny

punktowej nieefektywności poprzez wartość oczekiwaną rozkładu zmiennej  $z_i$  warunkowego względem złożonego składnika losowego  $\varepsilon_i$ . W przypadku rozkładu wykładniczego dla  $z_i$  otrzymuje się:

$$E(z_i|\varepsilon_i) = \sigma_v \left( \frac{\varphi(A_i)}{\Phi(A_i)} + A_i \right), \quad (33)$$

gdzie:  $\varphi(\cdot)$  oznacza funkcję gęstości rozkładu  $N(0,1)$ ,  $\Phi(\cdot)$  to dystrybuanta standaryzowanego rozkładu normalnego, zaś  $A_i = \varepsilon_i/\sigma_v - \sigma_v/\mu$ . Natomiast nieznane parametry w powyższym wzorze zastępuje się ocenami uzyskanymi w sposób opisany wzorem (32) lub w inny sposób, np. stosując metodę największej wiarygodności (MNW), którą krótko opisano w dalszej części.

Koop i Mullahy [1989] zaproponowali wykorzystanie Uogólnionej Metody Momentów (ang. Generalized Method of Moments) do estymacji stochastycznych modeli granicznych. Wymaga ona jedynie, aby zmienna odzwierciedlająca nieefektywność  $z_i$  miała dowolny rozkład ucięty w punkcie zero określony na dodatniej półosi, zaś zakłócenie losowe  $v_i$  miało rozkład symetryczny. Podstawowym założeniem jest, że pierwsze sześć momentów zmiennych  $v_i$  i  $z_i$  ma być skończonych, co jednak powoduje, że przy tak ogólnych założeniach metoda ta nie ma praktycznego znaczenia, bo nie daje zadawalających estymatorów nieefektywności. Dodatkowo wadą tej metody jak i obu przedstawionych modyfikacji metody MNK jest to, że nie pozwalają one na narzucenie warunków regularności ekonomicznych na parametry wektora  $\beta$ , co możliwe jest w przypadku MNW, którą zaproponowali Pitt i Lee [1981]. W popularyzacji tej metody mieli duży udział też Jondrow, Lovell, Materov i Schmidt. W tym przypadku utrzymane są w mocy założenia z modelu (29), a dodatkowo przyjmuje się konkretny rozkład prawdopodobieństwa dla  $z_i$  (w praktyce przyjmuje się ucięty normalny lub wykładniczy). MNW polega na znalezieniu takich wartości wektora parametrów  $\beta$  i parametrów struktury stochastycznej, które gwarantują największą gęstość prawdopodobieństwa próbkowego dla wektora obserwacji zmiennej objaśnianej  $y=(y_1, \dots, y_N)$ . Niezależność  $v_i$  i  $z_i$  pozwala przedstawić łączną gęstość  $v_i$  i  $z_i$  jako iloczyn gęstości tych zmiennych, odpowiednio  $f_v(v_i)$  i  $f_z(z_i)$ . Zatem gęstość złożonego składnika losowego  $\varepsilon_i=v_i+z_i$  otrzymuje się jako rozkład brzegowy iloczynu  $f_v(v_i)$  i  $f_z(z_i)$  względem  $z_i$ :

$$f_\varepsilon(\varepsilon_i) = \int_0^\infty f_v(\varepsilon_i - z_i) f_z(z_i) dz_i .$$



Funkcja podcałkowa jest mieszkanką warunkowego rozkładu normalnego i odpowiedniego rozkładu brzegowego zmiennej  $z_i$ , określonego na dodatniej półosi. W efekcie funkcję wiarygodności można zapisać:

$$\ell(\theta|y) = \prod_{i=1}^N f_{\varepsilon}(\varepsilon_i) = \prod_{i=1}^N \int_0^{\infty} f_v(\varepsilon_i - z_i) f_z(z_i) dz_i,$$

gdzie  $\theta$  oznacza wektor parametrów strukturalnych rozpatrywanego modelu, składający się z wektora  $\beta$  i parametrów charakteryzujących rozkłady prawdopodobieństwa  $z_i$  i  $v_i$ , a  $\varepsilon_i = y_i - (\beta_0 + x_i \beta_{(-0)})$ .

Warto zauważyć, że postać funkcji wiarygodności w przypadku nawet najprostszych rozkładów dla  $z_i$  jest na tyle skomplikowana, że całki występujące w iloczynie można szacować jedynie przy pomocy metod numerycznych. Należy je liczyć z dużą precyzją, ponieważ w przypadku, gdy jedna z nich przyjmuje bardzo małe wartości (bliskie zero), to zaokrąglenie tej wartości do zera powoduje, iż funkcja wiarygodności przyjmuje wartość zero bez względu na wartość innych składowych tego iloczynu. Własności asymptotyczne estymatora MNW w przypadku tych modeli nie są w pełni zbadane, ale przypuszcza się, że jest on zgodny i asymptotycznie efektywny, gdy  $N \rightarrow \infty$  (por. Schimdt i Sickles [1984]).

W przypadku szczególnym, gdy  $z_i \sim \text{EXP}(\mu^{-1})$  funkcja wiarygodności ma postać:

$$\ell(\beta, \mu, \sigma_v^2|y) = \mu^{-N} \prod_{i=1}^N \left[ \exp\left(\frac{\sigma_v^2}{2\mu^2} - \frac{\varepsilon_i}{\mu}\right) \Phi(A_i) \right],$$

gdzie  $A_i$  i  $\Phi(\cdot)$  oznaczają to samo co we wzorze (33). Zatem problemy numeryczne napotyka się przy obliczaniu dystrybuanty standaryzowanego rozkładu normalnego  $\Phi(\cdot)$ , zwłaszcza dla argumentu  $A_i$  znacznie mniejszego niż -5. Pewne rozwiązanie tego problemu zaproponowali Van den Broeck, Koop, Osiewalski i Steel [1994], stosując aproksymację tej dystrybuanty za pomocą ułamków łańcuchowych rzędu 20, gwarantujących dużą dokładność obliczeń.

Oceny wektora parametrów  $\theta$  uzyskane metodą MNW pozwalają na dokonanie pomiaru nieefektywności w sposób zaproponowany przez Jondrowa, Lovella, Materova i Schmidta [1982], poprzez wartość oczekiwaną rozkładu zmiennej  $z_i$  warunkowego względem złożonego składnika losowego  $\varepsilon_i$ , co w omawianym przykładzie sprowadza się do wykorzystania formuły (33). Metodę MNW dla stochastycznych modeli granicznych w badaniu efektywności kosztowej banków stosują m.in. Cebenoyan, Cooperman, Register, Hudgins [1993], Ferrier i Lovell [1990], Kaparakis, Miller i Noulas [1994], Lawrence [1989], Hughes i Mester [1993] oraz Mester [1993].

### 3.2. Model efektów stałych w przypadku danych przekrojowo-czasowych.

W przypadku posiadania danych przekrojowo-czasowych częstym założeniem (przyjmowanym zwłaszcza przy małym  $T$ ) jest stałość efektywności firmy w czasie ( $z_{it}=z_i$  dla  $t=1, \dots, T$ ), zatem efektywność jest czystym efektem indywidualnym (*individual effect*) każdej z firm. Oznacza to, że efektywność  $i$ -tej firmy w badanym okresie jest modelowana przez tę samą zmienną losową niezależnie od okresu czasu. Przyjęcie tego założenia umożliwia bardziej precyzyjny szacunek przeciętnej efektywności każdej firmy, gdyż odbywa się on na podstawie  $T$  obserwacji, a nie jednej. Ponadto przyjęcie tego założenia jest konieczne, aby równanie (34) było identyfikowalne, jeżeli tylko  $T > 1$ . W efekcie graniczny model funkcji kosztów można zapisać w postaci:

$$y_{it} = \beta_0 + x_{it}\beta_{(-0)} + v_{it} + z_i, \quad (i=1, \dots, N; t=1, \dots, T), \quad (34)$$

gdzie przyjęte oznaczenia są identyczne jak w równaniach (29).

Najprostszy sposób estymacji powyższego równania zaproponowali Schmidt i Sickles [1984]. Zauważyli, że skoro  $E(z_i)=\mu > 0$  w modelu (34), to w modelu  $y_{it} = (\beta_0 + \mu) + x_{it}\beta_{-0} + (v_{it} + z_i - \mu)$  składnik losowy  $v_{it} + z_i - \mu$  ma wartość oczekiwaną zero. Nie zalecają jednak estymacji takiego modelu zwykłą MNK. Estymator MNK parametrów  $\beta_{\mu} = \beta_0 + \mu$  i  $\beta_{-0}$  jest co prawda zgodny, jeżeli tylko  $N \rightarrow \infty$  i brak jest korelacji między  $z_i$  i regresorami  $x_{it}$  (choć nie jest zgodny w przypadku jeżeli  $T \rightarrow \infty$  dla ustalonego  $N$ ), ale istnieją lepsze, efektywniejsze metody estymacji równania (34) w celu uzyskania granicznej funkcji kosztu. W konsekwencji autorzy ci proponują traktować składniki  $z_i$  jako nieznane stałe i estymować w równaniu (34) obok  $K$  parametrów przy regresorach także  $N$  indywidualnych wyrazów wolnych  $\beta_{0,i}$  zdefiniowanych jako  $\beta_{0,i} = \beta_0 + z_i$ , które rozróżniają indywidualną efektywność kosztową firm. Zatem równanie (34) reprezentuje tzw. **model efektów stałych** (*fixed effects model*), który na gruncie klasycznym charakteryzuje się tym, że składniki  $z_i$  traktuje się jako nieznane stałe podlegające estymacji. Natomiast w klasycznym **modelu efektów losowych** traktuje się je jako zmienne losowe o pewnych rozkładach indeksowanych przez nieznane parametry. Model efektów stałych podlegający estymacji przybiera formę:

$$y_{it} = \underbrace{\beta_{0,i}}_{\beta_0 + z_i} + x_{it}\beta_{(-0)} + v_{it}, \quad (i=1, \dots, N; t=1, \dots, T), \quad (35)$$

Szacowanie  $K+N$  parametrów odbywa się na podstawie  $T \cdot N$  obserwacji. Estymacja modelu, w którym występują wyrazy wolne  $(\beta_{0,i})$  różne dla każdej z firm, polega na zastosowaniu MNK z uwzględnieniem w zbiorze zmiennych objaśniających dodatkowych  $N-1$  zmiennych zero—jedynkowych obok stałej  $\beta_0$ . Równoważną algebraicznie jest procedura estymacji równania (35), w którym wartość każdej ze zmiennych: objaśnianej i objaśniających (np.  $y_{it}$ ) zastępuje się odchyleniem od średniej po czasie (wyrażeniem  $y_{it} - \frac{1}{T} \sum_{t=1}^T y_{it}$ ). W literaturze estymator ten jest znany jako estymator wewnątrzgrupowy (ang. *within estimator*; zob. Judge, Griffiths, Hill, Lütkepohl, Lee [1985]). Dla wektora  $\beta_{(-0)}$  jest on zgodny w obu przypadkach:  $N \rightarrow \infty, T \rightarrow \infty$ . Natomiast dla wyrazów wolnych  $\beta_{0,i}$  estymator ten jest zgodny wraz ze wzrostem liczby obserwacji po czasie dla każdej firmy (a więc tylko przy  $T \rightarrow \infty$ ). Ponieważ proponowany model ma być wykorzystywany jedynie przy małym  $T$ , więc własności asymptotyczne związane ze wzrostem liczby okresów czasu nie mają w tym przypadku praktycznego znaczenia, zwłaszcza, że w przypadku, gdy  $T \rightarrow \infty$  trudno utrzymać założenie o stałości technologii. Ponadto stosowanie tego estymatora wymaga, aby wartości wszystkich zmiennych egzogenicznych (oprócz stałej) były zmienne po czasie. Po dokonaniu estymacji parametrów modelu (35) przyjmuje się:

$$\hat{z}_i = \hat{\beta}_{0,i} - \hat{\beta}_{0,\min}, \quad (36)$$

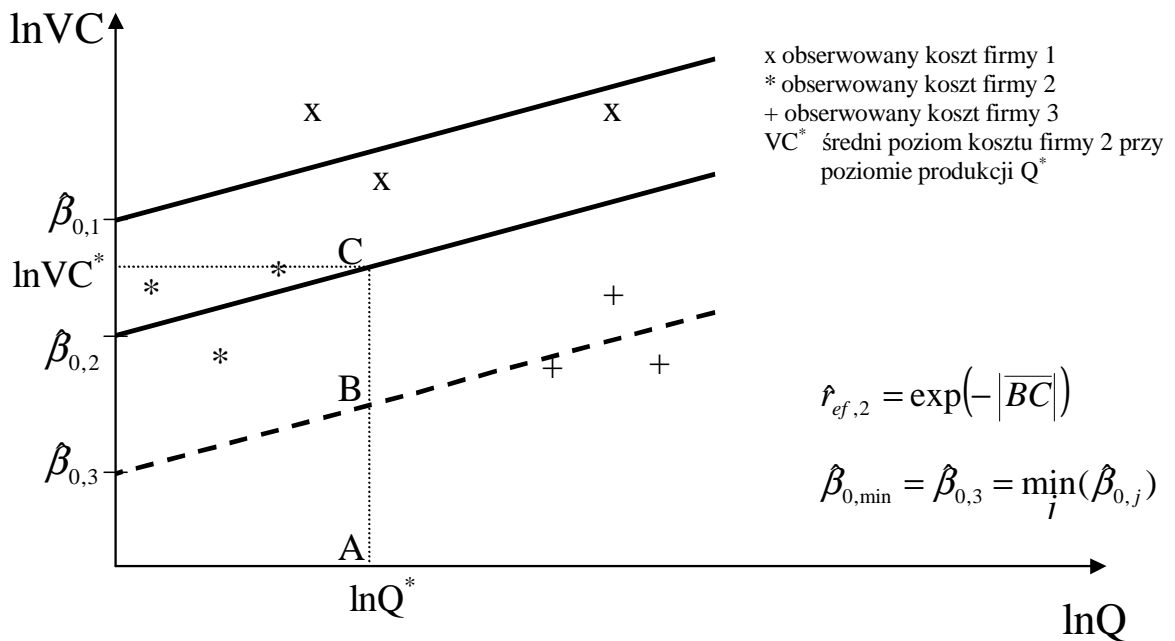
gdzie  $\hat{\beta}_{0,i}$  jest oszacowaniem stałej odpowiadającej  $i$ -tej firmie a  $\hat{\beta}_{0,\min} = \min_j (\hat{\beta}_{0,j})$ , jako ocenę wskaźnika nieefektywności  $z_i$  (określonego na skali  $[0, +\infty)$ ). Miernikiem efektywności na skali  $(0, 1]$  dla  $i$ -tej firmy jest:

$$\hat{r}_{ef,i} = \frac{\exp(\hat{y}_{it} + \min_j (\hat{\beta}_{0,j}))}{\exp(\hat{y}_{it})} = \frac{\exp(\hat{y}_{it} + \hat{\beta}_{0,\min})}{\exp(\hat{y}_{it} + \hat{\beta}_{0,i})} = \exp(-\hat{z}_i), \text{ gdzie:} \quad (37)$$

$\hat{y}_{it} = x_{it} \hat{\beta}_{(-0)} + \hat{v}_{it}$ ,  $\hat{\beta}_{(-0)}$  jest estymatorem wewnątrzgrupowym, a  $\hat{v}_{it}$  resztą z estymacji.

Tak zdefiniowany miernik zakłada, że firma o najmniejszym  $\hat{\beta}_i$  jest w pełni efektywna a nieefektywność pozostałych firm jest mierzona w relacji do niej. Dla estymatora wskaźnika nieefektywności, określenie jego błędu oszacowania na gruncie teorii klasycznej jest bardzo trudne (możliwe jedynie poprzez symulacje komputerowe). Graficznie tę metodę estymacji wskaźników efektywności opisuje Rysunek 4.

Rysunek 4. Wyznaczanie wskaźnika efektywności na podstawie danych panelowych: zindywidualizowane wyrazy wolne.



Koop, Osiewalski i Steel [1997] wykazali, że estymacja wskaźnika efektywności dla danych przekrojowo-czasowych za pomocą estymatora wewnątrzgrupowego odpowiada wnioskowaniu bayesowskiemu, w którym przyjmuje się bardzo silną wiedzę wstępną o wysokiej względnej nieefektywności firm (mierzonej względem firmy najefektywniejszej). Wiedzę tę można opisać formalnie za pomocą niewłaściwego rozkładu a priori postaci  $p(r_i) \propto 1/r_i$  gdzie  $r_i = \exp(\min_j(\beta_{0,j}) - \beta_{0,i})$ , a zatem  $r_i \in (0,1]$ . Rozkład ten ma następujące własności: szansa a priori, że firmy są bardzo nieefektywne ( $r_i$  bliskie 0) jest nieskończenie wielka, natomiast że są bardzo efektywne ( $r_i$  bliskie 1) jest tylko skończona. Należy się więc spodziewać, że zaprezentowana metoda nie doszacowuje w skończonych próbach względnej efektywności firm (rozrzut efektywności między firmą - wzorcem a pozostałymi firmami jest zbyt duży).

Stosując estymator wewnątrzgrupowy nie wykorzystuje się założenia o postaci rozkładu  $z_i$ , bo traktuje się  $z_i$  jako nieznanne, nie powiązane ze sobą stałe liczbowe (efekty stałe). Natomiast w modelach efektów losowych dla konkretnych typów rozkładów nieefektywności i przy założeniu normalności błędu symetrycznego  $v_{it}$  proponuje się w literaturze metodę największej wiarygodności (zob. Pitt i Lee [1981] oraz Schmidt i Sickles [1984]), która została skrótowo opisana dla przypadku  $T=1$  oraz wnioskowanie bayesowskie (zob. Koop, Osiewalski i Steel [1997]). Szczególnie to ostatnie podejście wydaje się naturalne

w przypadku szacowania efektywności; wymaga ono nowoczesnych technik Monte Carlo do obliczania gęstości brzegowych i momentów rozkładu a posteriori (zob. Osiewalski i Steel [1998]).

Posiadanie danych panelowych o dużym interwale czasowym, np. jeden rok, i dłuższym horyzoncie  $T$  pozwala założyć, że efektywność firm jest zmienna w czasie (por. Cornwell, Schmidt i Sickles [1990] oraz Kumbhakar [1990]). Jeżeli przyjmie się dodatkowe założenie, że  $z_{it}$  są niezależnymi zmiennymi losowymi o wartości oczekiwanej  $\mu$  (dla  $i=1, \dots, N$ , i  $t=1, \dots, T$ ), to nie nakłada się na dane żadnej struktury przekrojowo—czasowej (por. Pit i Lee [1981]). W tym przypadku szczególnym graniczna funkcja kosztu przyjmie postać:

$$y_{it} = \underbrace{\beta_0 + \mu}_{\beta_{0,\mu}} + x_{it}\beta_{-0} + \underbrace{v_{it} + z_{it} - \mu}_{\varepsilon_{it,\mu}}, \quad (i=1, \dots, N; t=1, \dots, T). \quad (38)$$

Najprostszy sposób estymacji wskaźników efektywności jest identyczny jak w przypadku danych przekrojowych, czyli równania (29). Uzyskane wskaźniki efektywności na podstawie reszt MNK w liczbie  $T \cdot N$ , oszacowane na podstawie  $T \cdot N$  obserwacji, charakteryzują efektywność każdej z firm w kolejnych badanych okresach. Grupując je po firmach można ponadto uzyskać dodatkowe informacje o średniej efektywności firm w okresie, z którego pochodzi próba. Oczywiście, do tego sposobu estymacji należy odnieść te same uwagi krytyczne co w przypadku estymacji w oparciu o dane przekrojowe: szacunek pojedynczego wskaźnika efektywności oparty jest w zasadzie na jednej obserwacji.

W części następnej zostanie przedstawione podejście bayesowskie do estymacji granicznej funkcji kosztu, wraz z modelowaniem i pomiarem efektywności kosztowej. Będzie to podstawowa metodologia statystyczna dla empirycznych badań w rozdziale piątym tej pracy.

## 4. Bayesowska analiza efektywności.

### 4.1. Elementy podejścia bayesowskiego.

Niech wektor  $y \in R^n$  oznacza możliwy przebieg (możliwą realizację) zjawiska empirycznego. Rozważa się pewną rodzinę rozkładów prawdopodobieństwa indeksowanych przez wektor  $\omega \in \Omega \subset R^s$ , zdefiniowanych przez rodzinę  $P$  funkcji gęstości  $p(y/\omega)$  dla  $\omega \in \Omega$ , czyli  $P = \{p(y/\omega), \omega \in \Omega\}$ . Rozkład  $p(y/\omega)$  jest określany mianem tzw. rozkładu próbkowego. Na gruncie niebayesowskim (klasycznym) rodzina  $P$  określa jednoznacznie parametryczny model statystyczny, w ramach którego można rozważać zagadnienie estymacji wektora  $\omega$ . Na gruncie bayesowskim rodzina  $P$  nie stanowi jeszcze kompletnego modelu statystycznego, którym jest dopiero łączny rozkład prawdopodobieństwa wektora obserwowanego i wektora parametrów:  $p(y, \omega) = p(y/\omega) \cdot p(\omega)$ , gdzie  $p(\omega)$  jest gęstością brzegowego rozkładu wektora parametrów, czyli tzw. rozkładu a priori.

Przyjęcie rozkładu a priori jest cechą wyróżniającą – a wg Lindleya [1972] definiującą – podejście bayesowskie. Jest to jednak rozróżnienie czysto formalne. Tym, co w sposób zasadniczy różnicuje oba podejścia jest odmienna interpretacja prawdopodobieństwa. Na gruncie klasycznym przyjmuje się interpretację częstościową - prawdopodobieństwo jest idealizacją częstości. Natomiast podejście bayesowskie opiera się na interpretacji logicznej bądź subiektywnej – prawdopodobieństwo jest miarą przekonania co do prawdziwości danego zdania logicznego (wystąpienia danego zdarzenia).

Zatem na gruncie bayesowskim każdą nieznaną wielkość traktuje się jako zmienną losową, której rozkład odzwierciedla wiedzę o tej nieznannej wielkości. Różne elementy modelu statystycznego (obserwacje i parametry) traktowane są symetrycznie, więc wstępną wiedzę o wszystkich nieznanach wielkościach (obserwowalnych i nieobserwowalnych) przed zaobserwowaniem zjawiska empirycznego modeluje łączny rozkład prawdopodobieństwa o funkcji gęstości  $p(y, \omega)$ . W bayesowskim modelu statystycznym ogólne zasady estymacji parametrów sprowadzają się do wyznaczenia z rozkładu łącznego  $p(y, \omega)$  warunkowej

gęstości dla wektora parametrów  $\omega$ , przy zaobserwowanym wektorze  $y$ , czyli funkcji gęstości tzw. rozkładu a posteriori, danej wzorem Bayesa:

$$p(\omega|y) = \frac{p(y, \omega)}{p(y)} = \frac{p(y|\omega)p(\omega)}{\int_{\Omega} p(y|\omega)p(\omega)d\omega} \propto l(\omega|y)p(\omega), \quad (39)$$

gdzie symbol  $\propto$  oznacza proporcjonalność, zaś  $l(\omega|y)$  jest funkcją wiarygodności, która jest identyczna z gęstością rozkładu próbkowego ( $l(\omega|y)=p(y|\omega)$ ). Różnica między nimi polega na tym, że funkcję wiarygodności traktuje się jako funkcję parametrów przy danych obserwacjach. Istota estymacji bayesowskiej polega więc na zastosowaniu wzoru Bayesa w celu wyznaczenia rozkładu a posteriori w oparciu o funkcję wiarygodności i rozkład a priori dla parametrów. Gęstość rozkładu a posteriori pozwala formalnie ująć wszelką informację o parametrach posiadanych przed wglądem w dane; w szczególności pozwala modelować warunki regularności ekonomicznej. Natomiast gęstość a posteriori odzwierciedla końcową wiedzę o  $s$ -wymiarowym wektorze parametrów  $\omega$  posiadaną po zaobserwowaniu wektora  $y$ .

O ile zasada estymacji bayesowskiej jest prosta, o tyle praktyczne jej zastosowanie w przypadkach wielowymiarowych może prowadzić do poważnych problemów analitycznych i numerycznych, gdyż obliczanie brzegowych funkcji gęstości oraz ich momentów polega na obliczaniu całek wielokrotnych. W praktyce w przypadku wielowymiarowym ( $s>1$ ), gdy przedmiotem zainteresowania badacza jest składowa wektora  $\omega$  (np.  $i$ -ta,  $i=1, \dots, s$ ), to konieczne jest wyznaczenie brzegowego rozkładu względem pozostałych składowych z rozkładu a posteriori danego wzorem (39) lub momentów i innych charakterystyk tego rozkładu. Rozkład brzegowy dla  $i$ -tej składowej wektora  $\omega$  ( $\omega_i$  dla  $i=1, \dots, s$ ) otrzymuje się poprzez obliczenie następującej  $s-1$  krotnej całki:

$$p(\omega_i|y) = \int_{\Omega_{(-i)}} p(\omega|y)d\omega_{(-i)}, \quad (40)$$

gdzie  $\omega_{(-i)}$  to wektor  $[\omega_1, \dots, \omega_{i-1}, \omega_{i+1}, \dots, \omega_s]$ , a  $\Omega_{(-i)}$  jest zbiorem możliwych jego wartości.

Jeżeli badacza interesują charakterystyki tego rozkładu (o ile one istnieją), to oblicza je poprzez wartość oczekiwaną (według rozkładu a posteriori) odpowiedniej funkcji zmiennej  $\omega_i$ , czyli  $g(\omega_i)$ , co prowadzi się do obliczenia całki:

$$E[g(\omega_i)|y] = \int_{\Omega_{(i)}} g(\omega_i)p(\omega_i|y)d\omega_i. \quad (41)$$

Jeżeli  $g(\omega) = \omega$  lub  $g(\omega) = \omega^2$  to  $E[g(\omega_i)|y]$  oznacza momenty zwykłe a posteriori rzędu 1 i 2 zmiennej  $\omega$ , zaś wariancje a posteriori uzyskuje się bezpośrednio ze wzoru:

$$Var[\omega_i|y] = E[(\omega_i)^2|y] - E[\omega_i|y]^2, \text{ gdzie } E[\omega_i|y] \text{ to wartość oczekiwana } \omega. \quad (42)$$

Oczywiście często w praktyce analityczne wyznaczenie powyższych całek nie jest możliwe, zatem oblicza się je za pomocą metod numerycznych. Jedną z tych metod, tzw. losowanie Gibbsa, zostanie wykorzystane w badaniach empirycznych prezentowanych w dalszej części pracy.

#### 4.2. Bayesowski stochastyczny model graniczny o wspólnym rozkładzie efektywności.

Do analizy stochastycznych modeli granicznych zostaną wykorzystane w części piątej przede wszystkim metody bayesowskie, które zaproponowali i których skutecznie używają Van den Broeck, Koop, Osiewalski i Steel [1994], Koop, Steel i Osiewalski [1995], i Koop, Osiewalski i Steel [dalej KOS 1994, 1997, 1999, 2000]. Podejście bayesowskie było z powodzeniem wykorzystywane w zastosowaniach empirycznych, od badań nad efektywnością szpitali (KOS [1997]) do analizy wzrostu gospodarczego państw (zob. KOS [1999, 2000]). Autorzy ci proponując te metody dowodzą, iż mają one przewagę nad podejściem klasycznym do wnioskowania w tych modelach, ponieważ m.in. są w stanie dostarczyć dokładnych (nieasymptotycznych) procedur wnioskowania oraz w pełni uwzględnić informacje a priori i niepewność co do poprawności specyfikacji modelu. Podejście bayesowskie dostarcza małopróbkowe i probabilistycznie interpretowalne wyniki za cenę wielowymiarowych całkowań numerycznych, których efektywną realizację zapewnia tzw. losowanie Gibbsa, metoda Monte Carlo typu łańcuchów Markowa. W niniejszym rozdziale zostaną przedstawione bayesowskie stochastyczne modele graniczne oraz losowanie Gibbsa jako numerycznej metody wyznaczania rozkładu a posteriori i jego charakterystyk.

Podstawowy stochastyczny model granicznej funkcji kosztów w przypadku obserwacji dla  $N$  firm z  $T$  okresów czasu można zapisać w formie:

$$y_{it} = h(x_{it}, \beta) + v_{it} + z_{it}, \quad (i=1, \dots, N; t=1, \dots, T), \quad (43)$$

gdzie przyjęte oznaczenia są identyczne jak w modelu (1), czyli:  $y_{it}$  – jest wartością zaobserwowanego logarytmu naturalnego kosztu zmiennego ( $VC_{it}$ )  $i$ -tej firmy w okresie  $t$



( $i=1, \dots, N; t=1, \dots, T$ ),  $x_{it}$  to wektor-wiersz zmiennych egzogenicznych (będących funkcjami produktów, cen czynników zmiennych i nakładów czynników stałych),  $h$  oznacza odpowiednio wyspecyfikowaną graniczną funkcją kosztu, liniową względem  $K+1$  nieznanych parametrów  $\beta$ , składniki  $v_{it}$  i  $z_{it}$  są zmiennymi losowymi, z których pierwszy jest symetryczny względem zera, a drugi nieujemny. Zakłada się, że składniki  $z_{it}$  i  $v_{it}$  są niezależne od siebie, zarówno po czasie jak i po firmach. **Miarą efektywności kosztowej jest  $r_{it}=\exp(-z_{it})$** , czyli łatwo interpretowalna wielkość o wartościach w przedziale  $(0,1]$ . W równaniu (43)  $z_{it}$  reprezentuje wzrost logarytmu kosztu spowodowany nieefektywnością techniczną (zbyt duża skala nakładów) lub alokacyjną (proporcje nakładów niezgodne z relacją cen rynkowych czynników). Wprowadzona notacja dopuszcza wykorzystanie danych przekrojowo-czasowych do estymacji przedstawionego modelu, lecz nie wyklucza użycia wyłącznie danych przekrojowych, co odpowiada przyjęciu  $T=1$ .

Pełna bayesowska analiza parametryczna wymaga określenia  $1^0$  rozkładu próbkowego parametryzowanego przez skończenie wymiarowy wektor  $\omega \in \Omega$ ,  $2^0$  rozkładu a priori dla  $\omega$ . W celu spełnienia warunku  $1^0$  zakłada się, że zmienne  $v_{it}$  mają identyczne, niezależne rozkłady normalne ze średnią zero i stałą wariancją  $\sigma_v^2$  (co oznacza się jako  $N(0, \sigma_v^2)$ ), oraz że  $z_{it}$  ma rozkład wykładniczy ze średnią (i odchyleniem standardowym)  $\lambda_{it}$ , czyli  $\text{Exp}(\lambda_{it}^{-1})$ . Czasem, zwłaszcza dla dużego  $T$ , przyjmuje się założenie, że  $z_{it}$  są zmiennymi niezależnymi po  $i$  oraz po  $t$ , jak w KOS [1999, 2000]; zob. także Pitt i Lee [1981]. W KOS [1997] zakłada się, że **efektywność jest czystym efektem indywidualnym dla każdej z firm**, czyli że  $z_{it}=z_i$  ( $t=1, \dots, T$ ); zob. także Pitt i Lee [1981] oraz Schmidt i Sickles [1984]. Argumentami za przyjęciem tego założenia jest fakt, że  $1^0$  w części empirycznej wykorzystane zostaną dane kwartalne z jednego tylko roku,  $2^0$  umożliwi to bardziej precyzyjny szacunek przeciętnej efektywności każdej firmy. Zatem w dalszej części rozważania będą dotyczyły głównie tego przypadku.

W podejściu bayesowskim wyróżnia się tzw. bayesowski model efektów stałych i model efektów losowych. Jednakże kryterium tego podziału jest odmienne niż w ujęciu klasycznym, a sprowadza się ono do różnicy między hierarchiczną i niehierarchiczną budową bayesowskiego parametrycznego modelu statystycznego. W klasycznym modelu efektów stałych składniki  $z_i$  są traktowane jako nieznanne stałe, w odróżnieniu do modelu efektów losowych, w którym dla  $z_i$  przyjmuje się konkretny rozkład, którego parametry są nieznanymi stałymi podlegającymi estymacji. Bayesowskie modele efektów losowych tworzy się przyjmując odpowiednią hierarchiczną strukturę modelu, która powoduje, iż brzegowe

rozkłady każdej ze zmiennych reprezentujących nieefektywność są stochastycznie zależne. Poszczególne przypadki klasycznych i bayesowskich modeli efektów stałych i losowych zawiera poniższa tabela (dla uproszenia przyjęto, że zarówno składniki  $z_i$  jak i parametry tych rozkładów mają rozkłady wykładnicze):

Zał. $z_{it}=z_i$	Model klasyczny	Model bayesowski
Model efektów stałych	$z_i$ traktowane są jako nieznanne stałe podlegające estymacji (model jest identyfikowalny dla $T>1$ ; estymacja omówiona w podrozdziale 5.2)	$z_i   \lambda_1^{-1}, \dots, \lambda_N^{-1} \sim \text{Exp}(\lambda_i^{-1})$ a $\lambda_i^{-1} \sim \text{iiExp}(c)$
Model efektów losowych	$z_i   \lambda_1^{-1}, \dots, \lambda_N^{-1} \sim \text{Exp}(\lambda_i^{-1})$ a $\lambda_i^{-1}$ to nieznanne parametry podlegające estymacji	$z_i   \lambda^{-1} \sim \text{Exp}(\lambda^{-1})$ a $\lambda^{-1} \sim \text{Exp}(c)$ - tzw. prosty model efektów losowych
	$z_i   \lambda_1^{-1}, \dots, \lambda_N^{-1} \sim \text{Exp}(\lambda_i^{-1})$ a $\lambda_i^{-1} = \prod_{j=1}^m \phi_j^{-w_{ij}}$ , gdzie $\phi_j$ to nieznanne parametry podlegające estymacji	$z_i   \lambda_1^{-1}, \dots, \lambda_N^{-1} \sim \text{Exp}(\lambda_i^{-1})$ a $\lambda_i^{-1} = \prod_{j=1}^m \phi_j^{-w_{ij}}$ , gdzie $\phi_j \sim \text{Exp}(g_j)$ , - tzw. uogólniony model efektów losowych

$w_{ij}$  oznacza wartość  $j$ -tej zmiennej egzogenicznej (będącej potencjalną przyczyną nieefektywności)  $i$ -tej firmy.

W modelu bayesowskim efektów stałych, składniki  $z_i$  mają niezależne, warunkowe względem parametrów, rozkłady np. wykładnicze z parametrem  $\lambda_i^{-1}$  ( $z_i | \lambda_1^{-1}, \dots, \lambda_N^{-1} \sim \text{Exp}(\lambda_i^{-1})$ ) dla  $i=1, \dots, N$ ), a parametry  $\lambda_i^{-1}$  mają niezależne znane rozkłady wykładnicze  $\text{Exp}(c)$ . W konsekwencji brzegowe rozkłady dla  $z_i$  (po wycalkowaniu  $\lambda_i^{-1}$ ) są niezależne. Natomiast w modelu efektów losowych, gdy  $z_i | \lambda^{-1} \sim \text{Exp}(\lambda^{-1})$  i  $\lambda^{-1} \sim \text{Exp}(c)$ , to choć  $z_i$  posiadają identyczne, niezależne rozkłady (warunkowe względem wspólnego parametru  $\lambda^{-1}$ ), to ich rozkłady brzegowe względem  $\lambda^{-1}$  ( $p(z_i)$ ) są rozkładami zależnymi (gęstość łączna  $p(z_1, \dots, z_N)$  nie da się przedstawić jako iloczyn niezależnych rozkładów brzegowych  $p(z_i)$ , por. KOS [1997]). Ten przypadek nazywany jest modelem o wspólnym rozkładzie efektywności (Common Efficiency Distribution - CED) i stanowi propozycję KOS [1997]. Bayesowski model efektów losowych uzyskuje się poprzez przyjęcie odpowiedniej hierarchicznej struktury modelu, czyli poprzez przyjęcie założenia, że nieefektywności każdej z firm jest realizacją z tego samego rozkładu parametryzowanego (w przypadku rozkładu wykładniczego) przez wspólny parametr  $\lambda^{-1}$ . Taka hierarchizacja modelu pozwala na wnioskowanie o wielkościach  $z_i$  na podstawie całej próby, a nie jednej obserwacji, ponadto umożliwia uwzględnienie innych czynników mających wpływ na efektywność, np. podobnego sposobu zarządzania w poszczególnych przedsiębiorstwach danej branży (w tym przypadku w oddziałach jednego banku).

Innym przykładem bayesowskiego modelu efektów losowych jest model VED (*Varying Efficiency Distribution*, KOS [1997]), w którym co prawda przyjmuje się  $z_i/\lambda_i^{-1}, \dots, \lambda_i^{-1} \sim \text{Exp}(\lambda_i^{-1})$ , lecz efekt losowy uzyskuje się przyjmując się bardziej złożoną hierarchię modelu tj. rozkład dla parametrów  $\lambda_i^{-1}$  jest indeksowany przez wspólny parametr  $\phi=(\phi_1, \dots, \phi_m)'$  i zmienne egzogeniczne  $(w_{ij})$ , które potencjalnie wyjaśniają nieefektywność. Fernández, Osiewalski i Steel [1997] omawiają dokładnie typy struktur możliwych do rozważania w przypadku danych przekrojowo-czasowych i podstawowe problemy wnioskowania bayesowskiego na podstawie takich modeli granicznych, w szczególności zagadnienia istnienia rozkładu a posteriori.

W przypadku modelu CED, będącego dalej przedmiotem szczególnego zainteresowania, zakłada się, że obserwowane są wartości  $y_{it}$  przy danych  $x_{it}$ , natomiast  $z_i$  są zmiennymi nieobserwowalnymi, a  $\theta=(\beta', \sigma_v^{-2}, \phi_1)'$   $\theta \in \Theta \subset \mathbb{R}^{K+1} \times \mathbb{R}_+ \times \mathbb{R}_+$  jest wektorem parametrów. Zarówno zmienne ukryte, jak parametry są przedmiotem wnioskowania. Rozkład a priori wektora  $\theta$  może być dowolny, jednak zwykle nie wprowadza się zbyt dużo subiektywnej informacji o  $\theta$ . Zakłada się, że rozkłady a priori dla wektora parametrów  $\beta$ , parametru precyzji  $\sigma_v^{-2}$  i  $\phi_1=\lambda^{-1}$  są niezależne między sobą. W rozważanym przypadku przyjmuje się następującą strukturę a priori:

$$p(\theta) = p(\sigma_v^{-2})p(\beta)p(\phi_1) \propto f_G(\sigma_v^{-2} | \frac{n_0}{2}, \frac{s_0}{2})I(\beta \in B)f_G(\phi_1 | 1, g_1), \quad (44)$$

gdzie  $f_G(\cdot | a, b)$  oznacza gęstość rozkładu Gamma ze średnią  $a/b$  i wariancją  $a/b^2$ , przy czym  $a=1$  odpowiada rozkładowi wykładniczemu. Struktura (44) odzwierciedla brak wstępnej wiedzy o wektorze  $\beta$  innej niż warunki regularności,  $\beta \in B$ , narzucone przez teorię ekonomii. Zakłada się więc, że albo  $I(\beta) \equiv 1$  jeżeli nie uwzględnia się warunków regularności, albo  $I(\beta)=1$  dla  $\beta$  spełniających te warunki i  $I(\beta)=0$  dla pozostałych  $\beta$ . Odpowiedni wybór wartości  $n_0 > 0$  i  $s_0 > 0$  pozwoli na przyjęcie nikłej wstępnej informacji o parametrze precyzji  $\sigma_v^{-2}$ , zapewniając zarazem istnienie rozkładu a posteriori, a więc możliwość wnioskowania bayesowskiego nawet dla danych przekrojowych ( $T=1$ ); zob. Fernández, Osiewalski i Steel [1997]. Tę wstępną informację można interpretować jako odpowiadającą hipotetycznej uprzedniej próbie o „liczebności”  $n_0$  i „sumie kwadratów reszt”  $s_0$ , na podstawie której wcześniej uzyskano informację o  $\sigma_v^{-2}$ . Przyjęcie np.  $n_0=s_0=10^{-6}$  prowadzi do bardzo rozproszonego rozkładu Gamma (ze średnią 1 i wariancją  $2 \cdot 10^6$ ). W przypadku modelu CED dla danych przekrojowo-czasowych można zastosować niewłaściwy rozkład a priori Jeffreysa postaci  $p(\sigma_v^{-2}) \propto \sigma_v^{-2}$ ,

który jest przypadkiem granicznym jądra gęstości rozkładu gamma dla  $n_0=s_0=0$ . Dla parametru  $\phi_1$  przyjmuje się rozkład wykładniczy z parametrem  $g_1=-\ln(r^*)$ , gdzie  $r^* \in (0, 1)$  jest stałą zadawaną przez badacza. W tym modelu wielkość  $r^*$  może być interpretowana jako mediana rozkładu a priori efektywności, ponieważ jest medianą brzegowego rozkładu a priori dla  $r_i=\exp(-z_i)$ ; zob. Van den Broeck, Koop, Osiewalski i Steel [1994]. Oznacza to, że z prawdopodobieństwem 0.5 losowo wybrana firma charakteryzuje się – zdaniem badacza - efektywnością poniżej wartości  $r^*$  albo powyżej tej wartości. Postać rozkładu a priori można modyfikować i badać wrażliwość wyników a posteriori na przyjęte założenia; w szczególności można zmieniać wartość  $n_0$ ,  $s_0$  i  $r^*$ .

Ostatecznie, pełny bayesowski stochastyczny model graniczny o wspólnym rozkładzie efektywności, przy założeniu stałości efektywności w badanym okresie, jest określony przez następującą łączną funkcję gęstości, określoną na przestrzeni wszystkich  $T \cdot N$  obserwacji  $y_{it}$  oraz wszystkich wielkości nieobserwowalnych, tj.  $N$  wskaźników nieefektywności  $z_i$  i  $K+3$  parametrów:

$$p(y, z, \theta | X) = p(\theta) p(z | \phi_1) p(y | \theta, z, X) \propto f_G(\sigma_v^{-2} | \frac{n_0}{2}, \frac{s_0}{2}) I(\beta \in B) f_G(\phi_1 | 1, g_1) \quad (45)$$

$$\prod_{i=1}^N f_G(z_i | 1, \phi_1) \prod_{i=1}^N \prod_{t=1}^T f_N^1(y_{it} | h(x_{it}, \beta) - z_i, \sigma_v^2),$$

gdzie  $y$  i  $X$  oznaczają odpowiednio: wektor  $T \cdot N \times 1$  obserwacji  $y_{it}$  i macierz  $T \cdot N \times (K+1)$  z  $x_{it}$  jako wierszami; a wektor  $z$  definiuje się jako wektor o  $N \times 1$  elementach  $z_i$ . Ponadto zapis  $f_N^1(|a, b)$  oznacza jednowymiarowy rozkład normalny o wartości oczekiwanej  $a$  i wariancji  $b$ . Podejście bayesowskie pozwala opisać w sposób probabilistyczny wszystkie informacje o wielkościach nieobserwowalnych (wstępne i pochodzące z danych) w postaci rozkładu a posteriori o gęstości  $p(z, \theta | y, X)$  proporcjonalnej do (45). Dogodnym sposobem sumaryzacji wiedzy zawartej w łącznym rozkładzie a posteriori jest obliczenie takich jego charakterystyk, jak np. kwantyle czy podstawowe momenty (wartości oczekiwane, wariancje i kowariancje - jeśli istnieją). Wnioskowanie o pojedynczych parametrach czy indywidualnych wskaźnikach efektywności wymaga również wyznaczenia z rozkładów  $p(z, \theta | y, X)$  odpowiednich jednowymiarowych brzegowych funkcji gęstości a posteriori. Zarówno momenty jak i kwantyle czy gęstości brzegowe są całkami, które nie mogą być znalezione analitycznie ze względu na zbyt skomplikowaną postać rozkładu a posteriori. Tradycyjne kwadratury i metody Monte Carlo z niezależnymi losowaniami nie stanowią efektywnych narzędzi przybliżania tych całek. Jednakże Koop, Osiewalski i Steel [1997] zauważyli, że rozkłady

warunkowe:  $p(\beta/\sigma_v^{-2}, \phi_l, z, y, X)$ ,  $p(\sigma_v^{-2}/\beta, \phi_l, z, y, X)$ ,  $p(\phi_l/\beta, \sigma_v^{-2}, z, y, X)$  oraz  $p(z/\beta, \sigma_v^{-2}, \phi_l, y, X)$  posiadają znane rozkłady, dlatego można zastosować nowoczesną metodę całkowania Monte Carlo (typu łańcuchów Markowa), znaną jako *Gibbs Sampling*. Losowanie Gibbsa jest to sposób uzyskiwania próbek z łącznego rozkładu wektora losowego na podstawie losowań z pełnych rozkładów warunkowych; szczegóły podają m.in. Casella i George [1992] oraz Tierney [1994], a także Koop, Osiewalski i Steel [1997], których algorytm zostanie wykorzystany w obliczeniach w niniejszej pracy.

#### 4.3. Próbki Gibbsa jako numeryczna metoda wyznaczania rozkładów a posteriori.

Gwałtowny rozwój narzędzi obliczeniowych, jaki obserwuje się od kilkunastu lat, umożliwił zarówno testowanie i pogłębianie wiedzy z zakresu analizy numerycznej, jak i jej implementację. W przypadku wielowymiarowych przestrzeni parametrów praktycznie jedynymi metodami obliczania charakterystyk rozkładu a posteriori są techniki Monte Carlo (MC). Ze względu na typ procesu stochastycznego symulującego konfiguracje parametrów rozważa się dwie grupy metod. Jeśli losuje się niezależnie z tego samego rozkładu, to mówi się o tzw. metodach Monte Carlo z funkcją ważności (ang. *Monte Carlo with Importance Sampling*). Jeśli mechanizm symulacji jest łańcuchem Markowa, to mówi się o markowskich metodach Monte Carlo (ang. *Markov Chain Monte Carlo*). Wśród tych ostatnich prostotą i automatyzmem wyróżnia się losowanie Gibbsa (ang. *Gibbs Sampling*).

Rozważa się łączny rozkład a posteriori  $p(\omega|y)$ , gdzie  $\omega = (\omega_1, \dots, \omega_d)$  jest wektorem parametrów i wielkości nieobserwowalnych; w modelu z podrozdziału 4.2  $\omega = (\theta'z)'$ . Każdy z elementów  $\omega_i$  może być skalarą lub wektorem. Algorytm Gibbsa przebiega następująco (Casella i George [1992], O'Hagan [1994], Tierney [1994]):

1. Przyjmuje się arbitralnie wektor wartości początkowych  $\omega^{(0)} = (\omega_1^{(0)}, \dots, \omega_d^{(0)})$ .
2. Wygenerowanie jednej realizacji wektora  $\omega^{(q)} = (\omega_1^{(q)}, \dots, \omega_d^{(q)})$  w  $q$ -tym cyklu Gibbsa ( $q=1,2,3,\dots$ ) składa się z  $d$  kroków, polegających na sekwencyjnym losowaniu z pełnych warunkowych rozkładów a posteriori:

$$\omega_1^{(q)} \sim p(\omega_1/\omega_2=\omega_2^{(q-1)}, \dots, \omega_d=\omega_d^{(q-1)}, y)$$

$$\omega_2^{(q)} \sim p(\omega_2/\omega_1=\omega_1^{(q)}, \omega_3=\omega_3^{(q-1)}, \dots, \omega_d=\omega_d^{(q-1)}, y)$$

...

$$\omega_{d_l}^{(q)} \sim p(\omega_{d_l} | \omega_{l=1} = \omega_{l-1}^{(q)}, \dots, \omega_{d_{l-1}} = \omega_{d_{l-1}}^{(q)}, y).$$

3. Powtarza się  $S$  razy krok 2 do uzyskania zbieżności algorytmu (cykle spalone), a następnie  $M$  razy - aż do otrzymania zadowalającej aproksymacji obliczanych całek.
4. Oszacowanie interesujących charakterystyk (wartości oczekiwanych, kwantyli a posteriori) dla dowolnych funkcji wektora  $\theta$ , np.  $g(\omega)$ , uzyskuje się poprzez ich odpowiedniki empiryczne, obliczone na podstawie realizacji uzyskanych w kolejnych  $M$  cyklach Gibbsa po odrzuceniu cykli wstępnych.

Przy pewnych ogólnych założeniach (aperiodyczność i nieredukowalność; por. Tierney [1994], O'Hagan [1994]), rozkład wylosowanego wektora  $\omega^{(q)}$  dąży do rozkładu stacjonarnego  $p(\omega|y)$ , gdy  $q \rightarrow \infty$ .

Posługując się rozkładami warunkowymi, uzyskuje się próbę z rozkładu łącznego (choć tylko w sensie asymptotycznym). Wartość oczekiwana a posteriori funkcji  $g(\omega)$  może być więc aproksymowana przez średnią wartość tej funkcji dla  $M$  wylosowanych wektorów  $\omega^{(q)}$ :

$$E[g(\omega)|y, X] \approx \frac{1}{M} \sum_{q=S+1}^{S+M} g(\omega^{(q)}). \quad (46)$$

Histogramy oparte na realizacjach wektorów  $\omega^{(q)} = (\omega_1^{(q)}, \dots, \omega_d^{(q)})$  dla  $q=S+1, \dots, S+M$  mogą posłużyć do narysowania wykresów gęstości brzegowych a posteriori  $p(g(\omega)|y)$ .

Najprostszą praktyczną metodą badania zbieżności jest empiryczna obserwacja aproksymowanych charakterystyk (46) uzyskiwanych w poszczególnych cyklach Gibbsa. Stabilizacja wyników losowania oznacza, że zbieżność w przypadku tej funkcji wektora  $\omega$  została osiągnięta.

Aby zastosować schemat Gibbsa w praktyce, pełne rozkłady warunkowe dla  $\omega^{(q)}$  muszą mieć postać umożliwiającą łatwe losowanie. Jeśli spełniony jest ten warunek, to algorytm Gibbsa staje się metodą automatyczną i może być wówczas stosowany w modelach o bardzo wysokim wymiarze przestrzeni parametrów.

W przypadku granicznego modelu funkcji kosztów należy zatem wyznaczyć warunkowe rozkłady a posteriori dla poszczególnych składowych wektora  $\theta$  i nieobserwowalnych składników  $z_i$ . Sposób ich otrzymania można omówić na przykładzie rozkładu dla parametru precyzji. Zgodnie ze wzorem Bayesa:

$$p(\sigma_v^{-2} | y, z, \beta, \phi_1, X) = \frac{p(y, z, \sigma_v^{-2}, \beta, \phi_1 | X)}{p(y, z, \beta, \phi_1 | X)}.$$

Licznik równania -  $p(y, z, \beta, \phi_1 | X)$  - rozkład brzegowy względem  $\sigma_v^{-2}$  jest całką, którą można przedstawić jako:

$$\begin{aligned} p(y, z, \beta, \phi_1 | X) &= \int_{\Theta_{\sigma_v^{-2}}} p(\sigma_v^{-2}, y, z, \beta, \phi_1 | X) d\sigma_v^{-2} = \int_{\Theta_{\sigma_v^{-2}}} p(y | z, \sigma_v^{-2}, \beta, \phi_1, X) p(\sigma_v^{-2}) p(z, \beta, \phi_1) d\sigma_v^{-2} = \\ &= c_1(y, z, \beta, \phi_1, X) \cdot p(z, \beta, \phi_1), \end{aligned}$$

gdzie  $c_1(\cdot)$  jest odpowiednią całką. Ponadto rozkład próbkowy dla  $y$  zależy od  $\sigma_v^{-2}$ , zatem łączny rozkład dla wielkości obserwowanych, nieobserwowalnych i parametrów ma postać:

$$p(y, z, \sigma_v^{-2}, \beta, \phi_1 | X) = p(y, \sigma_v^{-2} | z, \beta, \phi_1, X) p(z, \beta, \phi_1) = p(y | z, \sigma_v^{-2}, \beta, \phi_1, X) p(\sigma_v^{-2}) p(z, \beta, \phi_1).$$

W efekcie prowadzi to do warunkowej gęstości a posteriori dla  $\sigma_v^{-2}$  określonej wzorem:

$$p(\sigma_v^{-2} | y, z, \beta, \phi_1, X) = \frac{1}{c_1} p(y | z, \sigma_v^{-2}, \beta, \phi_1, X) p(\sigma_v^{-2}), \quad (47)$$

gdzie  $c_1$  jest wartością całki  $c_1(\cdot)$  określonej we wzorze powyżej.

Oczywiście struktura równania (47) jest analogiczna do ogólnego wzoru Bayesa danego formułą (39). Po odpowiednich przekształceniach warunkowy rozkład a posteriori dla parametru precyzji jest rozkładem gamma postaci:

$$p(\sigma_v^{-2} | y, X, z, \beta, \phi_1) = f_G \left( \sigma_v^{-2} \left| \frac{n_0 + T \cdot N}{2}, \frac{1}{2} \left\{ s_0 + \sum_{i,t} (y_{it} + z_i - h(x_{it}, \beta))^2 \right\} \right. \right). \quad (48)$$

Pozostałe warunkowe rozkład a posteriori dla poszczególnych zmiennych otrzymuje się przekształcając wyrażenia:

$$p(\beta | y, z, \sigma_v^{-2}, \phi_1, X) = \frac{1}{c_2} p(y | z, \sigma_v^{-2}, \beta, \phi_1, X) I(\beta) \quad (49)$$

$$p(z | y, \sigma_v^{-2}, \beta, \phi_1, X) = \frac{1}{c_3} p(y | z, \sigma_v^{-2}, \beta, \phi_1, X) p(z) \quad (50)$$

$$p(\phi_1 | y, z, \sigma_v^{-2}, \beta, X) = \frac{1}{c_4} p(z | \phi_1) p(\phi_1), \quad (51)$$

gdzie  $c_2, \dots, c_4$  są odpowiednimi stałymi normującymi.

Po odpowiednich przekształceniach, warunkowe rozkłady a posteriori (wykorzystywane w algorytmie Gibbsa) uzyskują postać:

$$p(\beta | y, X, z, \sigma_v^{-2}, \phi_1) \propto I(\beta \in B) \exp \left[ -\frac{1}{2} \sigma_v^{-2} \sum_{it} (y_{it} - h(x_{it}, \beta) - z_i)^2 \right], \quad (52)$$

$$p(z | y, X, \theta) = \prod_{i=1}^N \frac{1}{\Phi(-m_i/\sigma_v)} f_N^1(z_i | m_i, \sigma_v^{-2}) I(z_i \geq 0), \quad (53)$$

gdzie:  $m_i = \sum_{t=1}^T [y_{it} - h(x_{it}, \beta)] - \sigma_v^2 \phi_1$ , zaś  $\Phi(\cdot)$  oznacza dystrybuantę rozkładu  $N(0,1)$ . Pełny

rozkład warunkowy wektora  $z$  jest więc iloczynem uciętych niezależnych rozkładów normalnych. Natomiast rozkład wspólnego parametru rozkładu zmiennych  $z_i$  ma postać:

$$p(\phi_1 | z) = f_G \left( \phi_1 \left| 1 + N, \sum_{i=1}^N z_i - g_1 \right. \right). \quad (54)$$

Postać rozkładu warunkowego a posteriori dla  $\beta$  zależy od postaci funkcji  $h(x_{it}, \beta)$ . Schemat Gibbsa można łatwo stosować, gdy jest ona liniowa względem  $\beta$ , ponieważ wtedy wzór (52) przedstawia gęstość  $(K+1)$ -wymiarowego rozkładu normalnego, uciętego w przypadku restrykcyj nałożonych na parametry. Otrzymuje się zatem:

$$p(\beta | y, X, z, \sigma_v^{-2}, \phi_1) \propto I(\beta \in B) f_N^{K+1}(\beta | \beta, \sigma_v^2 (X' X)^{-1}), \quad (55)$$

gdzie

$$\beta = (X' X)^{-1} X'(y - z \otimes \mathbf{1}_T), \quad \mathbf{1}_T - \text{wektor jedynek } T \times 1, \quad (56)$$

zaś  $z \otimes \mathbf{1}_T$  oznacza iloczyn Kroneckera wektorów  $z$  i  $\mathbf{1}_T$ .

Przestrzeń możliwych wartości  $\beta$  nie jest całym zbiorem  $R^{K+1}$ , więc restrykcyjne realizuje się losując z  $(K+1)$  - wymiarowego rozkładu normalnego (nie uciętego), a następnie odrzucając te punkty z przestrzeni parametrów, które nie spełniają warunków określonych przez funkcję  $I(\beta)$ .

Próbkowanie Gibbsa polega zatem na cyklicznym losowaniu z rozkładów warunkowych danych wzorami (48), (53)-(55). W pierwszej kolejności podaje się wartości startowe dla zmiennych  $z_i$  i  $\beta$ . W praktyce przyjmuje się w tym celu, że wszystkie firmy charakteryzują się efektywnością na poziomie około np. 0.75, co odpowiada przyjęciu



$z_i^{(0)}=0.3$ . Następnie wyznacza się wstępne oszacowanie wektora nieznanymi parametrów  $\beta$  ze wzoru (56). Zwykle startuje się z różnych wartości  $z_i^{(0)}$  w celu sprawdzenia braku ich wpływu na wyniki końcowe. Innym sposobem inicjowania algorytmu Gibbsa jest wyznaczenie ocen parametrów wektora  $\beta$  za pomocą wstępnej metody, np. estymatora wewnątrz grupowego opisanego w rozdziale 3.2, a wartości dla  $z_i^{(0)}$  otrzymana się ze wzoru (36). Jeden pełny cykl Gibbsa w przypadku modelu CED, oparty na wynikach uzyskanych w  $q$ -tym losowaniu (lub wartościach startowych gdy  $q=0$ ), składa się z następujących kroków (por. KOS [1997]):

$$\begin{aligned} \phi_l^{(q+1)} &\text{ jest losowane z rozkładu o gęstości (54) dla } z=z^{(q)}, \\ (\sigma_v^{-2})^{(q+1)} &\text{ jest losowane z rozkładu o gęstości (48) dla } \beta=\beta^{(q)} \text{ i } z=z^{(q)}, \\ \beta^{(q+1)} &\text{ jest losowane z rozkładu o gęstości (55) dla } \sigma_v^{-2}=(\sigma_v^{-2})^{(q+1)} \text{ i } z=z^{(q)}, \\ z^{(q+1)} &\text{ jest losowane z rozkładu o gęstości (53) dla } \phi_l=\phi_l^{(q+1)}, \sigma_v^{-2}=(\sigma_v^{-2})^{(q+1)} \text{ i } \beta=\beta^{(q+1)}. \end{aligned} \quad (57)$$

Prezentowana technika wyznaczania rozkładów a posteriori i ich charakterystyk okazuje się bardzo skutecznym i prawie automatycznym narzędziem estymacji stochastycznych modeli granicznych. Ze względu na dużą liczbę parametrów inne techniki Monte Carlo, np. z funkcją ważności, nie dają w tym przypadku zadawalających wyników. Wydaje się, iż w dobie gwałtownego wzrostu mocy obliczeniowej komputerów osobistych, metody Monte Carlo mają szansę upowszechnić się i stać się standardowymi narzędziami estymacji zarówno modeli klasycznych jak i bayesowskich.

W następnym rozdziale zostanie przedstawione uogólnienie modelu CED, w którym uwzględnia się wpływ egzogenicznych zmiennych na różnice w poziomie efektywności w poszczególnych firmach (oddziałach banku). Dodatkowo zostaną opisane podstawowe testy bayesowskie, które umożliwią testowanie istotności wpływu tych zmiennych.

#### 4.4. Specyfikacja i testowanie egzogenicznych przyczyn nieefektywności.

W bayesowskim modelu efektów losowych CED rozkłady dla nieefektywności  $z_i$  miały co prawda niezależne rozkłady warunkowe względem wspólnego parametru rozkładu efektywności  $\phi_l$ , lecz przyjęcie hierarchicznej struktury modelu poprzez specyfikację rozkładu dla  $\phi_l$  powoduje, że rozkłady a priori dla efektów indywidualnych  $z_i$  są zależne. W efekcie wnioskowanie o  $z_i$  odbywa się na podstawie całej próby. Zależność a priori efektów

indywidualnych  $z_i$  można uzyskać wprowadzając dodatkowe zmienne egzogeniczne, które wyjaśniałyby potencjalne różnice w poziomie efektywności poszczególnych grup firm.

W podejściu klasycznym egzogeniczne różnice w efektywności firm specyfikuje się poprzez równania regresji, w którym empiryczne reszty otrzymane wcześniej w wyniku estymacji granicznej funkcji kosztu wyjaśnia się poprzez potencjalne zmienne egzogeniczne (zob. Berger i Mester [1993], Cebenoyan, Cooperman, Register i Hudgins [1993], Mester [1993], Berger i De Young [1997], Kraft i Tirtiroglu [1998]). Ze statystycznego punktu widzenia jest to niepoprawne podejście, ponieważ uzyskane reszty otrzymano wcześniej przy założeniu braku tychże egzogenicznych przyczyn nieefektywności. Alternatywnym rozwiązaniem była łączna estymacja modelu składającego się z dwóch równań, tj. granicznej funkcji kosztu i równania regresji dla  $z_i$  specyfikującego egzogeniczne zróżnicowanie poziomu efektywności w firmach. Estymację takiego modelu można dokonać za pomocą MNW zgodnie z ideą zaproponowaną przez Jondrowa, Lovella, Materova i Schimda [1982], lecz napotyka się na problemy z analitycznym wyznaczeniem odpowiednich rozkładów warunkowych i numerycznym znajdowaniem ekstremum funkcji wiarygodności.

Natomiast w przypadku modeli bayesowskich możliwość uwzględnienia potencjalnych różnic w efektywności uzyskuje się poprzez odpowiednią parametryzację średniej  $\lambda_i$  rozkładu prawdopodobieństwa dla efektu indywidualnego  $z_i$ , co jest krokiem zmierzającym do uogólnienia modelu CED. KOS [1997] zaproponowali model, w którym  $v_{it}$  tak jak wcześniej posiada rozkład normalny ze średnią zero i stałą wariancją  $\sigma_v^2$  natomiast  $z_i$  ma rozkład wykładniczy ze średnią  $\lambda_i$ , która może zależeć od kilku (przyjmuje się, że  $m-1$ ) egzogenicznych zmiennych  $w_{ij}$  ( $j=2, \dots, m$ ), wyjaśniających systematyczne różnice w efektywności firm. W szczególności zakłada się

$$\lambda_i = \prod_{j=1}^m \phi_j^{-w_{ij}} \quad (58)$$

gdzie  $\phi_j > 0$  są nieznanymi parametrami, a  $w_{i1}=1$ . Jeżeli  $m > 1$ , rozkład  $z_i$  może być inny dla różnych  $i$ , więc Koop, Osiewalski i Steel ten uogólniony przypadek modelu CED nazwali modelem o zmiennym rozkładzie efektywności (*Varying Efficiency Distribution* - VED). Jeżeli  $m=1$ , wtedy  $\lambda_i = \phi_1^{-1}$ , co oznacza, że wszystkie składniki reprezentujące nieefektywność są niezależnymi zmiennymi o tym samym rozkładzie prawdopodobieństwa i otrzymuje się przypadek CED omówiony we wcześniejszym podrozdziale.

Model VED można także stosować w przypadku, gdy nie zakłada się, że efektywność jest efektem indywidualnym, czyli dopuszcza się, aby rozkład dla  $z_{it}$  był różny po obu indeksach, zwłaszcza gdy  $T$  jest stosunkowo duże. Wykorzystanie takiego modelu w badaniach empirycznych dotyczących przyczyn wzrostu gospodarczego wybranych państw prezentowane jest w pracy KOS [2000]. Natomiast wykorzystywany schemat losowania Gibbsa jest opisany w pracy Osiewalski i Steel [1998]; zob. też Osiewalski i Marzec [1998a, 1998b] oraz Osiewalski, Marzec i Pipień [1999].

W omawianym tutaj modelu podtrzymane są wszystkie założenia co do rozkładów a priori przyjęte w modelu CED, lecz dodatkowo wprowadza się rozkład dla nowych składowych wektora  $\phi$ . W efekcie przyjmuje się następującą strukturę a priori:

$$p(\theta) = p(\sigma_v^{-2})p(\beta)p(\phi) \propto f_G(\sigma_v^{-2} | \frac{n_0}{2}, \frac{s_0}{2})I(\beta \in B) \prod_{j=1}^m f_G(\phi_j | 1, g_j). \quad (59)$$

Dobór wartości dla  $n_0$  i  $s_0$  w celu przyjęcia nikłej wiedzy o parametrze precyzji został omówiony w modelu CED. Dla  $m$  parametrów  $\phi_j$  przyjmuje się niezależne rozkłady wykładnicze z  $g_j=1$  dla  $j>1$  i  $g_1 = -\ln(r^*)$ , gdzie  $r^* \in (0, 1)$  jest stałą zadawaną przez badacza, która w modelu CED ( $m=1$ ) jest interpretowana jako mediana rozkładu a priori efektywności  $r_i = \exp(-z_i)$ . W przypadku modelu VED ( $m>1$ ) rozkład a priori dla  $\phi = (\phi_1, \dots, \phi_m)'$  jest scentrowany wokół rozkładu a priori dla przypadku CED i prawie nieinformacyjny jeśli chodzi o rolę zmiennych  $w_{ij}$ .

Zatem pełny bayesowski model graniczny VED określony jest przez następującą łączną funkcję gęstości, określoną na przestrzeni wszystkich  $TN$  obserwacji  $y_{it}$  - przy danych  $x_{it}$ ,  $w_i = (w_{i1}, \dots, w_{im})$  - oraz wszystkich wielkości nieobserwowalnych, tj.  $N$  wskaźników nieefektywności  $z_i$  i  $K+2+m$  parametrów składających się na wektor  $\theta = (\beta', \sigma_v^{-2}, \phi_1, \dots, \phi_m)'$ :

$$p(y, z, \theta, \phi | X, W) \propto f_G(\sigma_v^{-2} | \frac{n_0}{2}, \frac{s_0}{2})I(\beta \in B) \prod_{j=1}^m f_G(\phi_j | 1, g_j) \prod_{i=1}^N f_G\left(z_i | 1, \prod_{j=1}^m \phi_j^{w_{ij}}\right) \prod_{i=1}^N \prod_{t=1}^T f_N^1(y_{it} | h(x_{it}, \beta) - z_i, \sigma_v^2), \quad (60)$$

gdzie  $W$  jest macierzą  $N \times m$  składająca się z wektorów-wierszy  $w_i$ , które zawierają wartości zmiennych egzogenicznych wyjaśniających potencjalną nieefektywność  $i$ -tej firmy w badanym okresie, przy czym pierwsza kolumna macierzy  $W$  składa się z samych jedynek. Zatem wartość oczekiwaną rozkładu a priori dla  $z_i$  można przedstawić jako:

$$\lambda_i = \exp\left(-\sum_{j=1}^m \gamma_j w_{ij}\right), \text{ gdzie } \gamma_j = \ln(\phi_j).$$

W celu wykorzystania losowania Gibbsa należy podobnie jak w przypadku CED wyznaczyć z pełnego rozkładu łącznego rozkłady warunkowe. Te rozkłady w przypadku parametru precyzji i wektora  $\beta$  mają identyczną postać jak w przypadku CED, tj.:

$$p(\sigma_v^{-2} | y, X, W, z, \beta, \phi) = f_G\left(\sigma_v^{-2} \left| \frac{n_0 + TN}{2}, \frac{1}{2} \left\{ s_0 + \sum_{t,i} (y_{ti} + z_i - h(x_{ti}, \beta))^2 \right\} \right.\right), \quad (61)$$

$$p(\beta | y, X, z, \sigma_v^{-2}, \phi) \propto I(\beta \in B) f_N^{K+1}(\beta | \beta, \sigma_v^2 (X'X)^{-1}), \quad (62)$$

gdzie  $\beta$  jest dane wzorem (56). Natomiast pełny rozkład warunkowy wektora  $z$  ma  $N$ -wymiarowy rozkład ucięty o skalarnej macierzy wariancji i kowariancji - jako iloczyn warunkowych, uciętych i niezależnych rozkładów normalnych:

$$p(z | y, X, W, \theta) = \prod_{i=1}^N \frac{1}{\Phi(-m_i/\sigma_v)} f_N^1(z_i | m_i, \sigma_v^{-2}) I(z_i \geq 0), \quad (63)$$

gdzie:  $m_i = \sum_{t=1}^T [y_{ti} - h(x_{ti}, \beta)] - \sigma_v^2 / \lambda_i$ , a pozostałe oznaczenia są identyczne jak w wzorze (53).

Natomiast warunkowy rozkład a posteriori dla wektora  $\phi$  ma postać:

$$p(\phi | W, z) \propto \exp\left\{-\left(\sum_{j=1}^m g_j \phi_j + \sum_{i=1}^N \left(z_i \prod_{j=1}^m (\phi_j)^{w_{ij}}\right)\right)\right\} \prod_{j=1}^m (\phi_j)^{1 + \sum_{i=1}^N w_{ij} - 1} \quad (64)$$

Powyższy wzór nie jest jądrem jakiegokolwiek znanego  $m$ -wymiarowego rozkładu lub jądrem iloczynów niezależnych jednowymiarowych rozkładów. Przy założeniu  $w_{ij}=1$  dla  $i=1, \dots, N$  rozkład dla dowolnej współrzędnej wektora  $\phi$ , np.  $h$ -tej, ma prostszą postać dogodną do wyprowadzenia rozkładu warunkowego dla  $\phi_h$ :

$$p(\phi_h | W, z, \phi_{(-h)}) \propto f_G\left(\phi_h \left| 1 + \sum_{i=1}^N w_{ij}, g_h \right.\right) \exp\left\{-\phi_h \sum_{i=1}^N z_i \prod_{j>1}^m (\phi_j)^{w_{ij}}\right\}, \quad (65)$$

gdzie  $\phi_{(-h)} = [\phi_1, \dots, \phi_{h-1}, \phi_{h+1}, \dots, \phi_m]$  oznacza wektor  $\phi$  bez  $h$ -tej współrzędnej. Zatem pełny rozkład warunkowy dla  $\phi_h$ , tj. pierwszego elementu wektora  $\phi$ , jest rozkładem gamma, którego jądro ma postać:

(66)

$$p(\phi_1 | W, z, \phi_{(-1)}) \propto \phi_1^{(1+N)-1} \exp \left\{ -\phi_1 \left( g_1 + \sum_{i=1}^N z_i \prod_{j=2}^m (\phi_j)^{w_{ij}} \right) \right\}.$$

Innymi słowy rozkład warunkowy dla  $\phi_l$  jest rozkładem gamma o parametrach  $1+N$  i  $z_1 D_{1l} + z_2 D_{2l} + \dots + z_N D_{Nl}$ , gdzie:

$$D_{ir} = \prod_{\substack{j \\ j \neq r}}^m \phi_j^{w_{ij}}$$

Stosowanie losowania Gibbsa dla niestandardowego rozkładu pozostałych współrzędnych wektora  $\phi$  wymaga stosowania specjalnych technik, tzw. losowania z odrzucaniem lub algorytmu Metropolisa-Hastingsa (Tierney [1994], O'Hagan [1994]). KOS w pracach [1994] i [1995] podjęli się próby stosowania tych metod, lecz zauważyli oni, że dychotomiczny charakter zmiennych egzogenicznych  $w_{ij}$  (dla  $j > 1$ ), wyjaśniających różnice w efektywności, upraszcza postać funkcji określonych wzorami (64) i (65). W tym przypadku rozkład warunkowy  $h$ -tej współrzędnej wektora  $\phi$  względem pozostałych składowych ( $\phi_{(-h)}$ ) jest rozkładem gamma postaci:

$$p(\phi_h | W, z, \phi_{(-h)}) = f_G \left( \phi_h \left| 1 + \sum_{i=1}^N w_{ih}, g_h + \sum_{i=1}^N w_{ih} z_i D_{ih} \right. \right). \quad (67)$$

W ten sposób rozkład warunkowy dla  $\phi_h$  jest prostym rozkładem gamma, więc można stosować algorytm Gibbsa dla wyznaczenia rozkładów a posteriori wszystkich wielkości nieznanymi i nieobserwowalnymi za cenę jedynie częściowej utraty informacji (zwłaszcza, gdy zmienna  $w_{ij}$  dla  $j > 1$  jest zmienną ciągłą). Teoretycznie każdą zmienną ciągłą można bez utraty informacji zastąpić  $N-1$  zmiennymi zero-jedynkowymi. W praktyce jednakże dopuszcza się ujęcie wartości zmiennej ciągłej za pomocą tylko kilku zmiennych dychotomicznych, ponieważ wymaga się, aby łączna liczba zmiennych  $w_{ij}$  była dużo mniejsza niż liczba obiektów (firm). Algorytm Gibbsa w przypadku modeli VED polega na cyklicznym losowaniu z rozkładów warunkowych danych wzorami (61)-(63) i (67), a jeden cykl Gibbsa jest analogiczny jak w przypadku CED - wzór (57) - z tą tylko różnicą, że na początku tego cyklu dokonuje się losowania z odpowiednich rozkładów warunkowych dla składowych wektora  $\phi$ :

$$\phi_l^{(q+1)} \text{ jest losowane z rozkładu o gęstości (67) dla } z = z^{(q)}, \phi_{(-l)} = \phi_{(-l)}^{(q)}, \quad (68)$$

$$\phi_2^{(q+1)} \text{ jest losowane z rozkładu o gęstości (67) dla } z = z^{(q)}, \phi_l = \phi_l^{(q+1)} \text{ i } \phi_j = \phi_j^{(q)} \text{ dla } j > 2,$$

.....

$\phi_m^{(q+1)}$  jest losowane z rozkładu o gęstości (67) dla  $z=z^{(q)}$ ,  $\phi_{(-m)} = \phi_{(-m)}^{(q+1)}$ .

Na gruncie bayesowskim jednym ze sposobów sumaryzacji wiedzy o parametrach (lub wielkościach nieobserwowalnych  $z_i$ ) zawartej w rozkładzie a posteriori są podzbiory o największej gęstości a posteriori (ang. *subset of highest posterior density; H.P.D. subsets*). Bayesowskie podzbiory o największej gęstości a posteriori są odpowiednikami klasycznych przedziałów ufności. O ile te ostatnie często mają w przypadku modeli próbkowych charakter przybliżony (prawdziwy poziom ufności nie jest znany w przypadku małej próby), o tyle podzbiory bayesowskie są dokładne (znane jest prawdopodobieństwo a posteriori, że parametr jest w tym przedziale). Niech  $p(\theta | \text{dane})$  będzie gęstością rozkładu a posteriori, gdzie  $\omega \in \Omega$ . Zbiór  $R_\alpha \subset \Omega$ , jest określany jako podzbiór (lub przedział w przypadku jednowymiarowym, czyli gdy  $\Omega \subset \mathbb{R}$ , gdzie  $\mathbb{R}$  to zbiór liczb rzeczywistych) o największej gęstości a posteriori, odpowiadający prawdopodobieństwu  $1-\alpha$ , jeżeli:

1.  $\Pr\{\omega \in R_\alpha | \text{dane}\} = 1-\alpha$
2. Dla  $\omega_1 \in R_\alpha$  i  $\omega_2 \notin R_\alpha$ ,  $p(\omega_1 | \text{dane}) \geq p(\omega_2 | \text{dane})$ .

Zatem podzbiór H.P.D. jest to taki obszar przestrzeni wektora parametrów  $\omega$ , że przy ustalonym prawdopodobieństwie  $1-\alpha$  przedział H.P.D. jest najmniejszą podprzestrzenią o tej własności, że gęstość prawdopodobieństwa jakiegokolwiek punktu z jej wnętrza nie jest mniejsza od gęstości prawdopodobieństwa punktu spoza tej przestrzeni. Ponadto element  $\omega^*$  przestrzeni  $\Omega$  należy do  $R_\alpha$  przy ustalonym prawdopodobieństwie  $1-\alpha$ , jeżeli dla wektora losowego  $\omega$  zachodzi nierówność  $\Pr\{p(\omega | \text{dane}) \leq p(\omega^* | \text{dane}) | \text{dane}\} \geq \alpha$ , gdzie  $p(\omega | \text{dane})$  jest traktowana jako zmienna losowa (funkcja wektora losowego  $\omega$ ); Box i Tiao [1973]. Podzbiory H.P.D. służą do wnioskowania o wektorze  $\omega$ , a w szczególności służą do badania prawdziwości hipotezy  $\omega = \omega^*$ . W ten sposób można narzucić restrykcje na wektor parametrów lub jego składowe, testować ich prawdziwość w celu ewentualnej redukcji modelu. Ten typ testu opartego na analizie podzbiorów o największych wartościach gęstości a posteriori (przy ustalonej masie prawdopodobieństwa) określa się mianem testu typu H.P.D. bądź testem Lindleya. Nie jest to „standardowy” test bayesowski, gdyż nie wprowadza się w nim prawdopodobieństwa a priori prawdziwości restrykcji i nie oblicza się prawdopodobieństw a posteriori konkurencyjnych modeli (pełnego i po redukcji).

W przypadku modelu VED przedział H.P.D. można wykorzystać do badanie istotności egzogenicznych przyczyn nieefektywności, co sprowadza się do zbadania czy  $\gamma_j^* = 0$  ( $j=2, \dots, m$ ) należy do przedziału o największej gęstości a posteriori, którego

prawdopodobieństwo a posteriori wynosi  $1-\alpha$  przy ustalonym  $\alpha$  (np.  $\alpha=0.05$ ) Niska wartość  $p(\gamma_j^*=0 \mid \text{dane})$  wskazuje, że  $\gamma_j^*=0$  znajduje się w ogonie podzbioru przestrzeni parametrów  $\gamma$ , czyli w podzbiorze mało prawdopodobnym a posteriori. Jeśli natomiast omawiana gęstość a posteriori jest dostatecznie wysoka, to są podstawy przyjąć, że  $\gamma_j=0$ , czyli dokonać stosownej redukcji modelu wyjściowego VED i rozważać warunkowy rozkład a posteriori o gęstości  $p(\gamma_{(-j)} \mid \gamma_j=0, \text{dane})$ .

Test H.P.D. jest trudno wykorzystać (z przyczyn numerycznych) do określenia zasadności redukcji modelu VED do modelu CED poprzez narzucenie  $m-1$  restrykcji postaci  $\gamma_2^* = \dots = \gamma_m^* = 0$ . W praktyce stosuje się więc pewne uogólnienie testu H.P.D., tj. inny nieformalny test będący bayesowskim odpowiednikiem testu F. Opierając się na idei klasycznego testu F rozważa się formę kwadratową postaci (por. Zellner [1971]):

$$\tau = \tau(\gamma^*; y, X, W) = (\gamma^* - E(\gamma^*)) (D^2(\gamma^*))^{-1} (\gamma^* - E(\gamma^*)), \quad (69)$$

gdzie  $\gamma^*$  to wektor-kolumna o wymiarze  $m^* \times 1$  ( $m^* \leq m$ ) testowanych elementów wektora  $\gamma$ . Natomiast  $\tau$  jako funkcja składowych wektora losowego  $\gamma$  jest także zmienną losową, dla której można wyznaczyć jej jednowymiarowy rozkład a posteriori. Jeżeli rozkład  $p(\gamma^* \mid y, X, W)$  byłby  $m^*$ -wymiarowym rozkładem t-Studenta o wektorze wartości oczekiwanych  $E(\gamma^*)$ , macierzy kowariancji  $D^2(\gamma^*)$  i stopniach swobody  $\pi > 2$ , to zmienna losowa  $\tau/m^*$  miałaby rozkład F o parametrach  $m^*$  i  $\pi$ , więc przedstawiona dalej procedura byłaby numerycznie równoważna zastosowaniu klasycznego testu F. Z uwagi na nieznaną postać rozkładu formy  $\tau$ , należy zbadać czy wybrany element z podzbioru wartości wektora  $\gamma$ , tj.  $\gamma^*=0$  (gdzie 0 oznacza wektor zawierający zera o wymiarze  $m^* \times 1$ ), należy do zbioru o największej gęstości a posteriori przy ustalonym prawdopodobieństwie  $1-\alpha$ . Wartości zmiennej  $\tau(\gamma^*; y, X, W)$  odpowiadające  $\gamma^*=0$ , które znajdują się w ogonach rozkładu  $p(\tau \mid y, X, W)$  wskazują, iż ten punkt przestrzeni parametrów nie należy do zbioru o największej gęstości a posteriori. Taka sytuacja oznacza, że ewentualna redukcja modelu VED jest bezzasadna.

W kolejnym rozdziale przedstawione zostanie wykorzystanie bayesowskich stochastycznych modeli graniczne typu VED (ewentualnie CED) do pomiaru efektywności kosztowej oddziałów wybranego polskiego banku komercyjnego.

## 5. Empiryczna analiza efektywności kosztowej oddziałów.

### 5.1. Charakterystyka danych źródłowych.

Zaprezentowane i omówione modele - klasyczny i bayesowski - zostaną wykorzystane w empirycznej analizie efektywności kosztowej oddziałów jednego z polskich banków komercyjnych. Wykorzystane w badaniach dane kwartalne ( $T=4$ ) obejmują rok 1997 i pochodzą z 58 oddziałów banku ( $N=58$ ). Z próby zostały wyeliminowane oddziały nowo powstałe, a filie traktowane są łącznie z oddziałami. Zakres i przekrój danych otrzymanych z centrali banku był w dużej części uzależniony od struktury zakładowego planu kont, wg którego wygenerowano potrzebne informacje przy użyciu istniejącego systemu informatycznego księgowości bankowej. Wykaz kont bankowych, z których pozyskano dane wykorzystywane w niniejszej pracy przedstawia Tabela 1. Sposób ewidencji danych znacząco ograniczył liczbę i rodzaj czynników finansowych, które poprzez ich ceny znalazły się w funkcji kosztu, oraz ilość produktów.

Tabela 1. Wykaz kont księgowych jako źródło danych.<sup>10</sup>

Nazwa konta	Numer konta
<b>ZESPÓŁ 0 – MAJĄTEK TRWAŁY</b>	
Majątek trwały materialny:	021
Tereny i budynki o charakterze eksploatacyjnym	0211
Inwestycje - koszty poniesione w obce środki trwałe w związku z przebudową, rozbudową i adaptacją obiektów na potrzeby ich użytkowników	0212
Wyposażenia, maszyny, narzędzia, środki transportowe inne	0213
<b>ZESPÓŁ 1 OPERACJE Z UDZIAŁEM ŚRODKÓW PIENIĘŻNYCH I OPERACJE MIĘDZYBANKOWE</b>	
Należności i zobowiązania	12-18
<b>ZESPÓŁ 2 OPERACJE Z PODMIOTAMI NIEFINANSOWYMI</b>	
Należności i zobowiązania	21-29
<b>ZESPÓŁ 3 OPERACJE Z JEDNOSTAMI BUDŻETOWYMI I POZA BUDŻETOWYMI</b>	
Należności i zobowiązania	31-34
<b>ZESPÓŁ 5 OPERACJE RÓŻNE</b>	
Wierzyciele różni (podatki, składki ZUS, opłaty PZU, PFRON)	552

<sup>10</sup> wg „Bankowego Planu Kont z Komentarzem i Słownikiem”, pod redakcją J. Urygi, wyd. Interfin, Kraków, 1994.



<b>ZESPÓŁ 7 KONTA KOSZTÓW</b>	
1. Koszty operacji z udziałem środków płynnych oraz operacji międzybankowych:	71
Odsetki i prowizje	711, 712
2. Koszty operacji z klientami:	72
Odsetki i prowizje	721, 722
3. Koszty operacji z jednostkami budżetowymi lub pozabudżetowymi:	73
Odsetki	7311, 7321
Inne koszty	7312, 7322
4. Koszty operacji papierami wartościowymi	74
5. Koszty ogólne	76
Wydatki na personel	761
Koszty eksploatacyjne i związane z działalnością pozaeksploatacyjną <sup>11</sup>	762, 763
Podatki i opłaty	764

W celu estymacji granicznej krótkookresowej funkcji kosztu zmiennego, określonej wzorem (27), należy dokonać pomiaru kosztu zmiennego, cen zmiennych czynników produkcji, wielkości produkcji i zaangażowania czynnika stałego. Wyjaśniany przez model koszt zmienny ( $VC$ ) jest sumą kosztów zaangażowania czynników produkcji: kapitału ludzkiego ( $C_L$ ), finansowego ( $C_D$ ) i kapitału fizycznego z wyłączeniem nieruchomości ( $C_M$ ). Koszty zatrudnienia pracowników w postaci funduszu płac powiększono o wydatki, które ponosi oddział w formie m.in. składek ZUS, PFRON, PFŚP i zaliczek na podatek od osób fizycznych, ponieważ mają one wpływ na rzeczywistą cenę tego czynnika, aczkolwiek nie stanowią kosztu w myśl ustawy o rachunkowości. Szczegółowa specyfikacja elementów kosztów przedstawiona jest poniżej:

Składniki kosztu	Opis
$C_L$	Koszty osobowe i bezosobowe oraz wydatki w postaci składek ZUS, PFRON, PFŚP i zaliczek na podatek od osób fizycznych.
$C_D$	Koszty z tytułu odsetek naliczonych, zapłaconych i prowizji z tytułu pozyskania depozytów od podmiotów finansowych, niefinansowych i budżetowych oraz hipotetyczne koszty pozyskania środków finansujących nadwyżkę kredytów nad depozytami.
$C_M$	Koszt zużycia materiałów i koszty usług obcych, amortyzacja wartości niematerialnych i prawnych oraz środków trwałych z wyłączeniem nieruchomości.

Ceny czynników liczono jako iloraz kosztu zaangażowania czynnika (ewentualnie innych wydatków nie będących jednak kosztem w myśl ustawy o rachunkowości) przez jego wielkość. Cena pracy jest wyznaczana jako iloraz kosztu wynagrodzeń i wydatków związanych z tym czynnikiem ( $C_L$ ) ponoszonych w kwartale i wielkości zaangażowania kapitału ludzkiego mierzonego przez przeciętną liczbę zatrudnionych w kwartale (w

<sup>11</sup> bez kosztów stałych, np. podatków od nieruchomości, środków transportu; ubezpieczeń majątkowych, kosztów usług remontowych i konserwacyjnych, usług komunalnych, zużycia opału, koszty usług w zakresie ochrony mienia itp.

przeliczeniu na etaty; z wyłączeniem przebywających na urloпах macierzyńskich, wychowawczych i bezpłatnych), która jest średnią arytmetyczną ze stanów na koniec trzech miesięcy kwartału. Ceną depozytów i innych pozyskanych środków jest iloraz kosztów finansowych ( $C_D$ ) i wielkości zaangażowania tego czynnika<sup>12</sup>, które wyrażono przez wartość sumy: 1<sup>0</sup> nadwyżki kredytów nad depozytami w przypadku oddziałów kredytowych (środków pozyskanych z centrali) oraz 2<sup>0</sup> średniego stanu środków zgromadzonych na następujących rachunkach<sup>13</sup>:

- bieżących i pomocniczych podmiotów gospodarczych osób prowadzących działalność gospodarczą (udział wartościowo w depozytach ogółem kształtował się na poziomie około 23.5%),
- oszczędnościowo-rozliczeniowych i popularnych osób fizycznych (10%),
- depozytu terminowego imiennego i na okaziciela (48%),
- depozytowych typu „call” (1%),
- zgromadzonych z tytułu sprzedaży bonów oszczędnościowych (0.5%),
- walutowych „A”, „C” (avista lub terminowy) i „K” (a'vista lub terminowy) (12.75%),
- środków funduszy specjalnego przeznaczenia i depozytów zablokowanych (3.25%),
- pozostałych (1%).

Środki na rachunkach loro i lokaty innych banków nie są uwzględniane, ponieważ występowały jedynie w centrali banku. Natomiast wartość uzyskanej produkcji wyrażono w postaci jednego produktu – agregatu - jako wartość udzielonych kredytów:

- w rachunku bieżącym (9.25%),
- obrotowych (33.25%),
- inwestycyjnych (19%),
- konsorcyjnych (7%),
- dyskontowych (3.5%),
- udzielonych firmom leasingowym na finansowanie aktywów będących przedmiotem leasingu finansowego i operacyjnego (4.5%),
- w formie pożyczki gotówkowej dla ludności (m.in. na zakup samochodów) (9%),

---

<sup>12</sup> Istniejący system informatyczny nie pozwalał na uzyskanie średniej ważonej ceny depozytów ogółem w oparciu o rzeczywiste oprocentowanie poszczególnych rachunków depozytowych.

<sup>13</sup> Przy określaniu nazw rachunków depozytowych i rodzajów kredytów posłużono się terminologią używaną w oddziałach badanego banku.

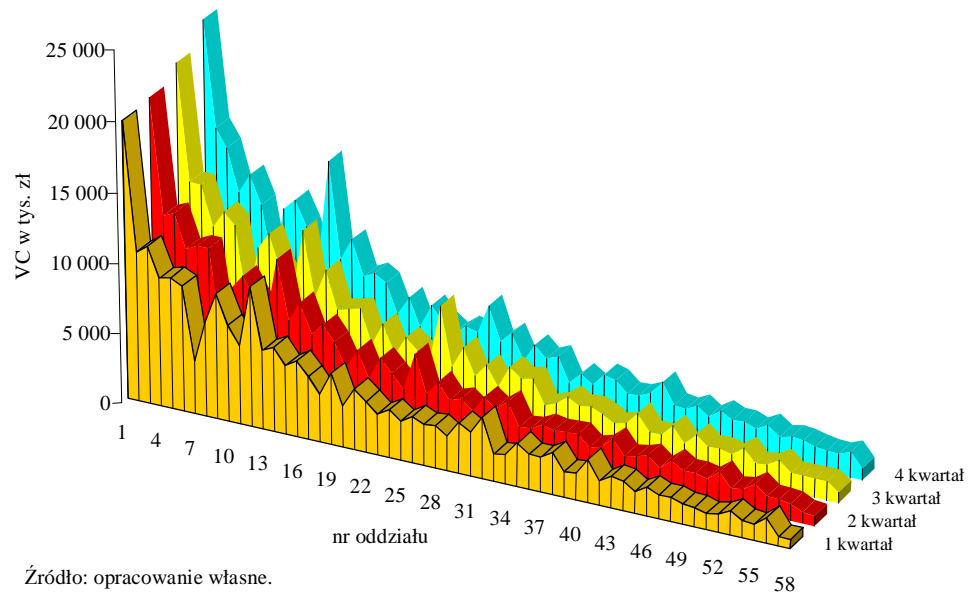
- faktoringu i cesji wierzytelności (6%)
- ratałnych (4.5%)
- pozostałych (4%)<sup>14</sup>.

Zarówno wartość pozyskanych depozytów jak i udzielonych kredytów w badanym kwartale wyrażono przez średnią ich wartość wg stanu na koniec każdego miesiąca w kwartale. Rysunki 5-9 przedstawiają kształtowanie się zmiennej objaśnianej i zmiennych objaśniających granicznej funkcji kosztu zmiennego w badanych kwartałach i w poszczególnych oddziałach, które są uporządkowane malejąco ze względu na średnią arytmetyczną dla wartości logarytmu (czyli średnią geometryczną) udzielonych kredytów w badanych czterech kwartałach. Zwraca uwagę duże zróżnicowanie w zaangażowaniu czynnika stałego przez poszczególne oddziały. Przypuszcza się, że informacje uzyskane z oddziałów, a dotyczące powierzchni własnych i najmowanych budynków, są obarczone dużymi błędami pomiaru, ponieważ wydaje się, że pomiar tych wielkości nie był przeprowadzony w identyczny sposób. Zmienna ta pozostanie jednak w modelu, ponieważ w literaturze zaleca się wyrażenie zaangażowania tego czynnika raczej przez powierzchnię budynków, niż przez np. wartość księgową. Rezultaty wstępnych badań z zakresu efektywności kosztowej oddziałów, w których przyjęto ten drugi sposób pomiaru czynnika, autor prezentuje dla funkcji Cobba-Douglasa i danych przekrojowych w pracy Marzec [1998a, 1998b, 1999].

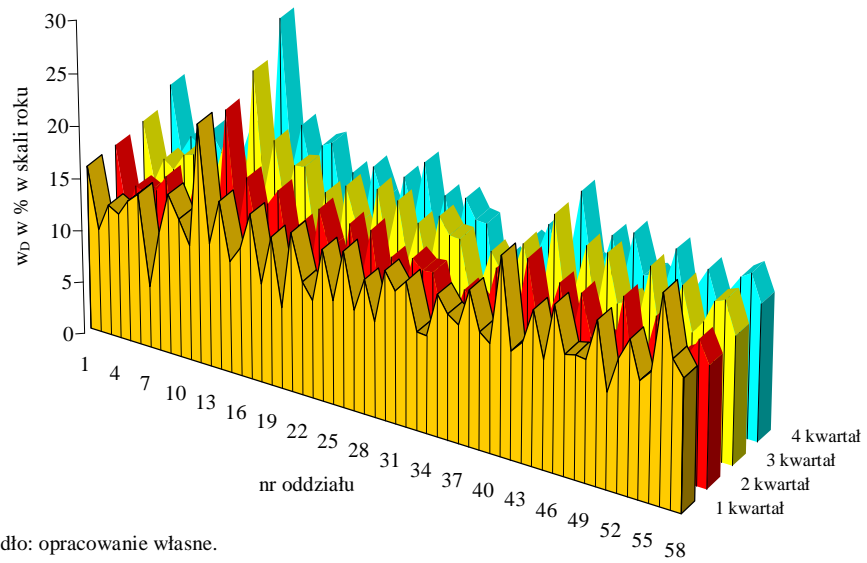
---

<sup>14</sup> z wyłączeniem rachunków nostro i lokat w innych bankach.

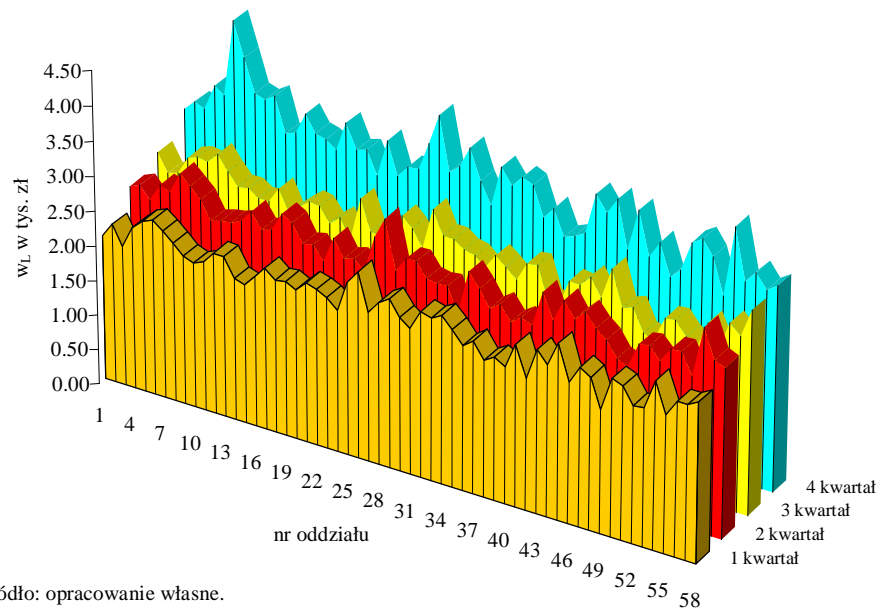
Rysunek 5. Koszty zmienne w poszczególnych kwartałach ( $VC$ ).



Rysunek 6. Cena depozytów i innych pozyskanych środków ( $w_D$ ).

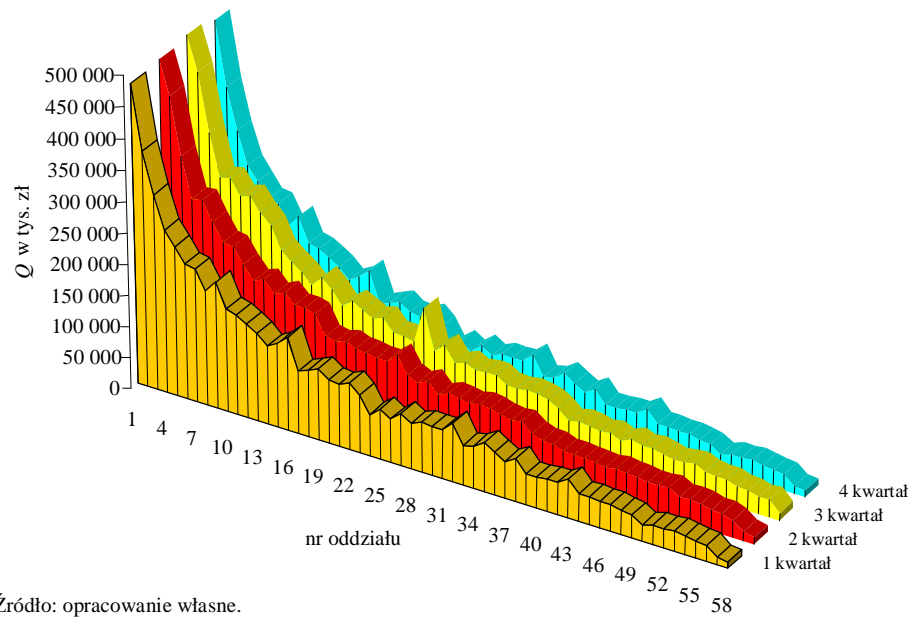


Rysunek 7. Cena pracy ( $w_L$ ) jako średnie miesięczne wynagrodzenie wraz z narzutami w skali kwartału.



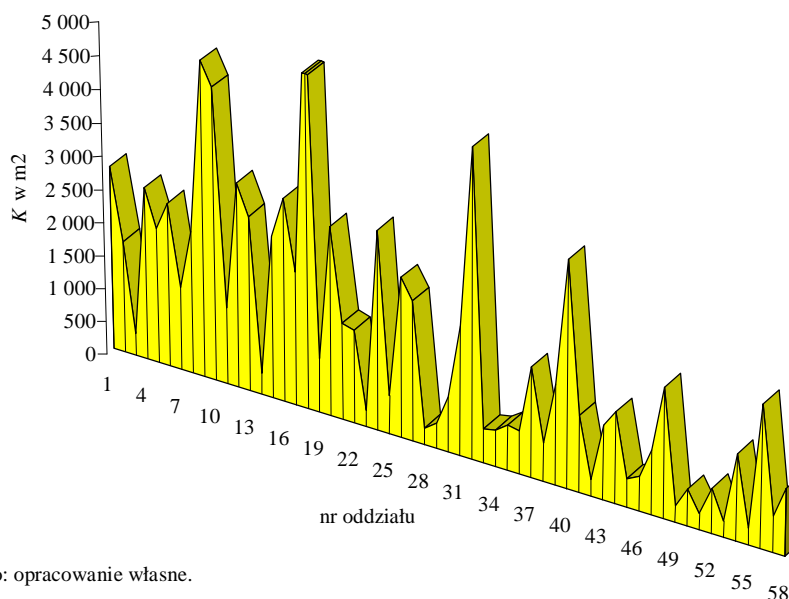
Źródło: opracowanie własne.

Rysunek 8. Kredyty i inne pożyczone środki ( $Q$ ).



Źródło: opracowanie własne.

Rysunek 9. Powierzchnia pomieszczeń biurowych ( $K$ ).



Źródło: opracowanie własne.

Przedstawione powyżej dane zostaną wykorzystane do estymacji klasycznego i bayesowskiego modelu granicznego. Jednakże w celu wyjaśnienia egzogenicznych przyczyn zróżnicowania poziomu efektywności za pomocą modelu VED należy określić potencjalne zmienne dychotomiczne, które wydają się być odpowiedzialne za zróżnicowanie poziomu efektywności w poszczególnych podgrupach oddziałów. Zaproponowano trzy takie zmienne ( $m=4$ ,  $w_{i1} \equiv 1$ ) odzwierciedlające 1<sup>o</sup> specjalizację oddziałów ( $w_{i2}$ ), 2<sup>o</sup> wielkość oddziału mierzona wartością produkcji ( $w_{i3}$ ) i 3<sup>o</sup> posiadanie przez oddziały filii mających szeroki zakres uprawnień ( $w_{i4}$ )<sup>15</sup>. Samodzielność filii jest wyróżniona przez fakt posiadania przez nią jednoznacznie identyfikującego numeru rozliczeniowego oraz pełną ofertę usługową. Zmienna  $w_{i3}$  ma charakter zmiennej ciągłej, ale dokonano jej dychotomizacji wyodrębniając dwie różne grupy oddziałów ze względu na średnią miesięczną wartość udzielonych kredytów. W szczególności:

- $w_{i2} = 1$ , gdy średnia miesięczna wartość kredytów udzielonych przez oddział była większa od wartości pozyskanych depozytów w badanych czterech kwartałach,  $w_{i2} = 0$  w przeciwnym przypadku,
- $w_{i3} = 1$ , gdy średnia miesięczna wartość kredytów oddziału była większa niż 100 mln złotych,

<sup>15</sup> Można też zaproponować inne zmienne, np. atrakcyjność regionu, w którym działa oddział, mierzona wielkością produktu narodowego w przeliczeniu na mieszkańca.

- $w_{i4} = 1$ , gdy oddział posiada filię mającą szeroki zakres podejmowania decyzji (tzn. świadczącą pełny zakres usług bankowych).

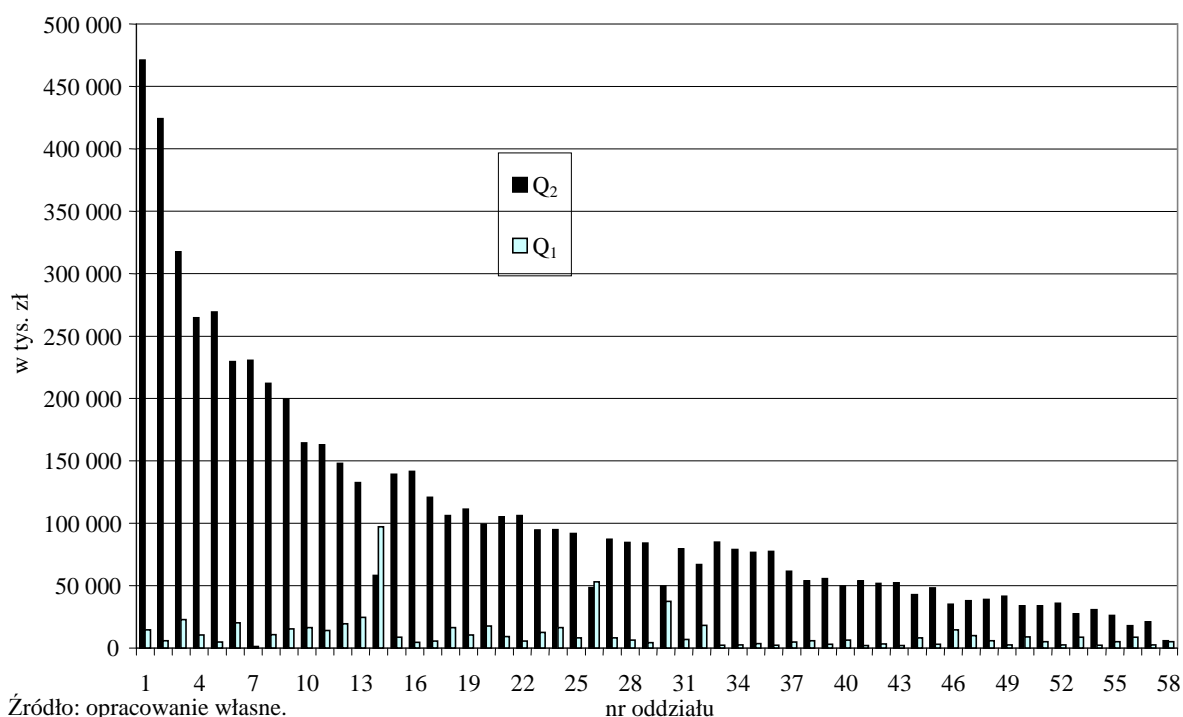
Wartości tych zmiennych dla poszczególnych oddziałów są prezentowane wraz z omawianymi wynikami estymacji mierników efektywności. Podział oddziałów na depozytowe i kredytowe, duże i małe ze względu na zakres prowadzonej działalności oraz ze względu na strukturę organizacyjną pozwoli na identyfikację potencjalnych przyczyn różnic w efektywności, co będzie przedmiotem formalnego testowania.

W celu analizy potencjalnych korzyści pełnego zakresu produkcji i estymacji dwuproduktowej granicznej funkcji kosztu (zob. wzór(28)) dokonano podziału całkowitej produkcji oddziałów ( $Q$ ) na dwa produkty, mianowicie:

- kredyty udzielone klientom komercyjnym (podmiotom finansowym, budżetowym i niefinansowym z wyłączeniem osób fizycznych nie prowadzących działalności gospodarczej) łącznie z nadwyżką depozytów nad kredytami w przypadku oddziałów specjalizujących się pozyskiwaniu środków (zmienna  $Q_1$  wyrażona w tys. złotych),
- kredyty udzielone klientom detalicznym tj. osobom fizycznym z wyłączeniem prowadzących działalność gospodarczą (zmienna  $Q_2$  wyrażona w tys. złotych).

Kształtowanie się średnich wartości tych zmiennych w badanym roku dla poszczególnych oddziałów przedstawia Rysunek 10. Wartość kredytów udzielonych klientom detalicznym była średnio około 10-krotnie mniejsza niż wartość kredytów udzielonych klientom komercyjnym. Jednakże kilka oddziałów, tj. o numerach 14, 26, 30 i 58, specjalizuje się udzielaniu kredytów detalicznych, których udział w kredytach ogółem jest znaczący (w przypadku dwóch pierwszych oddziałów wynosi ponad 50%).

Rysunek 10. Średnia wartość kredytów komercyjnych i detalicznych ( $Q_1$  i  $Q_2$ ).



## 5.2. Estymacja modelu granicznej funkcji kosztu z wykorzystaniem MNK.

W celu estymacji granicznej funkcji kosztu i pomiaru efektywności przy użyciu technik klasycznych, zostanie wykorzystana skorygowana metoda najmniejszych kwadratów dla danych przekrojowych. Wykorzystanie estymatora wewnątrzgrupowego (*within estimator*) w oparciu o dane przekrojowo-czasowe nie jest możliwe, ponieważ jedna ze zmiennych, tj. nakłady czynnika stałego mierzone powierzchnią pomieszczeń biurowych, nie wykazuje zmienności po czasie, co jest wymagane przy stosowaniu tej metody. Ponadto zastosowanie estymatora wewnątrzgrupowego (podobnie jak skorygowanej MNK) nie pozwala na określenie precyzji szacunku pomiaru efektywności. Należy przypomnieć, że stosując estymator wewnątrzgrupowy niejawnie zakłada się nieskończenie duże prawdopodobieństwo a priori, że oddział jest wysoce nieefektywny, a tylko skończone, że jest w pełni efektywny, co pokazali KOS [1997]. Natomiast stosowanie metody największej wiarygodności sprowadziłoby się do numerycznego wyznaczania ekstremum (maksimum) skomplikowanej funkcji wielu zmiennych, omówionej w podrozdziale 3.1.



Dla granicznej funkcji kosztu przyjęto postać zmodyfikowanej funkcji Cobba-Douglasa jako szczególnego przypadku funkcji translogarytmicznej. W przypadku danych przekrojowych przyjęcie translogarytmicznej postaci granicznej funkcji kosztu w celu jej estymacji skorygowaną MNK nie jest zalecane, ponieważ można spodziewać się dużych błędów szacunku otrzymanych ocen parametrów z powodu stosunkowo małej liczby stopni swobody. Ponadto stosowanie MNK dla funkcji translogarytmicznej nie pozwala narzucić skomplikowanych restrykcji na parametry, uwzględniających odpowiednie znaki dla elastyczności, więc można oczekiwać, że uzyskane wyniki będą trudno interpretowalne z ekonomicznego punktu widzenia. Zatem zastosowania uproszczonej formy funkcyjnej dla kosztu granicznego należy traktować jako podejście wstępne, ale wygodne z uwagi na prostotę obliczeń. Uzyskane wyniki stanowiąc będą nieformalny test na jakość danych, charakter procesu produkcji i stałość efektywności w badanym okresie czterech kwartałów. Ponadto wstępne wyniki pozwolą sprawdzić, czy dane liczbowe potwierdzą specyfikację przyjętego modelu mikroekonomicznego opisującego technologię i koszt oddziałów banku komercyjnego, na którym to modelu opiera się dalsza zaawansowana analiza efektywności kosztowej. Można podejrzewać, że otrzymane wstępne wyniki badań będą wskazywać na potrzebę zastosowania ogólniejszych form funkcyjnych dla funkcji kosztu i formalnych metod estymacji modeli granicznych w oparciu w dane przekrojowo-czasowe. Zmodyfikowana funkcja kosztu granicznego typu Cobba-Douglasa w przypadku danych przekrojowych ma postać:

$$\ln VC_i = \beta_0 + \beta_1 \ln w_{i,D} + \beta_2 \ln w_{i,L} + \beta_3 \ln Q_i + \beta_4 \ln K_i + \beta_{13} (\ln Q_i)^2 + z_i + v_i. \quad (70)$$

Forma funkcyjna dana wzorem (70) została użyta m.in. przez Christensena i Greene'a [1976] dla funkcji kosztu przedsiębiorstw branży energetycznej USA w 1970 roku. Postać jej uogólnia funkcję Cobba-Douglasa przez dopuszczenie zmiennego efektu skali, który zależy od wielkości produkcji. Współczynnik efektu skali wyznaczony z równania powyżej wynosi:

$$RTS_i = (\beta_3 + 2\beta_{13} \ln(Q_i))^{-1} \quad (i = 1, \dots, N), \quad (71)$$

i przedstawia zmiany skali produkcji spowodowane przez proporcjonalny wzrost nakładów czynników zmiennych.

Zakładając wstępnie  $z_i \equiv 0$ , na podstawie danych z czterech kwartałów 1997 roku otrzymano następujące oceny MNK parametrów równania (70), błędy średnie szacunku tych ocen, wartości statystyki t-Studenta oraz wartość współczynnika determinacji  $R^2$ :

I kwartał 1997				II kwartał 1997		
Zmienna	ocena parametru	Błąd oceny	statystyka t-Studenta	ocena parametru	Błąd oceny	statystyka t-Studenta
<i>Const</i>	5.221	2.942	1.775	6.313	3.327	1.897
$\ln(w_D * 100)$	0.805	0.034	23.456	0.796	0.036	21.899
$\ln w_L$	-0.010	0.100	-0.101	-0.022	0.101	-0.221
$\ln Q$	-0.055	0.312	-0.175	-0.157	0.347	-0.451
$\ln K$	0.035	0.010	3.646	0.030	0.010	3.085
$(\ln Q)^2$	0.026	0.009	2.997	0.029	0.010	3.005
$R^2=0.994$				$R^2=0.994$		

III kwartał 1997				IV kwartał 1997		
Zmienna	Ocena parametru	Błąd oceny	statystyka t-Studenta	Ocena parametru	Błąd oceny	statystyka t-Studenta
<i>Const</i>	8.037	3.851	2.087	7.669	3.899	1.967
$\ln(w_D * 100)$	0.811	0.036	22.404	0.789	0.041	19.411
$\ln w_L$	0.052	0.105	0.497	0.087	0.095	0.909
$\ln Q$	-0.412	0.394	-1.044	-0.377	0.422	-0.894
$\ln K$	0.028	0.010	2.819	0.037	0.011	3.416
$(\ln Q)^2$	0.036	0.011	3.301	0.034	0.011	2.970
$R^2=0.993$				$R^2=0.992$		

Dopasowanie danych empirycznych do danych teoretycznych dla każdego z okresów jest bardzo wysokie; współczynniki determinacji wynoszą ponad 0.99. Jednocześnie błędy szacunku parametrów są stosunkowo duże, co wskazuje na występowanie współliniowości, do pomiaru której zastosowano miernik uwarunkowania macierzy  $X'X$  (zob. Osiewalski [1992]). Dla macierzy niescentrowanych współczynników korelacji  $R_n$  między wartościami zmiennych objaśniających<sup>16</sup> obliczono wartości własne. Za miarę współliniowości przyjęto pierwiastek ilorazu maksymalnej i minimalnej wartości własnej macierzy  $R_n$ . Dla danych z poszczególnych kwartałów miara ta przyjmuje następujące wartości: 4.94E+06, 5.75E+06, 2670 i 2640, które wskazują - zwłaszcza dla danych z dwóch pierwszych okresów - na wysoki poziom współliniowości<sup>17</sup>. Może ona powodować zarówno zwiększanie błędów średnich szacunku dla pewnych parametrów lub ich kombinacji, jak i ich zmniejszania w stosunku do hipotetycznej sytuacji ortogonalnych zmiennych objaśniających. Zatem otrzymane błędy średnie szacunku należy interpretować z dużą ostrożnością.

<sup>16</sup>  $R_n = W^{-1}(X'X)W^{-1}$ , gdzie  $W$  jest macierzą diagonalną stopnia  $K+1$  zawierającą na przekątnej długości kolumn macierzy  $X$  (wektorów wartości zmiennych objaśniających).

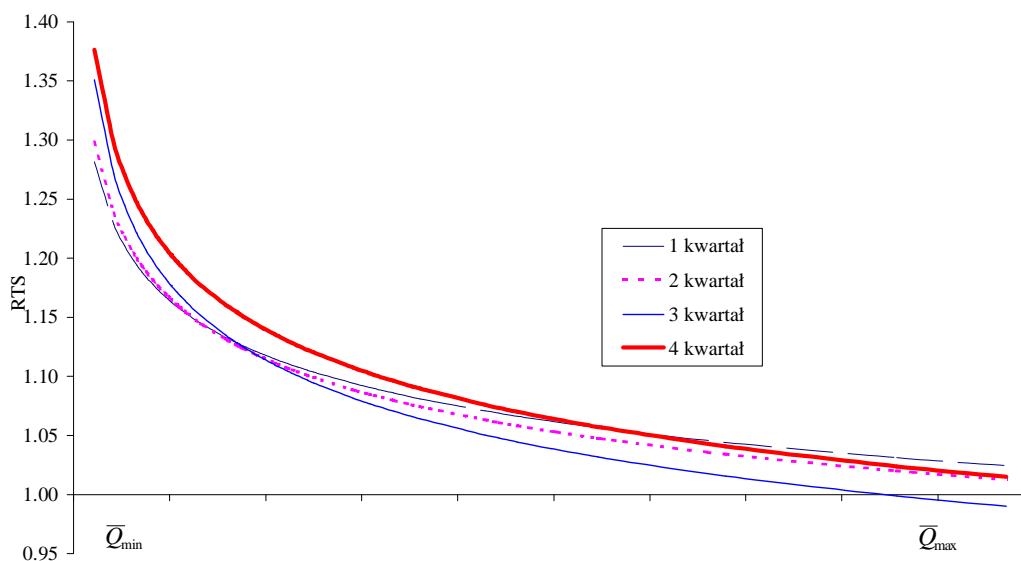
<sup>17</sup> Natomiast dla funkcji Cobba-Douglasa ( $\beta_{13}=0$ ) pierwiastek ilorazu maksymalnej i minimalnej wartości własnej macierzy  $R_n$  dla poszczególnych kwartałów wynosił odpowiednio 316, 303, 309 i 266, co wskazuje na niewielki stopień współliniowości. Zatem przyczyną silnej współliniowości w modelu (70) jest występowanie dwóch zmiennych: logarytmu produkcji i kwadratu tej wielkości.

Oceny parametrów uzyskano stosując MNK i nie nakładając żadnych restrykcji na parametry, zatem warunki regularności ekonomicznej funkcji kosztów nie muszą być spełnione (znak oceny parametru  $\beta_4$  przy  $\ln K$  powinien być ujemny, a  $\beta_2$  dodatni). Wpływ ceny pracy na koszty zmienne wydaje się być w tym modelu dużo mniejszy niż rola ceny depozytów, które w sposób znaczący wpływają na kształtowaniu się kosztu zmiennego. Wzrost tej ceny o 10% (tj. o 1 punkt procentowy przy wyjściowym oprocentowaniu depozytów 10%) powoduje wzrost kosztów o ponad 8%. Istotność parametrów przy kwadracie logarytmu wielkości produkcji wskazuje, że proces produkcji charakteryzuje się zmiennym efektem skali (względem zmiennych czynników produkcji). Sumaryczną informację o charakterze procesu produkcji zawiera tabela poniżej:

Efekt skali	I kwartał 1997	II kwartał 1997	III kwartał 1997	IV kwartał 1997
Sredni (po oddziałach)	1.14	1.13	1.13	1.15
Maksymalny	1.29	1.30	1.35	1.37
Minimalny	1.03	1.01	0.99	1.01

Proces produkcji badanych oddziałów charakteryzuje się współczynnikiem efektu skali (względem zmiennych czynników produkcji), wynoszącym średnio około 1.14. Wzrost kosztów zmiennych o 1%, spowodowany wzrostem nakładów czynników zmiennych o 1%, jest związany ze wzrostem produkcji o około 1.14%. Wskazuje to, przy ustalonych cenach, na możliwość zwiększenia zysku poprzez zwiększenie skali działalności większości oddziałów. Rysunek 11 przedstawia kształtowanie się współczynników efektu skali oddziałów banku w badanych kwartałach, obliczonych w oparciu o średnie wartości logarytmu produkcji ( $\ln Q$ ) w czterech kwartałach, zatem na osi odciętych prezentowane są średnie geometryczne wartości produkcji ( $Q$ ).

Rysunek 11. Współczynnik efektu skali dla zmodyfikowanej funkcji typu Cobba–Douglasa (dla średnich geometrycznych wartości produkcji w badanych kwartałach).



Powyższe charakterystyki uzyskano stosując zwykłą MNK, a więc rozważając przeciętną funkcję kosztów. Najprostszym sposobem uzyskania z funkcji granicznej (i wskaźników efektywności kosztowej) jest korekta wyrazu wolnego z wykorzystaniem najmniejszej reszty MNK. Procedura ta, zwana skorygowaną MNK została szczegółowo przedstawiona w podrozdziale 3.1.

Wskaźniki nieefektywności kosztowej oddziałów obliczone wg wzoru (31) zawarte są w Tabeli 2. Interpretując otrzymane wyniki można stwierdzić, że np. w pierwszym kwartale 1997 roku średnio 11% kosztu zmiennego poniesionego przez wszystkie oddziały banku było kosztem nadwyżkowym, nieuzasadnionym ani wielkością produkcji ani cenami czynników zmiennych czy wielkością zaangażowania czynnika stałego. Warto jednak zauważyć, że otrzymane oszacowania miar efektywności charakteryzują się znacznym, prawdopodobnie nadmiernym zróżnicowaniem, co wynika z uproszczonej metody estymacji funkcji kosztu i pomiaru samej nieefektywności (m.in. zakłada się brak odchyleń czysto losowych:  $v_{it} \equiv 0$ ). O ile ranking oddziałów może być wiarygodny, to poziom efektywności wydaje się być zaniżony. Pomiar efektywności dla funkcji kosztu typu Cobba-Douglasa wskazywał jeszcze większą nieefektywność w przypadku oddziałów najmniej efektywnych, ilustrując wpływ formy funkcyjnej na oceny efektywności (por. KOS [1994]). Rankingi oddziałów w badanych okresach czasu charakteryzują się dużym podobieństwem, co potwierdzają współczynniki korelacji rang Spearmana dla otrzymanych rankingów, które wynoszą między 0.867 a 0.951. Jednakże należy z całą ostrożnością traktować dokonany pomiar efektywności, ponieważ

zastosowana skorygowana MNK nie umożliwia określenia błędów średnich oszacowania tych wielkości, a jej wyniki są wrażliwe na błędy pomiaru kosztu zmiennego lub obserwacje nietypowe.

Tabela 2. Porównanie ocen kosztowej efektywności oddziałów (zmodyfikowana metoda MNK)<sup>18</sup>.

<i>i</i>	I kwartał 1997		II kwartał 1997		III kwartał 1997		IV kwartał 1997	
	$\hat{f}_i$	Ranking	$\hat{f}_i$	Ranking	$\hat{f}_i$	Ranking	$\hat{f}_i$	Ranking
1	0.932	14	0.947	9	0.919	11	0.921	13
2	0.927	16	0.953	6	0.936	4	0.935	6
3	0.841	45	0.855	41	0.841	41	0.831	43
4	0.902	29	0.901	32	0.888	24	0.881	30
5	0.888	33	0.908	22	0.884	30	0.886	28
6	0.899	31	0.902	30	0.898	23	0.910	20
7	1.000	1	1.000	1	0.993	2	0.962	2
8	0.843	43	0.849	43	0.843	40	0.815	45
9	0.910	25	0.908	23	0.886	28	0.888	27
10	0.866	36	0.847	44	0.833	47	0.815	46
11	0.857	38	0.875	35	0.841	42	0.842	39
12	0.844	42	0.856	40	0.835	44	0.838	41
13	0.822	56	0.820	54	0.787	57	0.789	56
14	0.824	54	0.813	55	0.806	54	0.792	55
15	0.830	50	0.841	46	0.820	49	0.802	49
16	0.934	13	0.907	25	0.887	25	0.914	19
17	0.950	4	0.943	13	0.929	6	0.938	5
18	0.908	26	0.898	33	0.886	26	0.898	25
19	0.837	47	0.833	52	0.816	52	0.809	47
20	0.843	44	0.856	39	0.845	39	0.838	40
21	0.943	6	0.936	19	0.915	14	0.948	3
22	0.914	23	0.907	24	0.864	36	0.870	34
23	0.825	53	0.841	47	0.817	50	0.796	53
24	0.824	55	0.872	37	0.873	33	0.801	50
25	0.830	52	0.838	51	0.838	43	0.859	35
26	0.901	30	0.901	31	0.916	13	0.921	14
27	0.830	51	0.812	57	0.814	53	0.799	51
28	0.839	46	0.841	45	0.834	46	0.836	42
29	0.856	39	0.872	36	0.849	38	0.848	38
30	0.848	41	0.839	50	0.832	48	0.822	44
31	0.928	15	0.938	16	0.907	18	0.919	16
32	0.849	40	0.841	48	0.798	56	0.767	58
33	0.920	20	0.962	3	0.907	19	0.938	4
34	0.907	27	0.913	21	0.903	22	0.910	21
35	0.950	5	0.947	8	0.927	7	0.903	22
36	0.938	10	0.957	4	0.919	10	0.928	10
37	0.821	57	0.812	56	0.801	55	0.794	54
38	0.943	7	0.941	15	0.905	21	0.916	18
39	0.938	9	0.946	11	0.916	12	0.924	11
40	0.865	37	0.853	42	0.851	37	0.857	36
41	0.891	32	0.866	38	0.864	35	0.872	32
42	0.918	22	0.881	34	0.870	34	0.851	37
43	0.936	11	0.906	26	0.885	29	0.883	29

<sup>18</sup> W pracy przyjęto, iż badane oddziały są uporządkowane za względu na malejącą średnią wartość logarytmu udzielonych kredytów w badanych czterech kwartałach.

<i>i</i>	I kwartał 1997		II kwartał 1997		III kwartał 1997		IV kwartał 1997	
	$\hat{r}_i$	Ranking	$\hat{r}_i$	Ranking	$\hat{r}_i$	Ranking	$\hat{r}_i$	Ranking
44	0.935	12	0.941	14	0.920	9	0.929	9
45	0.920	19	0.936	18	0.913	17	0.919	15
46	0.942	8	0.916	20	0.914	15	0.924	12
47	0.922	18	0.937	17	0.906	20	0.900	23
48	0.870	35	0.904	28	0.875	32	0.891	26
49	0.911	24	0.906	27	0.878	31	0.870	33
50	0.919	21	0.947	7	0.931	5	0.929	8
51	0.830	49	0.840	49	0.834	45	0.798	52
52	0.958	3	0.946	10	0.922	8	0.899	24
53	0.831	48	0.821	53	0.817	51	0.808	48
54	0.907	28	0.902	29	0.886	27	0.874	31
55	0.923	17	0.943	12	0.914	16	0.918	17
56	0.887	34	0.954	5	0.981	3	0.935	7
57	0.961	2	0.965	2	1.000	1	1.000	1
58	0.766	58	0.757	58	0.734	58	0.774	57
<b>Minimalne <math>\hat{r}_i</math></b>	<b>0.766</b>		<b>0.757</b>		<b>0.734</b>		<b>0.767</b>	
<b>Średnie <math>\hat{r}_i</math></b>	<b>0.889</b>		<b>0.892</b>		<b>0.876</b>		<b>0.873</b>	

W podsumowaniu można stwierdzić, że efektywność oddziałów w badanym okresie nie zmienia się w sposób systematyczny. Otrzymane oceny parametrów pozwalają założyć, że w badanym okresie charakter procesu produkcji oddziałów banku nie zmienia się w czasie i wskazują, że do jego opisu należy używać form funkcyjnych dopuszczających przynajmniej zmienność efektu skali. Ewentualne różnice ocen parametrów technologii w badanych kwartałach są spowodowane dużymi błędami średnimi szacunku, które wynikają z faktu estymacji modelu tylko dla danych przekrojowych, a także ze współliniowości.

Uzyskane wyniki dają podstawę do budowy modelu dla danych przekrojowo-czasowych, w którym można by przyjąć założenie o stałości technologii i stałości efektywności po czasie (wskaźnik nieefektywności  $z_{it}$  traktowany jako efekt indywidualny;  $z_{it} = z_i$ ), co pozwoli na precyzyjny szacunek parametrów technologii i przeciętnej efektywności każdego z oddziałów. Kolejnym krokiem będzie przyjęcie ogólniejszej formy funkcji kosztu, dopuszczającej nie tylko zmienność efektu skali, ale także zmienność elastyczności kosztu względem cen czynników zmiennych i nakładu czynnika stałego.

### 5.3. Estymacja bayesowskiego modelu losowych efektów indywidualnych (o zmiennym rozkładzie efektywność - VED).

Estymacja stochastycznego modelu granicznego i pomiar efektywności w przypadku danych przekrojowo-czasowych zostanie dokonany w oparciu o bayesowski model o zmiennym rozkładzie efektywność (VED), dla translogarytmicznej funkcji kosztu zmiennego postaci:

$$\begin{aligned} \ln VC_{it} = & \beta_0 + \beta_1 \ln w_{it,D} + \beta_2 \ln w_{it,L} + \beta_3 \ln Q_{it} + \beta_4 \ln K_i + \beta_5 \ln w_{it,D} \ln w_{it,L} \\ & + \beta_6 \ln w_{it,D} \ln Q_{it} + \beta_7 \ln w_{it,D} \ln K_i + \beta_8 \ln w_{it,L} \ln Q_{it} + \beta_9 \ln w_{it,L} \ln K_i \\ & + \beta_{10} \ln Q_{it} \ln K_i + \beta_{11} (\ln w_{it,D})^2 + \beta_{12} (\ln w_{it,L})^2 + \beta_{13} (\ln Q_{it})^2 + \beta_{14} (\ln K_i)^2 \\ & + v_{it} + z_i, \end{aligned} \quad (72)$$

wprowadzonej i omówionej w podrozdziale 2.3.

Zmienne egzogeniczne, które wyjaśniają potencjalne systematyczne różnice w poziomie efektywności oddziałów banku to specjalizacja na depozytowe i kredytowe ( $w_{i2}$ ), wielkość oddziału mierzona wartością udzielonych kredytów ( $w_{i3}$ ) oraz struktura organizacyjna ( $w_{i4}$ ), co zostało szerzej omówione w rozdziale 5.1. Model VED jest naturalnym uogólnieniem modelu CED, zatem istnieje możliwość statystycznego testowania istotności zmiennych  $w_{ij}$  ( $j=2,3,4$ ). W przypadku szczególnym - braku istotnego wpływu tych zmiennych - redukcja modelu VED doprowadziłaby do otrzymania modelu CED. W celu zastosowania próbkowania Gibbsa, jako metody wyznaczania charakterystyk rozkładów a posteriori omówionej w podrozdziale 4.3, należy w pierwszej kolejności ustalić parametry rozkładów a priori dla  $\sigma_v^{-2}$  i  $\phi$  oraz określić funkcję wskaźnikową  $I(\beta)$ , która specyfikuje warunki regularności ekonomicznej, a także podać wartości startowe dla  $z_i$  i  $\beta$ . Oczywiście uzyskane wyniki końcowe, w postaci rozkładów a posteriori dla parametrów i wielkości nieobserwowalnych  $z_i$  oraz odpowiednich charakterystyk tych rozkładów, mogą być wrażliwe na wartości parametrów rozkładów a priori i wartości startowe, stąd istnieje konieczność przeprowadzenia analizy wrażliwości. Zasadniczym problemem jest badanie zbieżności algorytmu Gibbsa, którą w teorii uzyskuje się po wykonaniu odpowiednio dużej liczby cykli – losowań Gibbsa. Można spodziewać się, że stabilizacja w kolejnych cyklach wartości oczekiwanych i odchyłeń standardowych a posteriori wybranych wielkości oznacza osiągnięcie zbieżności tej metody.

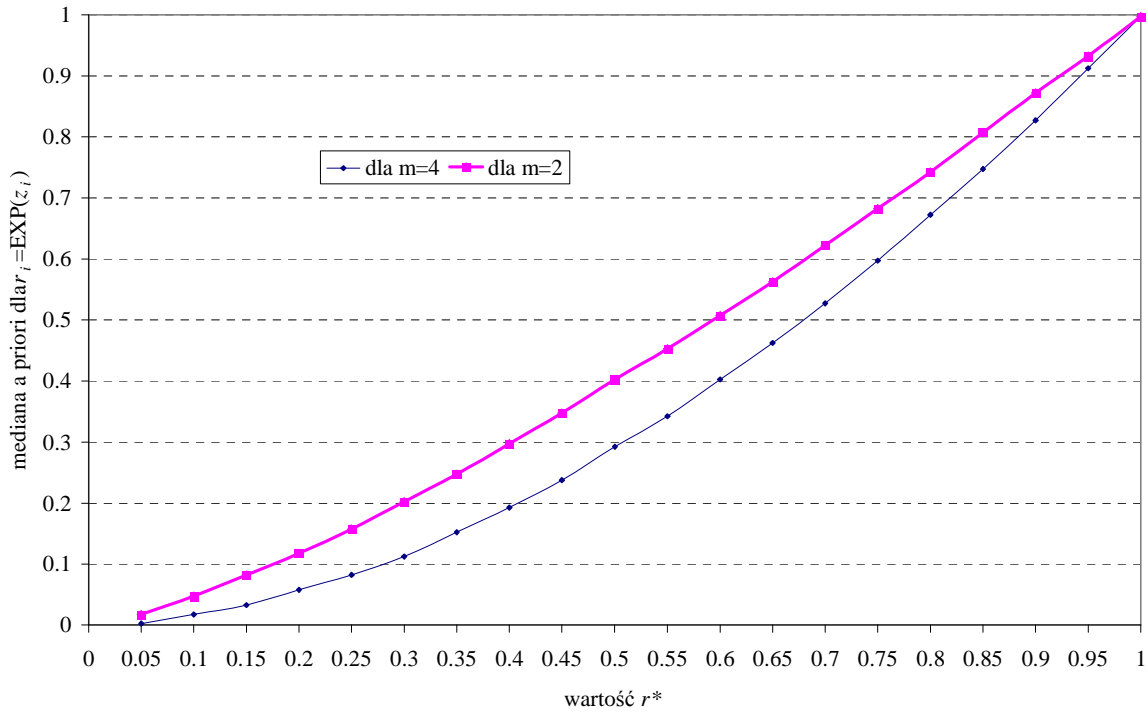
W przypadku granicznego bayesowskiego modelu efektów losowych VED przyjęto następujące wartości parametrów dla rozkładów a priori:

- dla parametru precyzji  $\sigma_v^{-2}$  przyjęto  $n_0=s_0=10^{-6}$  co prowadzi do prawie nieinformacyjnego rozkładu gamma (ze średnią 1 i wariancją  $2 \cdot 10^6$ ),
- dla składowych wektora  $\phi=(\phi_1, \dots, \phi_4)'$  przyjęto rozkłady wykładnicze o parametrach  $g_j=1$  dla  $j>1$  ( $j=1, \dots, 4$ ) i  $g_1=-\ln(r^*)$ , gdzie  $r^*=0.7$  bądź  $0.9$ .

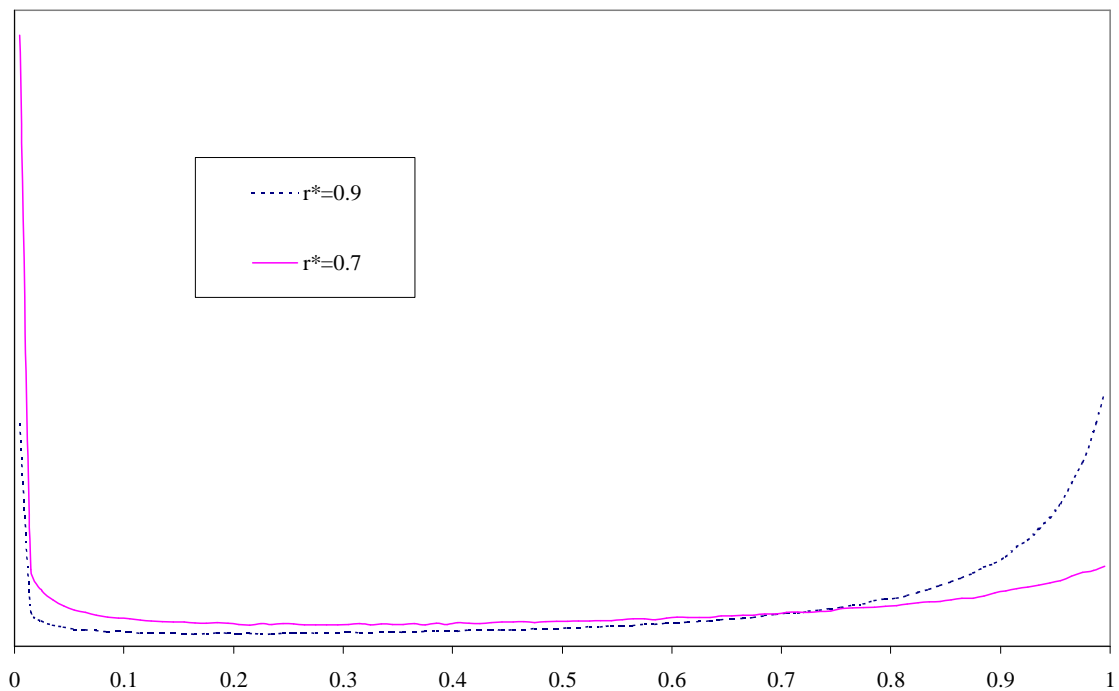
KOS [1997] zauważyli, że w przypadku modelu VED ( $m>1$ ) z nierówności Jensena wynika, że  $P(r_i \leq r^*) > 0.5$ , co oznacza, że mediana rozkładu a priori dla efektywności wynosi mniej niż  $r^*$ . Rysunek 12 przedstawia przybliżoną zależność między medianą rozkładu a priori dla efektywności a parametrem  $r^*$  dla dwóch przypadków VED, tj. dla  $m=4$  i  $m=2$ . Wyniki te uzyskano na drodze symulacji przyjmując średnie wartości po oddziałach zmiennych  $w_{ij}$ . Warto zauważyć, że im więcej występuje w modelu zmiennych wyjaśniających potencjalne, systematyczne różnice w poziomie efektywności, tym mediana rozkładu a priori przyjmuje niższe wartości. W omawianym modelu VED ( $m=4$ ) dla  $r^*=0.7$  i  $r^*=0.9$  wynosi ona odpowiednio 0.528 i 0.828. Zatem przyjęcie za  $r^*$  wartości niższych niż 0.7 nie wydaje się uzasadnione z ekonomicznego punktu widzenia, aczkolwiek pozwoli to na zbadanie wrażliwości wyników a posteriori na założenia a priori dotyczące wektora  $\phi$ . Gęstości rozkładu a priori jako funkcje parametru  $r^*$  tworzą rodzinę rozkładów tzw. u-kształtnych (mają kształt litery „U”; por. KOS [1997]), co przedstawia Rysunek 13. Prawdopodobieństwo a priori, że oddział jest w pełni nieefektywny ( $r$  bliskie 0) jest stosunkowo duże (zwłaszcza im mniejsze jest  $r^*$ ), lecz zawsze skończone, w przeciwieństwie do klasycznego estymatora wewnątrzgrupowego, dla którego odpowiednik tego prawdopodobieństwa jest nieskończenie duże.



Rysunek 12. Mediana brzegowego rozkładu a priori dla efektywności  $r_i$  jako funkcja parametru  $r^*$  (dla średnich wartości zmiennych  $w_{ij}$ ) – VED ( $m=2$  i  $m=4$ ).



Rysunek 13. Brzegowy rozkład a priori dla efektywności  $r_i$  (dla średnich wartości zmiennych  $w_{ij}$ ) – VED ( $m=4$ ).



Mikroekonomiczna teoria produkcji i kosztu precyzuje warunki regularności ekonomicznej, które w badanym modelu zostały narzucone przez funkcję wskaźnikową  $I(\beta)$  w rozkładzie a priori wektora parametrów  $\beta=[\beta_0, \dots, \beta_{14}]$ , zapewniając tym samym odpowiednie znaki dla elastyczności. W przypadku funkcji translogarytmicznej danej wzorem

(72) liczba wszystkich możliwych restrykcji na elastyczności wynosi 1160. Jednakże z uwagi na fakt, że funkcja translogarytmiczna jest jedynie lokalną aproksymacją dowolnej funkcji (wokół arbitralnie przyjętego punktu, który w praktyce można utożsamiać z „punktem ciężkości” zbioru danych), sensowne wydaje się narzucić te restrykcje tylko dla „przeciętnej” firmy (tzn. takiej, która charakteryzuje się średnimi z próby - w czasie i po oddziałach - wartościami zmiennych objaśniających) albo dla średnich wartości elastyczności po oddziałach. Oba sposoby narzucenia tych restrykcji są równoważne; w pracy przyjęto drugie podejście, zatem funkcja wskaźnikowa przyjęła następującą postać:

$$I(\beta) = 1, \text{ gdy } \begin{cases} \eta(VC/w_D) = \frac{1}{58} \sum_{i=1}^{58} \eta(VC/w_D)_i > 0 \\ \eta(VC/w_L) = \frac{1}{58} \sum_{i=1}^{58} \eta(VC/w_L)_i > 0 \\ \eta(VC/Q) = \frac{1}{58} \sum_{i=1}^{58} \eta(VC/Q)_i > 0 \\ \eta(VC/w_D) + \eta(VC/w_L) < 1 \end{cases} \quad (73)$$

$I(\beta) = 0$  w pozostałych przypadkach,

gdzie:

$$\begin{aligned} \eta(VC/w_D)_i &= \beta_1 + \beta_5 \overline{\ln w_{i,L}} + \beta_6 \overline{\ln Q_i} + \beta_7 \overline{\ln K_i} + 2\beta_{11} \overline{\ln w_{i,D}} \\ \eta(VC/w_L)_i &= \beta_2 + \beta_5 \overline{\ln w_{i,L}} + \beta_8 \overline{\ln Q_i} + \beta_9 \overline{\ln K_i} + 2\beta_{12} \overline{\ln w_{i,D}} \\ \eta(VC/Q)_i &= \beta_3 + \beta_6 \overline{\ln w_{i,D}} + \beta_8 \overline{\ln Q_i} + \beta_{10} \overline{\ln K_i} + 2\beta_{13} \overline{\ln Q_i} \\ \overline{\ln w_{i,D}} &= \frac{1}{4} \sum_{t=1}^4 \ln w_{it,D}, \overline{\ln w_{i,L}} = \frac{1}{4} \sum_{t=1}^4 \ln w_{it,L}, \overline{\ln Q_i} = \frac{1}{4} \sum_{t=1}^4 \ln Q_{it}, \overline{\ln K_i} = \ln K_i. \end{aligned} \quad (74)$$

W praktyce powyższe restrykcje nie okazały się wiążące oprócz tej dotyczącej sumy elastyczności kosztu zmiennego względem cen czynników zmiennych. Podjęto próbę narzucenia restrykcji danych wzorem (74) na elastyczności każdego z oddziałów. Okazało się, że dla większości oddziałów restrykcja  $\eta(VC/w_D)_i + \eta(VC/w_L)_i < 1$ , która jest konsekwencją jednorodności funkcji kosztu względem wszystkich cen, nie była wiążąca. Jednakże w przypadku kilku innych otrzymana suma wartości oczekiwanych a posteriori wynosiła nieznacznie więcej niż jeden, lecz suma elastyczności charakteryzowała się małym odchyleniem standardowym a posteriori. Zatem dla danych przekrojowo-czasowych próba utrzymanie warunku na sumę elastyczności kosztu względem cen czynników dla każdego oddziału powodowała duży spadek efektywność numerycznej w losowaniu Gibbsa (około dziesięciokrotny). Warto wspomnieć, że w przypadku danych przekrojowych narzucenie tych

warunków było łatwe i w żaden znaczący sposób nie miało wpływu na efektywność losowania Gibbsa.

W rozkładzie a priori dla wektora  $\beta$  nie uwzględniono restrykcji na elastyczność kosztu zmiennego względem nakładów czynnika stałego, ponieważ formalnie nie udowodniono, że w przypadku translogarytmicznej krótkookresowej funkcji kosztu zmiennego elastyczność kosztu względem czynnika stałego ma znak ujemny (co można pokazać dla funkcji kosztu typu Cobba-Douglasa). Ponadto posiadane dane empiryczne wyraźnie wskazywały na dodatni znak tej elastyczności. Próba narzucenia tych restrykcji spowodowała drastyczny spadek numerycznej efektywności losowania. W przypadku danych przekrojowych uwzględnienie tej restrykcji powodowało, że masa prawdopodobieństwa gęstości rozkładu a posteriori dla tej elastyczności skupiona była na lewo od zera, wartości oczekiwana a posteriori była bliska zero i odchylenie standardowe było stosunkowo duże. Zatem także z obawy, że rozkład a posteriori tej charakterystyki silnie zależałby od rozkładu a priori, omawiana restrykcja nie została uwzględniona we wzorze (73).

W celu uruchomienia algorytmu Gibbsa należało przyjąć wartości startowe dla nieobserwowalnych zmiennych  $z_i$  i wektora parametrów  $\beta$ . W tym celu przyjęto  $z_i^{(0)}=0.3$  dla  $i=1,\dots,58$ , co odpowiada przyjęciu, że wszystkie firmy charakteryzują się efektywnością na poziomie około 0.75. Można byłoby założyć inne punkty startowe, lecz zwykle dla osiągnięcia szybkiej zbieżności algorytmu Gibbsa przyjmuje się wartości, które wydają się być realne, np. na poziomie ocen uzyskanych przy pomocy skorygowanej MNK, która zaniża poziom efektywności.

Przyjmując wartości startowe  $z_i^{(0)}$  uruchomiono algorytm Gibbsa, wykonując w pierwszej kolejności  $S=10^5$  cykli wstępnych (próbnych lub spalonych) w celu osiągnięcia zbieżności algorytmu, czyli zbieżności mechanizmu losowania wektora  $\theta^{(q)}$  do rozkładu stacjonarnego, czyli rozkładu a posteriori  $p(\theta|y)$ . Następnie wykonano  $M=5\cdot 10^5$  cykli Gibbsa w celu otrzymania zadowalającej aproksymacji obliczanych charakterystyk a posteriori<sup>19</sup>. Wyniki uzyskane po  $S$  cyklach wstępnych, np. dla parametrów technologii, były prawie identyczne z tymi uzyskanymi po  $M$  cyklach, co potwierdza zbieżność algorytmu. Zachowanie się (zbieżność) estymatorów opartych na losowaniu Gibbsa w zależności od liczby cykli badano śledząc zachowanie się wybranych charakterystyk a posteriori

---

<sup>19</sup> W omawianym przypadku komputer klasy Pentium II 400 Mhz na wykonanie  $6\cdot 10^5$  cykli Gibbsa potrzebuje 8.5 godziny.

(stabilizacja wyników praktycznie oznacza zbieżność). W literaturze przedmiotu jest wiele propozycji formalnych metod testowania zbieżności, lecz z uwagi na zakres i charakter tej pracy nie są one wykorzystywane. Rysunek 14 i Rysunek 15 prezentują średnią (po oddziałach) z wartości oczekiwanych i odchyłeń standardowych dla efektywności oddziałów:

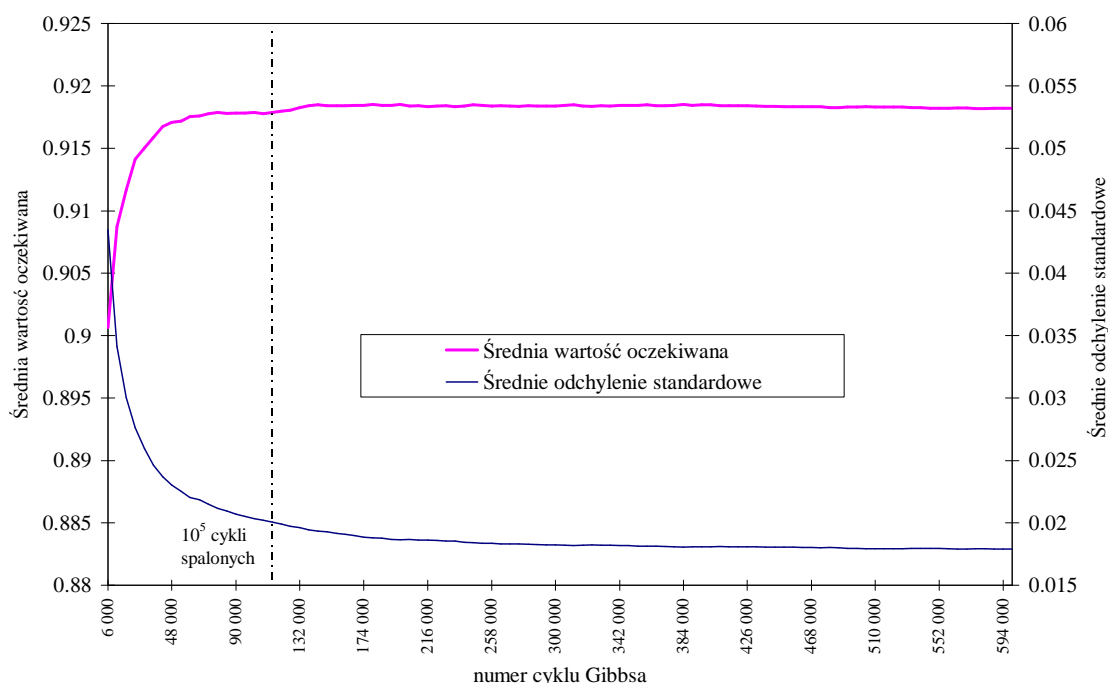
$$\bar{E}(r|y, X) = \frac{1}{58} \sum_{i=1}^{58} E(r_i|y, X), \quad \bar{D}(r|y, X) = \frac{1}{58} \sum_{i=1}^{58} D(r_i|y, X),$$

oraz dla elastyczności kosztu względem wielkości produkcji:

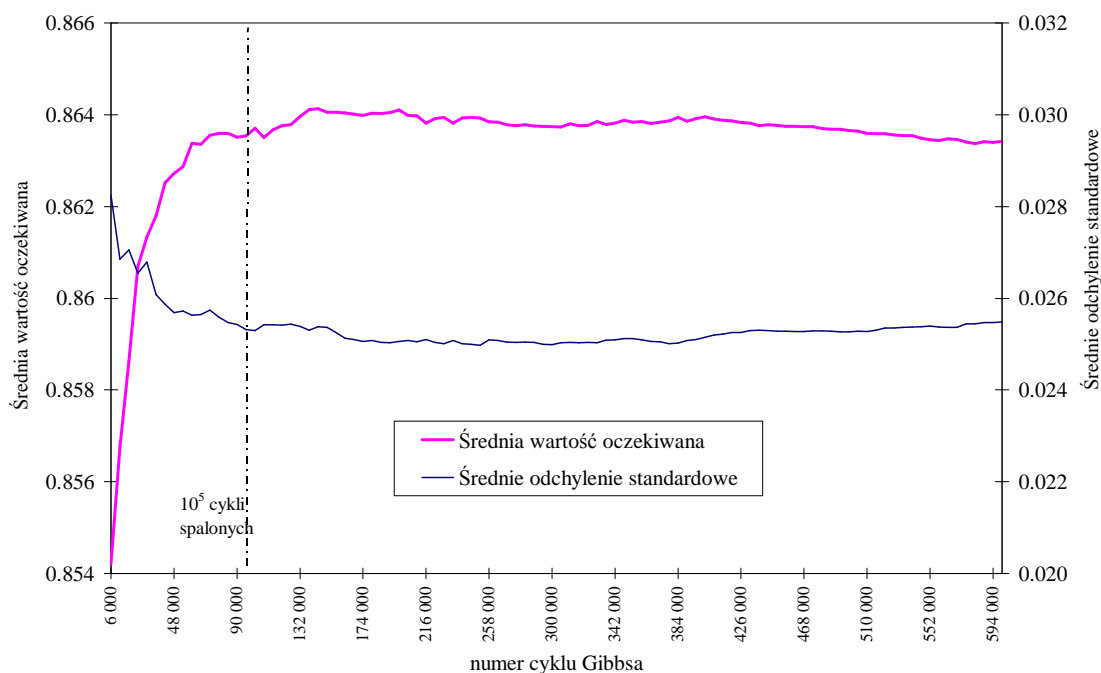
$$\bar{E}(\eta(VC/Q)|y, X) = \frac{1}{58} \sum_{i=1}^{58} E(\eta(VC/Q)_i|y, X), \quad \bar{D}(\eta(VC/Q)|y, X) = \frac{1}{58} \sum_{i=1}^{58} D(\eta(VC/Q)_i|y, X).$$

Szczególną uwagę zwraca prawidłowość, że w miarę zwiększania liczby cykli Gibbsa w ramach cykli spalonych, aproksymacja wartości oczekiwanej a posteriori rośnie, a odchylenie standardowe a posteriori maleje. Podsumowując można stwierdzić, iż w obu przypadkach estymatory oparte o losowanie Gibbsa (dla czterech badanych wielkości) osiągnęły zbieżność już po około 100 tysiącach cykli. Przyjęcie innych wartości dla parametru  $r^*$ , tj. 0.5 i 0.9, jak i nie uwzględnienie w modelu restrykcji na parametry danych równaniem (73), nie ma wpływu na szybkość zbieżności algorytmu.

Rysunek 14. Średnia wartość oczekiwana i średnie odchylenie standardowe efektywności ( $r$ )- VED ( $m=4$ ,  $r^*=0.7$ ).



Rysunek 15. Średnia wartość oczekiwana i średnie odchylenie standardowe elastyczności kosztu względem produkcji (Q) - VED ( $m=4, r^*=0.7$ ).



Po sprawdzeniu zbieżności algorytmu Gibbsa można przystąpić do prezentacji i interpretacji otrzymanych wyników, które będą prezentowane w całości dla  $r^*=0.7$  i częściowo dla  $r^*=0.9$ . Tabela 3 prezentuje wartości oczekiwane i odchylenia standardowe a posteriori parametrów granicznej funkcji kosztów danej równaniem (72) w przypadku modelu (60):

Tabela 3. Wartości oczekiwane i odchylenia standardowe a posteriori parametrów równania - VED ( $m=4, r^*=0.7$ ).

Parametr	Zmienna	$E(\cdot y, X)$	$D(\cdot y, X)$
$\beta_0$	Stała	1.619	3.017
$\beta_1$	$\ln w_D$	0.988	0.455
$\beta_2$	$\ln w_L$	-0.499	0.584
$\beta_3$	$\ln Q$	0.625	0.220
$\beta_4$	$\ln K$	-0.399	0.177
$\beta_5$	$\ln w_D \ln w_L$	-0.014	0.043
$\beta_6$	$\ln w_D \ln Q$	-0.014	0.025
$\beta_7$	$\ln w_D \ln K$	0.054	0.023
$\beta_8$	$\ln w_L \ln Q$	-0.040	0.015
$\beta_9$	$\ln w_L \ln K$	0.001	0.012
$\beta_{10}$	$\ln Q \ln K$	-0.012	0.011
$\beta_{11}$	$(\ln w_D)^2$	-0.039	0.036
$\beta_{12}$	$(\ln w_L)^2$	0.091	0.040
$\beta_{13}$	$(\ln Q)^2$	0.018	0.008
$\beta_{14}$	$(\ln K)^2$	0.036	0.009
$D^2(v_{it}) = \sigma_v^2$		0.000278	0.000037

Konstrukcja funkcji translogarytmicznej stwarza możliwość występowania silnej współliniowości zmiennych objaśniających, co potwierdzały uzyskane wyniki empiryczne. Miara współliniowości, omówiona w rozdziale 5.2 na stronie 88, w przypadku danych przekrojowo-czasowych wynosi 18700 i wskazuje na jeszcze wyższą współliniowość niż w przypadku zmodyfikowanej funkcji Cobba-Douglasa<sup>20</sup>. Jednakże mimo silnego wpływu skorelowania zmiennych objaśniających, interesujące charakterystyki procesu produkcji prezentowane poniżej, tzn. elastyczności kosztu zmiennego względem cen czynników zmiennych, wielkości produkcji i nakładu czynnika stałego, charakteryzują się dużą precyzją szacunku, w przeciwieństwie do parametrów równania translogarytmicznej funkcji kosztu.

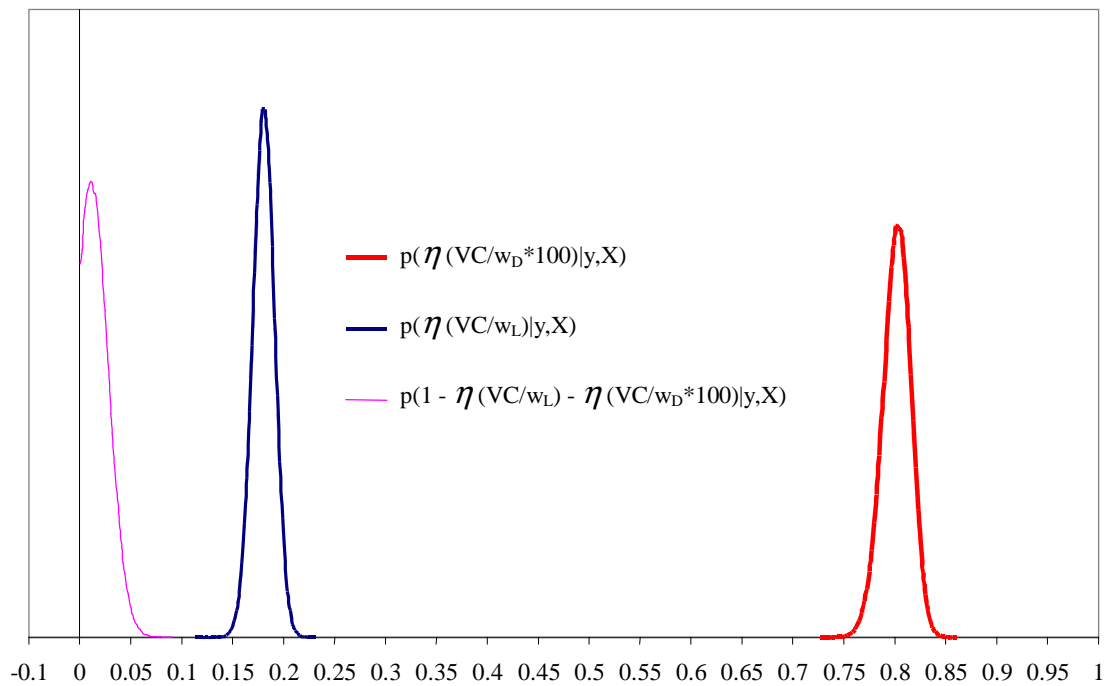
Przedstawione wartości oczekiwane a posteriori oryginalnych parametrów równania translogarytmicznej funkcji kosztu nie posiadają bezpośredniej ekonomicznej interpretacji. Źródłem wiedzy o kształtowaniu się kosztu zmiennego są m.in. elastyczności, które są liniowymi funkcjami logarytmów poszczególnych zmiennych mikroekonomicznej funkcji kosztu. Rysunki 16 i 17 przedstawiają sumaryzację tej wiedzy w postaci brzegowych gęstości a posteriori  $p(\eta(\cdot)|y, X)$  dla  $\eta(\cdot) = \frac{1}{58} \sum_{i=1}^{58} \eta(\cdot)_i$ , czyli średniej elastyczności kosztu względem cen czynników zmiennych, wielkości produkcji oraz zaangażowania czynnika stałego<sup>21</sup>. Wysmukłe i symetryczne kształty brzegowych funkcji gęstości a posteriori elastyczności kosztu względem ceny depozytów i innych pozyskanych środków, średniej płacy pracownika oddziały, wartości udzielonych kredytów i zaangażowania czynnika stałego tj. nieruchomości wskazują, że zebrane przekrojowo-czasowe dane pozwalają bardzo precyzyjnie szacować te charakterystyki technologii. Jedynie ucięty w zerze wykres gęstości a posteriori elastyczności względem nieobserwowanej zmiennej  $w_M$  wskazuje na nieistotną rolę ceny czynnika materiałowego.

---

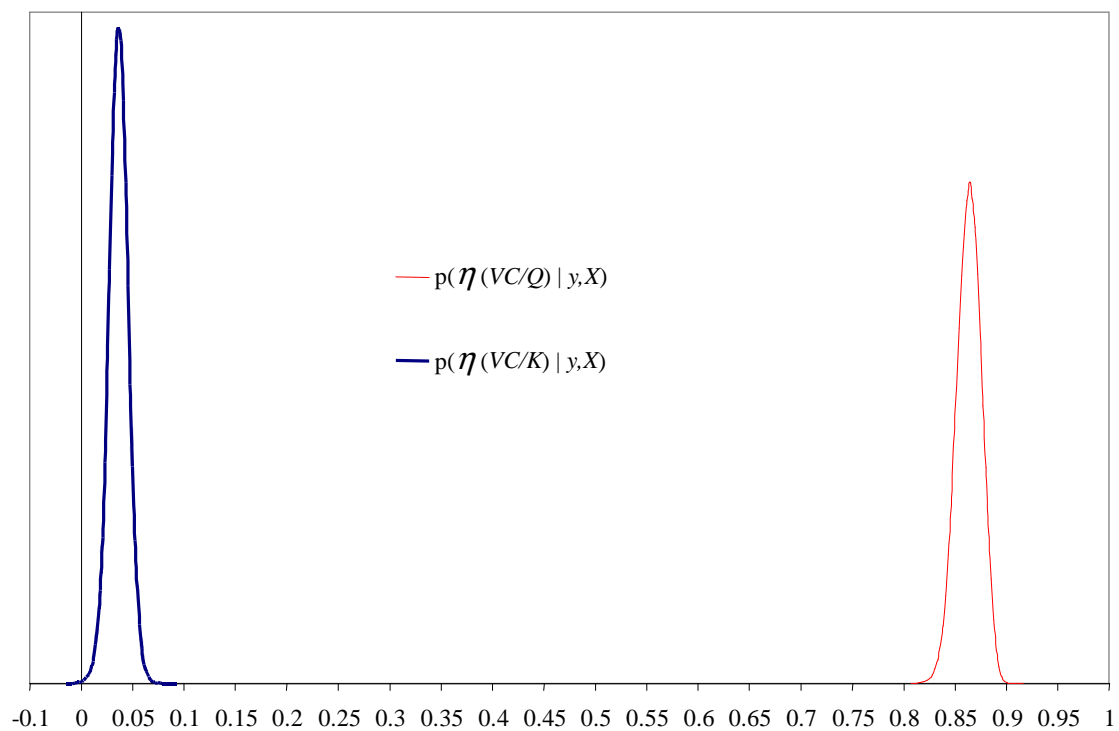
<sup>20</sup> Dokonano także pomiaru współliniowości dla danych przekrojowych w przypadku funkcji translogarytmicznej; odpowiedni miernik przyjmuje dla kolejnych kwartałów następujące wartości: 80635, 77056, 83981 i 52991. Zatem wykorzystanie danych przekrojowo-czasowych przyczyniło się częściowo do obniżenia stopnia silnej współliniowości zmiennych objaśniających w przypadku funkcji translogarytmicznej.

<sup>21</sup> Oba rysunki są przedstawione z zachowaniem identycznej skali zarówno na osi rzędnej, jak i odciętej.

Rysunek 16. Rozkłady a posteriori elastyczności względem cen czynników produkcji - VED ( $m=4, r^*=0.7$ ).



Rysunek 17. Rozkład a posteriori elastyczności względem wielkości produkcji ( $Q$ ) i względem nakładu czynnika stałego ( $K$ ) - VED ( $m=4, r^*=0.7$ ).



Natomiast Tabela 4 zawiera informacji o średniej wartości oczekiwanej elastyczności:

$$\bar{E}(\eta(\cdot)|y, X) = \frac{1}{58} \sum_{i=1}^{58} E(\eta(\cdot)_i | y, X)$$

i średnim odchyleniu standardowym a posteriori:

$$\bar{D}(\eta(\cdot)|y, X) = \frac{1}{58} \sum_{i=1}^{58} D(\eta(\cdot)_i|y, X),$$

które informują o charakterze technologii oddziałów banku komercyjnego.

Tabela 4. Średnia wartości oczekiwanych i odchyłeń standardowych a posteriori dla elastyczności - VED ( $m=4$ ,  $r^*=0.7$ ).

	$\eta(VC/w_D)$	$\eta(VC/w_L)$	$\eta(VC/Q)$	$\eta(VC/K)$	$1 - \eta(VC/w_L) - \eta(VC/w_D)$
Średnia wartość oczekiwana	0.802	0.180	0.863	0.036	0.018
Średnie odchylenie standardowe	0.037	0.026	0.025	0.023	0.063

Największy i statystycznie istotny wpływ na kształtowanie się kosztu zmiennego ma wielkość produkcji ( $Q$ ) i cena czynnika finansowego ( $w_D$ ). Wzrost wartości udzielonych kredytów o 1% powoduje wzrost kosztu zmiennego o 0.86% ( $\pm 0.025\%$ ) *ceteris paribus*. Natomiast wzrost ceny depozytów i innych pozyskanych środków powoduje wzrost kosztu zmiennego o około 0.8% ( $\pm 0.037\%$ ) *ceteris paribus*. Cena pracy ma mniejszy wpływ na poziom kosztu, aczkolwiek też odrywa istotną statystycznie rolę, tzn. wzrost płacy o 1% powoduje wzrost kosztów zmiennych o 0.18% ( $\pm 0.026\%$ ). Rola nieobserwowanej, stałej dla oddziałów ceny czynnika materiałowego jest znikoma lub żadna. Możliwe jest zatem usunięcie jej z modelu i powtórna estymacja bez nakładania warunku jednorodności funkcji kosztu względem cen pozostałych dwóch czynników zmiennych. Natomiast utrzymanie warunku jednorodności przez wydzielenie każdej z pozostałych dwóch cen przez jedną z nich spowodowałoby zmniejszenie liczby parametrów granicznej funkcji kosztu o 5 parametrów, co pozwoliłoby być może uzyskać wyniki o jeszcze większej precyzji. Dodatni znak przy elastyczności kosztu zmiennego względem powierzchni pomieszczeń biurowych oddziałów może wskazywać na zjawisko komplementarności tego czynnika produkcji.

W celu zbadania wrażliwości otrzymanych wyników na restrykcje, dokonano powtórnej estymacji modelu VED ( $m=4$ ) bez narzucania restrykcji na sumę elastyczności kosztu względem cen czynników zmiennych. Uzyskano prawie identyczne wyniki i jak przewidywano średnie wartości oczekiwane i odchylenia standardowe a posteriori dla elastyczności  $\eta(VC/w_D \cdot 100)$  i  $\eta(VC/w_L)$  były nieco większe niż w przypadku braku restrykcji, i wynosiły odpowiednio:

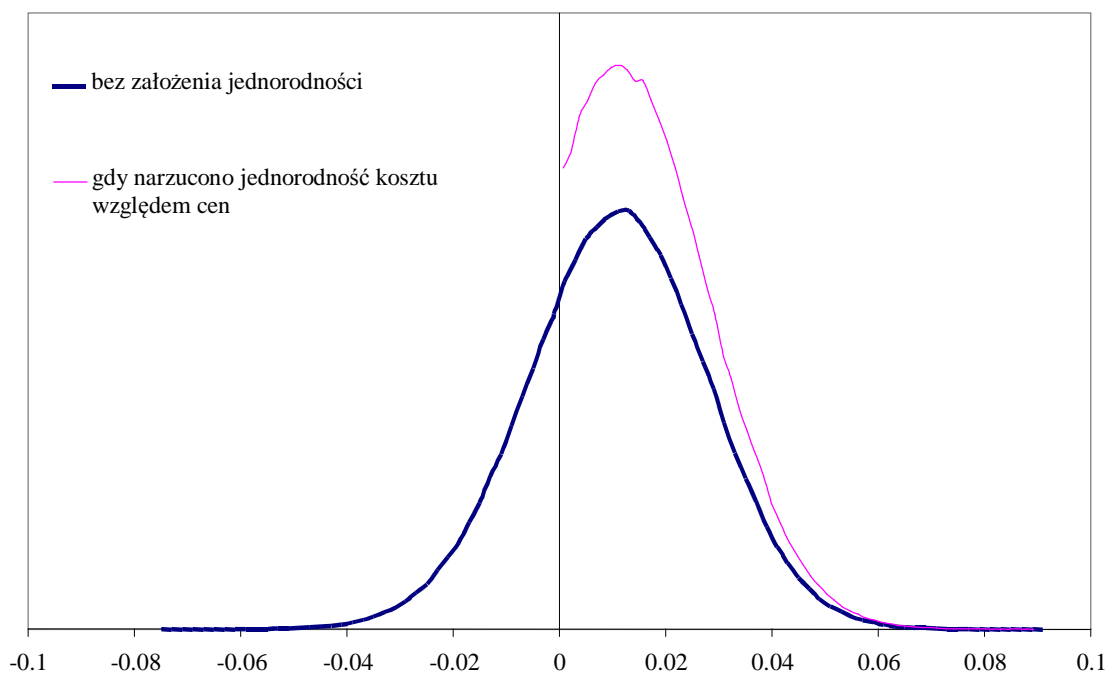
Tabela 5. Średnia wartości oczekiwane i odchylenia standardowe a posteriori elastyczności - VED ( $m=4$ ,  $r^*=0.7$ ) bez restrykcji:  $\eta(VC/w_D) + \eta(VC/w_L) < 1$ .

	$\eta(VC/w_D)$	$\eta(VC/w_D \cdot 100)$	$\eta(VC/Q)$	$\eta(VC/K)$	$1 - \eta(VC/w_L) - \eta(VC/w_D \cdot 100)$
Średnia wartość oczekiwana	0.807	0.182	0.862	0.036	0.011
Średnie odchylenie standardowe	0.039	0.027	0.026	0.024	0.065



Restrykcja na średnią arytmetyczną sumy elastyczności kosztu względem cen czynników zmiennych  $w_D$  i  $w_L$  jest wiążąca, co ilustruje Rysunek 18, lecz nie ma ona znaczenia z punktu widzenia ekonomicznej interpretacji roli cen czynników produkcji. Prawdopodobieństwo a posteriori, że średnia suma elastyczności względem zmiennych  $w_D$  i  $w_L$  jest większa niż jeden wynosi około 0.26. Wykres gęstości a posteriori wielkości  $1-\eta(VC/w_D \cdot 100)-\eta(VC/w_L)$ , która przy założeniu jednorodności (i tylko wtedy) może być interpretowana jako elastyczność względem nieobserwowanej ceny  $w_M$ , jest bardziej spłaszczony przy braku restrykcji niż przy jej narzuceniu. Jednak modalne rozkładu w obu przypadkach prawie się pokrywają i są w bliskim otoczeniu zera, wskazując na nieistotną rolę nieobserwowanej ceny w kształtowaniu się kosztów zmiennych. Podsumowując porównanie obu przypadków – z narzuconą a priori jednorodnością względem cen i bez niej – można stwierdzić, że dane wyraźnie potwierdzają sensowność tego niezbędnego z teoretycznego punktu widzenia warunku regularności ekonomicznej.

Rysunek 18. Gęstość a posteriori  $p(1-(\eta(VC/w_D \cdot 100)-\eta(VC/w_L)) - VED (m=4, r^*=0.7)$ .



Przyjęcie translogarytmicznej formy funkcyjnej dla funkcji kosztu pozwala na szczegółową analizę charakteru technologii badanych oddziałów. Główne charakterystyki tj. wartości oczekiwane i odchylenia standardowe brzegowych rozkładów a posteriori elastyczności kosztu przedstawia Tabela 6 (oddziały posortowane są malejąco ze względu na średnią wartość udzielonych kredytów w badanym okresie).

Warunki regularności ekonomicznej, choć zostały narzucone wyłącznie dla średnich wartości oczekiwanych a posteriori elastyczności względem cen czynników zmiennych, ich sumy i średniej elastyczności względem produkcji (por. wzór (73) i (74)), to spełnione są także dla większości oddziałów w próbie. W oddziałach, w których wartość oczekiwana a posteriori sumy elastyczności kosztu względem dwóch obserwowanych cen czynników zmiennych jest większa od jedności, oceny te z reguły charakteryzują się dużymi odchyleniami standardowymi, zatem z punktu zarówno interpretacji ekonomicznej jak i statystycznej można przyjąć, że suma tych elastyczności nie różni istotnie się od jedności. Jedynie w przypadku oddziałów o numerach 18, 32 i 58 wyrażenie  $E(\eta(w_D \cdot 100)_i + \eta(w_L)_i | y, X)$  przyjmuje wartości na tyle większe od jedności przy stosunkowo małym odchyleniu standardowym, iż losowanie Gibbsa stało się bardzo nieefektywne przy próbie narzucenia wspomnianego warunku dla każdego oddziału. Pewna grupa oddziałów charakteryzuje się ujemną i istotnie różną od zera wartością elastyczności kosztu względem nakładów czynnika stałego. Wnikliwa analiza pozwoliła stwierdzić, że prawidłowość ta dotyczy oddziałów mających najmniejsze zaangażowanie tego czynnika mierzone powierzchnią pomieszczeń budynków eksploatowanych przez oddział (por. także Rysunek 21). Przyjęcie translogarytmicznej funkcji kosztu uwolniło od założenia, iż najmniejsza (największa) firma pod względem wielkości produkcji charakteryzuje się największą (najmniejszą) wartością współczynnika skali produkcji; bezzasadność takiego założenia potwierdziły wyniki empiryczne.

Tabela 6. Wartości oczekiwane i odchylenia standardowe a posteriori dla elastyczności kosztu w poszczególnych oddziałach banku- VED ( $m=4, r^*=0.7$ ).

$i$	$E(\eta(w_D)_i   \cdot)$	$D(\eta(w_D)_i   \cdot)$	$E(\eta(w_L)_i   \cdot)$	$D(\eta(w_L)_i   \cdot)$	$E(\eta(w_D)_i + \eta(w_L)_i   \cdot)$	$D(\eta(w_D)_i + \eta(w_L)_i   \cdot)$	$E(\eta(Q)_i   \cdot)$	$D(\eta(Q)_i   \cdot)$	$E(\eta(K)_i   \cdot)$	$D(\eta(K)_i   \cdot)$
1	0.806	0.055	0.107	0.040	0.914	0.094	0.912	0.035	0.095	0.028
2	0.816	0.051	0.121	0.031	0.937	0.081	0.919	0.033	0.036	0.027
3	0.716	0.050	0.117	0.028	0.833	0.079	0.929	0.033	-0.067	0.027
4	0.833	0.053	0.149	0.028	0.981	0.081	0.893	0.033	0.082	0.027
5	0.811	0.034	0.149	0.032	0.960	0.066	0.894	0.031	0.070	0.022
6	0.820	0.035	0.171	0.029	0.991	0.064	0.884	0.032	0.086	0.022
7	0.839	0.037	0.176	0.030	1.015	0.067	0.901	0.033	-0.001	0.023
8	0.828	0.037	0.156	0.025	0.984	0.062	0.888	0.033	0.069	0.023
9	0.850	0.063	0.143	0.041	0.993	0.104	0.876	0.042	0.142	0.039
10	0.860	0.059	0.153	0.034	1.013	0.094	0.872	0.042	0.130	0.039
11	0.803	0.061	0.148	0.034	0.951	0.095	0.891	0.042	0.024	0.039
12	0.797	0.058	0.146	0.033	0.944	0.090	0.866	0.042	0.137	0.038
13	0.842	0.030	0.170	0.025	1.012	0.055	0.873	0.023	0.089	0.020
14	0.703	0.029	0.146	0.020	0.849	0.049	0.896	0.023	-0.045	0.020
15	0.843	0.029	0.156	0.019	0.999	0.048	0.876	0.023	0.082	0.020
16	0.837	0.029	0.162	0.022	0.999	0.051	0.867	0.022	0.111	0.019
17	0.808	0.031	0.174	0.021	0.982	0.052	0.863	0.023	0.088	0.019

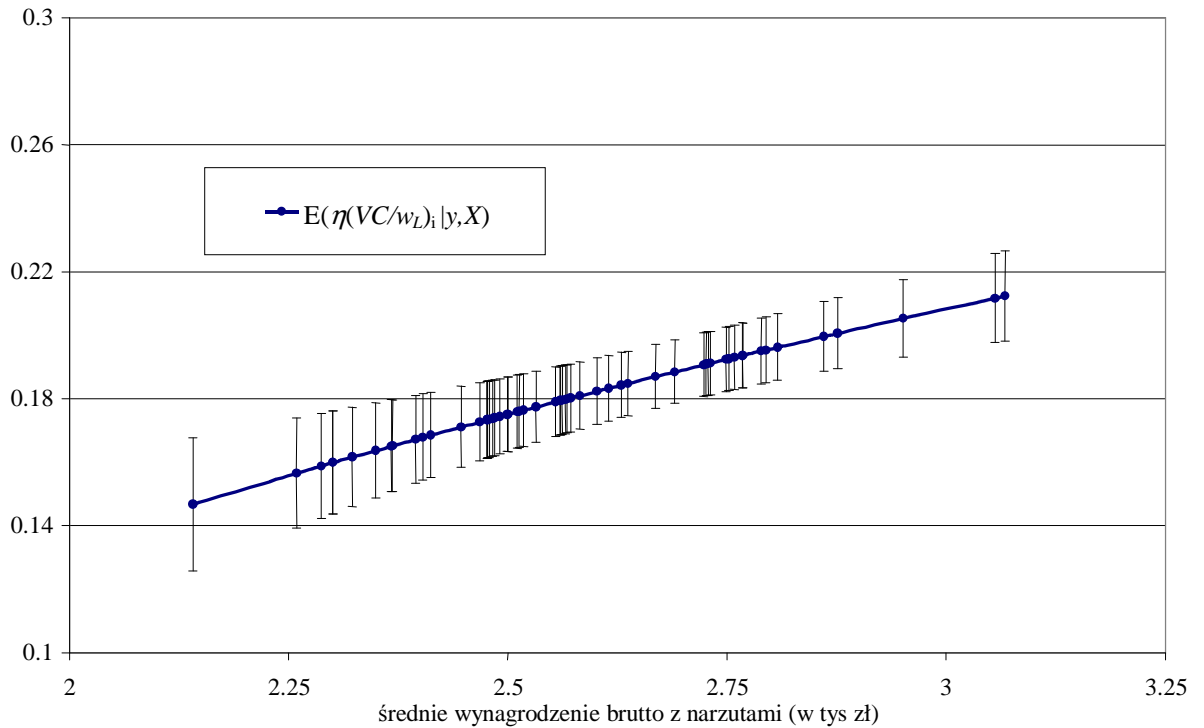
$i$	$E(\eta(w_D)_{i })$	$D(\eta(w_D)_{i })$	$E(\eta(w_L)_{i })$	$D(\eta(w_L)_{i })$	$E(\eta(w_D)_{i+})$	$D(\eta(w_D)_{i+})$	$E(\eta(Q)_{i })$	$D(\eta(Q)_{i })$	$E(\eta(K)_{i })$	$D(\eta(K)_{i })$
18	0.910	0.030	0.167	0.021	1.077	0.051	0.861	0.025	0.139	0.019
19	0.771	0.033	0.165	0.019	0.936	0.052	0.877	0.025	0.016	0.019
20	0.882	0.033	0.159	0.022	1.041	0.055	0.872	0.025	0.077	0.019
21	0.795	0.031	0.180	0.020	0.976	0.051	0.864	0.021	0.063	0.019
22	0.816	0.031	0.165	0.017	0.981	0.048	0.872	0.022	0.046	0.020
23	0.741	0.032	0.171	0.018	0.911	0.050	0.893	0.023	-0.090	0.020
24	0.839	0.031	0.159	0.037	0.998	0.069	0.858	0.021	0.116	0.018
25	0.777	0.045	0.198	0.037	0.976	0.082	0.873	0.028	-0.022	0.022
26	0.822	0.038	0.205	0.030	1.027	0.068	0.848	0.027	0.102	0.020
27	0.856	0.043	0.178	0.034	1.033	0.077	0.863	0.029	0.066	0.023
28	0.713	0.030	0.193	0.033	0.905	0.063	0.877	0.023	-0.068	0.018
29	0.764	0.023	0.200	0.023	0.964	0.046	0.877	0.019	-0.057	0.016
30	0.774	0.024	0.181	0.019	0.955	0.043	0.864	0.021	0.023	0.017
31	0.829	0.026	0.174	0.017	1.004	0.043	0.857	0.022	0.079	0.018
32	0.874	0.027	0.189	0.022	1.063	0.048	0.843	0.021	0.143	0.017
33	0.795	0.036	0.181	0.031	0.976	0.067	0.879	0.021	-0.043	0.028
34	0.784	0.035	0.200	0.025	0.984	0.060	0.869	0.019	-0.028	0.027
35	0.780	0.037	0.189	0.023	0.968	0.060	0.866	0.021	-0.001	0.028
36	0.780	0.037	0.172	0.023	0.952	0.060	0.869	0.020	-0.005	0.027
37	0.835	0.031	0.193	0.030	1.027	0.061	0.850	0.018	0.063	0.025
38	0.760	0.030	0.178	0.022	0.938	0.052	0.857	0.017	0.004	0.025
39	0.836	0.030	0.180	0.020	1.016	0.050	0.850	0.018	0.056	0.025
40	0.879	0.033	0.179	0.023	1.058	0.056	0.840	0.017	0.116	0.025
41	0.780	0.023	0.206	0.026	0.986	0.048	0.838	0.021	0.065	0.015
42	0.738	0.022	0.193	0.021	0.931	0.042	0.867	0.022	-0.073	0.016
43	0.812	0.022	0.215	0.019	1.027	0.041	0.843	0.021	0.039	0.016
44	0.804	0.021	0.192	0.016	0.997	0.037	0.839	0.021	0.069	0.016
45	0.770	0.051	0.223	0.032	0.994	0.083	0.853	0.022	-0.036	0.027
46	0.751	0.049	0.188	0.029	0.938	0.078	0.850	0.022	0.001	0.026
47	0.807	0.054	0.201	0.032	1.009	0.086	0.844	0.024	0.030	0.028
48	0.846	0.061	0.185	0.035	1.031	0.096	0.837	0.026	0.083	0.029
49	0.740	0.022	0.175	0.021	0.915	0.043	0.864	0.015	-0.069	0.017
50	0.756	0.024	0.205	0.017	0.961	0.041	0.840	0.017	0.009	0.017
51	0.754	0.022	0.216	0.014	0.969	0.035	0.855	0.015	-0.077	0.017
52	0.789	0.024	0.199	0.019	0.988	0.044	0.842	0.015	0.010	0.017
53	0.720	0.058	0.192	0.034	0.912	0.092	0.850	0.031	-0.051	0.031
54	0.831	0.058	0.232	0.030	1.063	0.088	0.825	0.033	0.053	0.032
55	0.736	0.065	0.200	0.033	0.936	0.098	0.847	0.034	-0.060	0.034
56	0.838	0.058	0.234	0.033	1.071	0.091	0.807	0.033	0.109	0.033
57	0.771	0.023	0.230	0.031	1.000	0.054	0.818	0.016	0.011	0.016
58	0.826	0.022	0.267	0.019	1.093	0.041	0.787	0.016	0.054	0.016

Przyjęcie translogarytmicznej funkcji kosztu powoduje m.in., że elastyczności kosztu zmiennego względem cen czynników zmiennych, produkcji bądź nakładów czynnika stałego są funkcjami wszystkich tych wielkości. W celu uproszczenia graficznej prezentacji na Rysunkach 19-22 przedstawiono wartości oczekiwane a posteriori dla poszczególnych elastyczności jako funkcje tylko jednej zmiennej, przyjmując pozostałe zmienne na poziomie średnich wartości ich logarytmów; por. wzór (74). Dodatkowo przedstawiono odchylenia

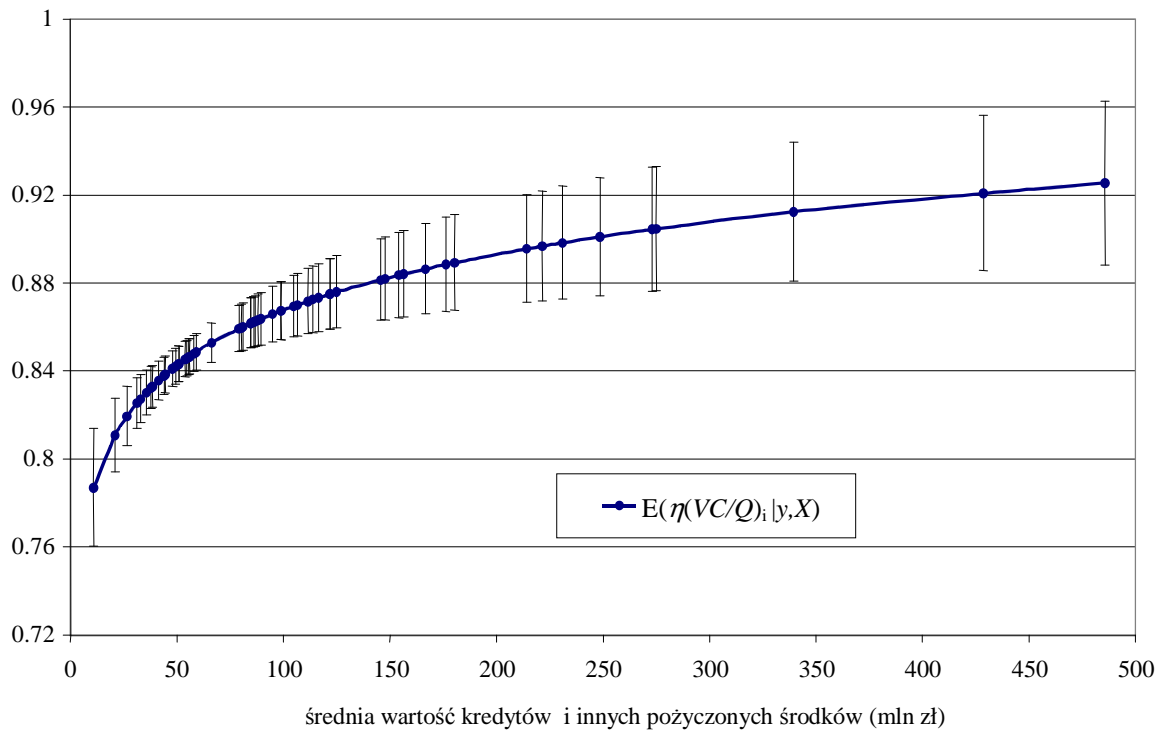
standardowe a posteriori elastyczności (*in plus, in minus*), które wykazują tę prawidłowość, że w przypadku oddziałów charakteryzujących się najmniejszymi i największymi wartościami zmiennych przyjmują one większą wartość niż w przypadku oddziałów, dla których wartości tych zmiennych oscylują wokół średniej z próby. Oznacza to, że najpewniej wnioskuje się o elastycznościach w przypadku oddziałów „przeciętnych”.

Analiza kształtu krzywych powstałych z połączenia punktów reprezentujących wartości oczekiwane a posteriori elastyczności dostarcza użytecznych informacji, które mogą być wykorzystane m.in. w tworzeniu polityki płacowej, depozytowej, kredytowej i inwestycyjnej banku. Elastyczności kosztu zmiennego względem ceny pracy, produkcji i nakładów czynnika stałego są rosnącymi funkcjami odpowiednio tych zmiennych (przy pozostałych ustalonych). Wraz ze wzrostem średniego wynagrodzenia w oddziale, procentowy wzrost średniego wynagrodzenia powoduje coraz większy wzrost względny kosztu zmiennego. Zatem wskazane jest kontrolowanie poziomu płac w oddziałach o przeciętnie najwyższych wynagrodzeniach, natomiast w przypadku oddziałów o najniższych płacach, ich wzrost może stanowić element motywujący pracowników do wydajniejszej pracy. Rysunek 20 wskazuje, że wszystkie oddziały charakteryzują się rosnącym efektem skali (odwrotność elastyczności kosztu względem produkcji), który maleje wraz ze wzrostem udzielonych kredytów. W efekcie oddziały najmniejsze, o średniej wartości udzielonych kredytów w kwartale poniżej około 125 mln złotych, (które stanowią ponad 72% wszystkich oddziałów) mogą w sposób znaczący zwiększyć produkcję ponosząc przy tym proporcjonalnie mniejsze koszty zmienne. Natomiast Rysunek 21 informuje, że elastyczność kosztu zmiennego względem nakładu czynnika stałego jest mniejsza od zera w przypadku oddziałów mających w dyspozycji najmniejszą powierzchnię posiadanych bądź najmowanych budynków. Oznacza to, że w tych oddziałach występuje zjawisko substytucji między czynnikiem stałym a czynnikami zmiennymi. Wstępna analiza Rysunku 22 przedstawiającego krzywą elastyczności kosztu zmiennego względem ceny depozytów i innych pozyskanych środków wskazuje, że jest ona malejącą funkcją tej ceny. Jednakże nachylenie krzywej do osi odciętych i duże odchylenia standardowe powodują, iż można postawić tezę, że omawiana elastyczność nie zależy od tej ceny, zatem jest stała i kształtuje się na poziomie oscylującym wokół 0.8.

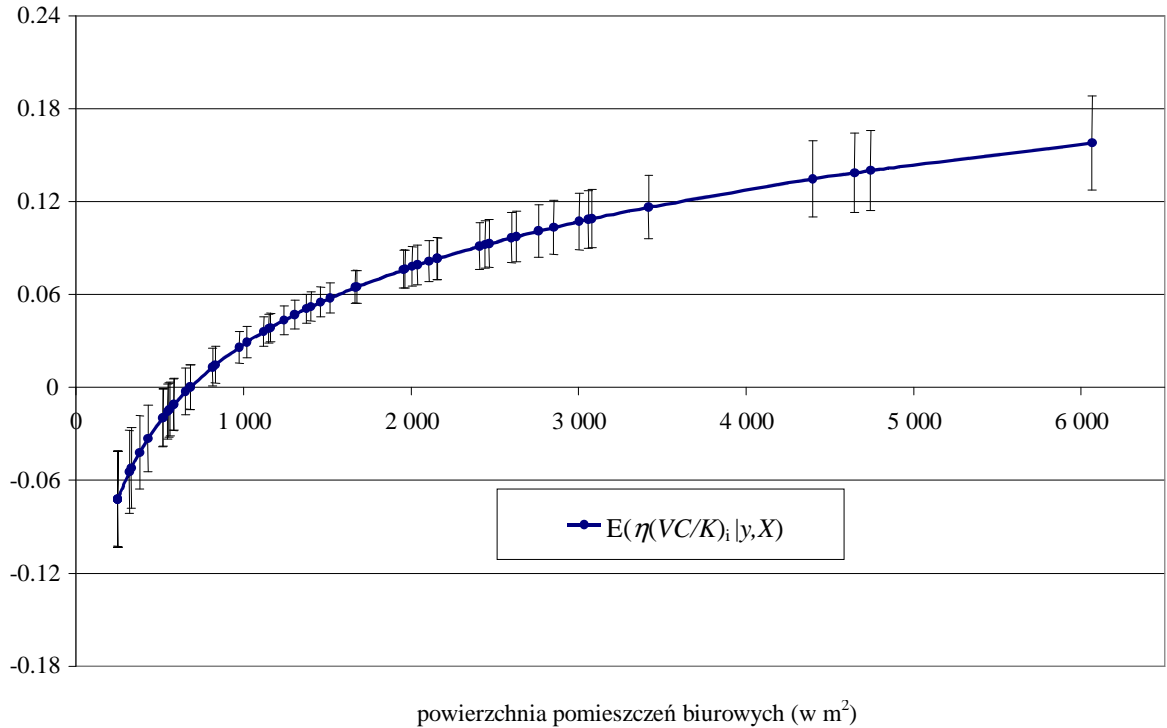
Rysunek 19. Wartości oczekiwane i odchylenia standardowe a posteriori elastyczności kosztu względem średniej ceny płacy dla poszczególnych oddziałów (przy ustalonych wartościach pozostałych zmiennych egzogenicznych na poziomie średniej geometrycznej z próby) - VED ( $m=4, r^*=0.7$ ).



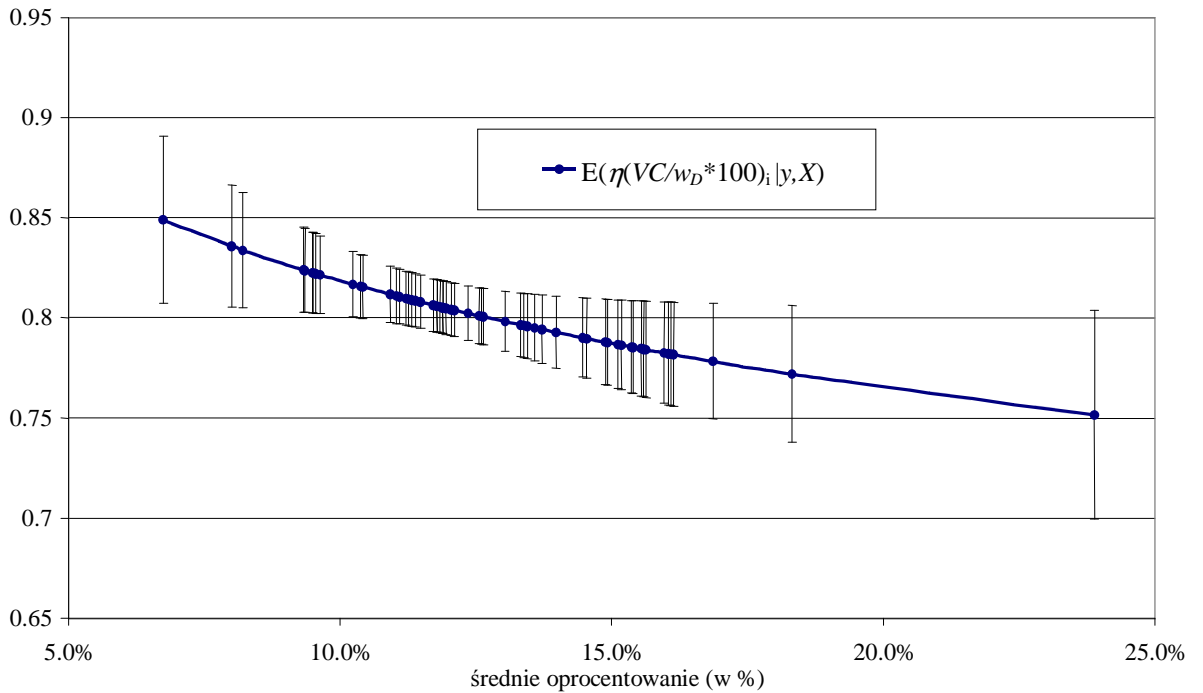
Rysunek 20. Wartości oczekiwane i odchylenia standardowe a posteriori elastyczności kosztu względem produkcji dla poszczególnych oddziałów (przy ustalonych wartościach pozostałych zmiennych egzogenicznych na poziomie średniej geometrycznej z próby) - VED ( $m=4, r^*=0.7$ ).



Rysunek 21. Wartości oczekiwane i odchylenia standardowe a posteriori elastyczności kosztu względem nakładów czynnika stałego dla poszczególnych oddziałów (przy ustalonych wartościach pozostałych zmiennych egzogenicznych na poziomie średniej geometrycznej z próby) - VED ( $m=4, r^*=0.7$ ).



Rysunek 22. Wartości oczekiwane i odchylenia standardowe a posteriori elastyczności kosztu względem ceny depozytów i innych pozyskanych środków dla poszczególnych oddziałów (przy ustalonych wartościach pozostałych zmiennych egzogenicznych na poziomie średniej geometrycznej z próby) - VED ( $m=4, r^*=0.7$ ).



Zasadniczym celem podjętych badań jest jednak ocena efektywności kosztowej oddziałów banku komercyjnego. Tabela 7 przedstawia wartości oczekiwane i odchylenia standardowe a posteriori dla wskaźników efektywności  $r_i = \exp(-z_i)$  poszczególnych oddziałów (uporządkowanych malejąco względem średniej wielkości produkcji w czterech kwartałach) oraz dla parametrów  $\lambda_i$ , będących wartością oczekiwaną rozkładu a priori zmiennej reprezentującej nieefektywność  $z_i$ . Natomiast dychotomiczne zmienne  $w_{i2}$ ,  $w_{i3}$  i  $w_{i4}$  niosą informacje o specjalizacji oddziału, jego wielkości i o tym czy posiada filie świadczące pełny zakres usług bankowych. Badane oddziały charakteryzują się przeciętnie dość wysoką efektywnością, skoro średnia wartość oczekiwana a posteriori dla efektywności,  $E(r|y, X) = \frac{1}{58} \sum_{i=1}^{58} E(r_i|y, X)$ , wynosi 0.918 (przy średnim odchyleniu standardowym 0.017). Z

ekonomicznego punktu widzenia oznacza to, że minimalny koszt zmienny konieczny do uzyskania danej wielkości produkcji przy obserwowanych cenach czynników zmiennych i nakładach czynnika stałego (który wynika z mikroekonomicznej funkcji kosztu) stanowi w badanych oddziałach banku przeciętnie 91.8% rzeczywiście poniesionego kosztu. Zatem nadwyżkowy (nieuzasadniony) koszt zmienny stanowi średnio 8.2% obserwowanego kosztu. Najmniejszą efektywnością - na poziomie 0.769 ( $\pm 0.016$ ) - charakteryzuje się oddział o numerze 58, natomiast największą - 0.996 ( $\pm 0.004$ ), czyli praktycznie pełną - oddział 56. Mediana rozkładu a posteriori dla efektywności kształtuje się na poziomie 0.93, gdy tymczasem dla mediany rozkładu a priori przyjęto 0.528 (przy  $r^* = 0.7$ ); dane wskazują jednoznacznie na wyższą niż spodziewana a priori efektywność kosztowa oddziałów. Zróżnicowanie indywidualnych wskaźników efektywności jest o wiele mniejsze niż wskazywałyby wyniki wstępne w przypadku modelu (70), a małe wartości odchylenia standardowego oznaczają dużą precyzję otrzymanych wyników. Sporządzony ranking efektywności oddziałów na podstawie tych wskaźników nie zależy od zarówno od założeń a priori tj. stałej  $r^*$ , jak i restrykcji na sumę elastyczności względem cen  $w_D$  i  $w_L$ . Współczynniki korelacji rang Spearmana, gdy  $r^*$  wynosi 0.5, 0.7 i 0.9 kształtują się na poziomie ponad 0.99. Jednakże ten ranking jest mało precyzyjny, ponieważ większość czyli 44 z 58 oddziałów (75% ogółu) charakteryzuje się efektywnością z przedziału (0.88; 1), a średnie odchylenie standardowe wynosi około 0.017, co powoduje, że dwa sąsiednie oddziały w rankingu nie są - ze statystycznego punktu widzenia - rozróżnialne pod względem poziomu efektywności. Natomiast z całą pewnością można wyodrębnić oddziały wzorcowe (o numerach 56, 55, 57 i 26) i mało efektywne (3, 8, 11, 13, 14, 15, 19, 20, 27, 37, 53 i zwłaszcza 58). Przedmiotem

szczegółowej analizy centrali banku powinno być porównanie stylu zarządzania i innych uwarunkowań oddziałów wzorcowych i tych oddziałów, w których nieuzasadniona część kosztów zmiennych przekracza np. 12%.

Na końcu omawianej tabeli znajdują się średnie wartości oczekiwane a posteriori efektywności dla różnych grup oddziałów. Trudno jest dostrzec wyraźne zależności między rodzajem oddziału (określonym przez zmienne dychotomiczne) a poziomem efektywności, zwłaszcza, że średnie odchylenia standardowe a posteriori są relatywnie duże w stosunku do różnic między średnimi wartościami oczekiwanymi a posteriori. Potwierdzeniem braku wyraźnej reguły jest fakt, że zarówno oddział o najwyższej ( $i=56$ ) jak i najmniejszej efektywności ( $i=58$ ) należy do tej samej podgrupy, tj. posiada nadwyżkę kredytów nad depozytami, wartość udzielonych kredytów wynosi poniżej 100 mln zł i nie posiada filii. W tej sytuacji wskazane jest użycie np. testów H.P.D., które pozwolą na formalne testowanie ewentualnych systematycznych różnic w efektywności kosztowej oddziałów.

Tabela 7. Wartości oczekiwane i odchylenia standardowe a posteriori wskaźników efektywności modelu VED ( $m=4, r^*=0.7$ ).

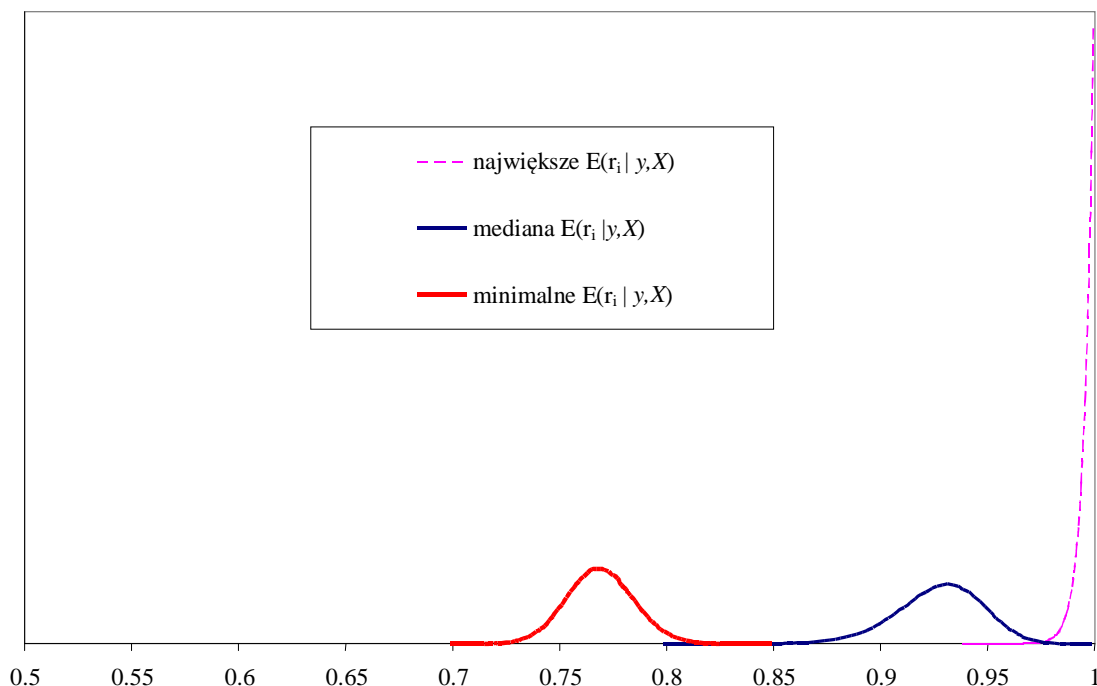
$i$	$w_{i2}$	$w_{i3}$	$w_{i4}$	$E(\lambda_i/y, X, W)$	$D(\lambda_i/y, X, W)$	$E(r_i/y, X)$	$D(r_i/y, X)$
1	1	1	1	0.082	0.042	0.901	0.036
2	0	1	0	0.126	0.038	0.898	0.030
3	0	1	0	0.126	0.038	0.872	0.051
4	0	1	1	0.130	0.062	0.898	0.021
5	0	1	0	0.126	0.038	0.890	0.022
6	0	1	0	0.126	0.038	0.927	0.021
7	0	1	0	0.126	0.038	0.983	0.015
8	0	1	0	0.126	0.038	0.844	0.017
9	0	1	1	0.130	0.062	0.963	0.024
10	0	1	0	0.126	0.038	0.897	0.020
11	0	1	0	0.126	0.038	0.852	0.017
12	0	1	1	0.130	0.062	0.888	0.026
13	0	1	0	0.126	0.038	0.834	0.015
14	1	1	0	0.080	0.030	0.839	0.033
15	0	1	1	0.130	0.062	0.840	0.015
16	0	1	0	0.126	0.038	0.959	0.017
17	1	1	0	0.080	0.030	0.976	0.016
18	0	1	0	0.126	0.038	0.976	0.021
19	0	1	0	0.126	0.038	0.836	0.016
20	0	1	0	0.126	0.038	0.852	0.016
21	0	1	0	0.126	0.038	0.964	0.017
22	1	1	0	0.080	0.030	0.901	0.016
23	0	1	0	0.126	0.038	0.919	0.027
24	1	1	1	0.082	0.042	0.886	0.015
25	0	0	0	0.108	0.031	0.891	0.014
26	0	0	0	0.108	0.031	0.982	0.013
27	0	0	0	0.108	0.031	0.831	0.014
28	0	0	1	0.117	0.068	0.930	0.024
29	0	0	0	0.108	0.031	0.935	0.016



$i$	$w_{i2}$	$w_{i3}$	$w_{i4}$	$E(\lambda_i/y, X, W)$	$D(\lambda_i/y, X, W)$	$E(r_i / y, X)$	$D(r_i / y, X)$
30	1	0	0	0.067	0.016	0.855	0.015
31	0	0	0	0.108	0.031	0.953	0.015
32	0	0	0	0.108	0.031	0.909	0.017
33	1	0	0	0.067	0.016	0.973	0.014
34	0	0	0	0.108	0.031	0.963	0.014
35	1	0	0	0.067	0.016	0.964	0.015
36	0	0	0	0.108	0.031	0.957	0.015
37	1	0	0	0.067	0.016	0.833	0.013
38	1	0	0	0.067	0.016	0.944	0.015
39	1	0	0	0.067	0.016	0.946	0.016
40	1	0	0	0.067	0.016	0.908	0.014
41	1	0	0	0.067	0.016	0.910	0.014
42	1	0	0	0.067	0.016	0.971	0.017
43	1	0	0	0.067	0.016	0.943	0.015
44	1	0	0	0.067	0.016	0.961	0.014
45	1	0	0	0.067	0.016	0.992	0.007
46	1	0	0	0.067	0.016	0.947	0.015
47	1	0	0	0.067	0.016	0.940	0.015
48	1	0	0	0.067	0.016	0.909	0.013
49	1	0	0	0.067	0.016	0.957	0.016
50	1	0	0	0.067	0.016	0.961	0.013
51	1	0	0	0.067	0.016	0.920	0.016
52	1	0	0	0.067	0.016	0.948	0.013
53	1	0	0	0.067	0.016	0.870	0.016
54	1	0	0	0.067	0.016	0.931	0.014
55	1	0	0	0.067	0.016	0.986	0.012
<b>56</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0.067</b>	<b>0.016</b>	<b>0.996</b>	<b>0.004</b>
57	1	0	0	0.067	0.016	0.995	0.005
<b>58</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0.067</b>	<b>0.016</b>	<b>0.769</b>	<b>0.016</b>
Średnia dla oddziałów z $w_{i2}=1, w_{i3}=1$ i $w_{i4}=1$						0.893	0.026
Średnia dla oddziałów z $w_{i2}=0, w_{i3}=0$ i $w_{i4}=0$						0.927	0.014
Średnia dla oddziałów z $w_{i2}=1, w_{i3}=0$ i $w_{i4}=0$						0.933	0.014
Średnia dla oddziałów z $w_{i2}=0, w_{i3}=1$ i $w_{i4}=0$						0.899	0.022
Średnia dla oddziałów z $w_{i2}=0, w_{i3}=0$ i $w_{i4}=1$						0.931	0.024
Średnia dla oddziałów z $w_{i2}=1, w_{i3}=1$ i $w_{i4}=0$						0.906	0.022
Średnia dla oddziałów z $w_{i2}=0, w_{i3}=1$ i $w_{i4}=1$						0.897	0.022
Średnia dla oddziałów z $w_{i2}=1$						0.928	0.015
Średnia dla oddziałów z $w_{i2}=0$						0.909	0.020
Średnia dla oddziałów z $w_{i3}=1$						0.900	0.022
Średnia dla oddziałów z $w_{i3}=0$						0.932	0.014
Średnia dla oddziałów z $w_{i4}=1$						0.901	0.023
Średnia dla oddziałów z $w_{i4}=0$						0.921	0.016
Średnia dla oddziałów						0.918	0.017

Rysunek 23 przedstawia gęstości brzegowe wybranych oddziałów o największej i najmniejszej efektywności oraz oddziału, którego wskaźnik efektywności jest medianą rozkładu efektywności. Z uwagi na unormowanie wskaźnika  $r_i$  masa gęstości a posteriori dla oddziału najefektywniejszego skupiona jest w bliskim lewostronnym sąsiedztwie jedności.

Rysunek 23. Brzegowe gęstości a posteriori dla efektywności na przykładzie wybranych oddziałów - VED ( $m=4, r^*=0.7$ ).



Prezentowane powyżej wyniki dotyczące charakteru procesu produkcji i elastyczności kosztu nie są wrażliwe na założenia w priori, tj. zmianę parametru  $r^*$ . Wyniki uzyskane zarówno dla wartości mniejszych niż 0.7 jak i 0.9 są prawie identyczne. Podstawowe wyniki dla  $r^*=0.9$  zawierają Załącznik 1. Parametr  $r^*$  ma nieznaczny wpływ na poziom efektywności oddziałów, zaś rozkłady a posteriori dla składowych wektora  $\gamma$  są bardzo wrażliwe na wielkość parametru  $r^*$  - zgodnie z intuicją.

W proponowanym modelu VED zmiany poziomu efektywności każdej z firm są wyjaśniane przez dychotomiczne zmienne egzogeniczne ( $w_{ij}$ ). O kierunku i sile wpływu tych przyczyn nieefektywności informują składowe wektora  $\gamma$ , których wartości oczekiwane i odchylenia standardowe a posteriori prezentowane są poniżej:

Tabela 8. Wartości oczekiwane i odchylenia standardowe a posteriori dla  $\gamma$ - VED ( $m=4, r^*=0.7$  i  $r^*=0.9$ ).

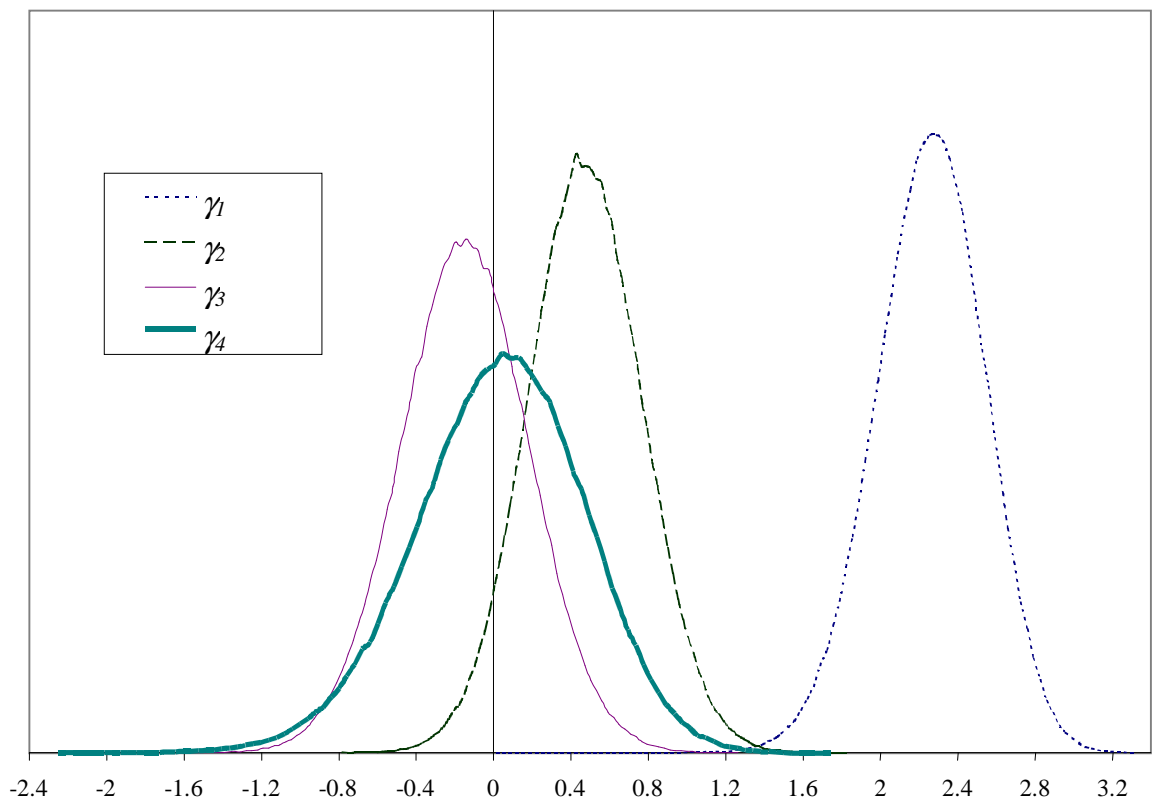
	$r^*=0.5$		$r^*=0.7$		$r^*=0.9$	
	$E(\cdot y,X,W)$	$D(\cdot y,X,W)$	$E(\cdot y,X,W)$	$D(\cdot y,X,W)$	$E(\cdot y,X,W)$	$D(\cdot y,X,W)$
$\gamma_1$	2.049	0.267	2.267	0.279	2.495	0.295
$\gamma_2$	0.623	0.288	0.467	0.293	0.307	0.302
$\gamma_3$	-0.005	0.329	-0.150	0.330	-0.303	0.336
$\gamma_4$	0.018	0.439	0.020	0.435	0.015	0.431

Interpretując powyższe wyniki należy zauważyć, iż niezależnie od przyjętej wartości  $r^*$  zarówno struktura organizacyjna oddziałów ( $w_{i4}=1$ , gdy oddział posiada filię) jak i skala prowadzonej działalności (oddział prowadzi działalność na dużą skalę, gdy łączna wartości

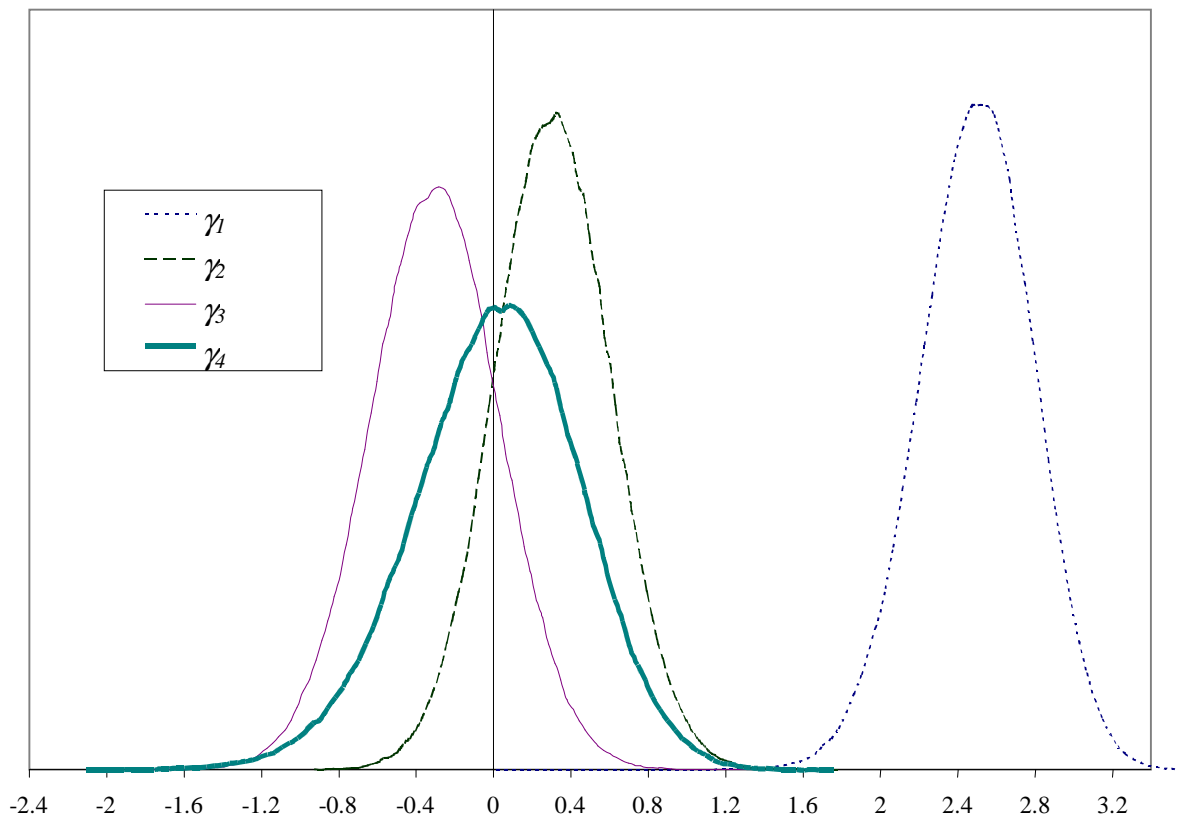
udzielonych kredytów wynosi powyżej 100 mln zł;  $w_{i3}=1$ ) nie mają statystycznie istotnego wpływu na zróżnicowanie efektywności w poszczególnych grupach oddziałów. Nieznaczące różnice w poziomie efektywności wydają się być spowodowane specjalizacją oddziałów, tj. większą efektywnością charakteryzują się oddziały kredytowe (posiadające nadwyżkę wartości kredytów nad depozytami;  $w_{i2}=1$ ) niż depozytowe, aczkolwiek im wyższą wartość przyjmuje  $r^*$ , tym wnioskowanie to jest słabsze. Pojawia się zatem problem wyboru odpowiedniego poziomu dla  $r^*$  i problem statystycznego testowania roli zmiennych  $w_{ij}$ . W modelu VED istnieje możliwość formalnego testowania istotności systematycznych różnic w efektywności kosztowej, co w świetle otrzymanych wyników wydaje się konieczne. Duże odchylenia standardowe a posteriori parametrów  $\gamma_j$  (dla  $j=3,4$ ) wskazują, iż zmienne  $w_{i3}$  i  $w_{i4}$  się nieistotne, co potwierdzają testy zastosowane poniżej, a omówione w podrozdziale 4.4.

Test H.P.D. pozwala zbadać, czy wybrana wartość parametru, tj.  $\gamma_j^*=0$  ( $j=2,3,4$ ), należy do przedziału o największej gęstości a posteriori przy ustalonym z góry prawdopodobieństwie wynoszącym  $1-\alpha$ . Odpowiednie brzegowe rozkłady a posteriori dla pojedynczych składowych wektora  $\gamma$  (dla  $r^*$  równego 0.7 i 0.9) przedstawiają Rysunki 24-25. Wynika z nich, że wartość  $\gamma_j^*=0$  ( $j=3,4$ ) należą z dużym prawdopodobieństwem do przedziałów o największej gęstości a posteriori. W przypadku parametru  $\gamma_3$  prawdopodobieństwo to wynosi przynajmniej 0.33 i 0.64, gdy odpowiednio  $r^*=0.7$  i  $r^*=0.9$ . Zaś dla  $\gamma_4$  jest ono zawsze bliskie jeden, gdyż punkt  $\gamma_4^*=0$  znajduje się w najbliższym sąsiedztwie modalnej rozkładu gęstości a posteriori. Rezultaty otrzymane dla  $r^*=0.3$  i  $r^*=0.5$  także jednoznacznie wskazują na nieistotną rolę tych zmiennych; jednak takie wartości  $r^*$  nie oddają wstępnej wiedzy. Natomiast położenie funkcji gęstości dla parametru  $\gamma_2$  względem punktu zero zależy silnie od założeń a priori, co przedstawia Rysunek 26. Przyjmując za  $r^*$  wartości 0.9 i 0.7 punkt  $\gamma_2^*=0$  należy do H.P.D. z prawdopodobieństwem przynajmniej na poziomie odpowiednio 0.71 i 0.89. Zatem w przypadku zmiennej  $w_{i2}$  zastosowany test nie rozstrzyga jednoznacznie czy pozostawić tę zmienną w modelu czy ją usunąć. Dokonana analiza pozwala stwierdzić, że im mniejsze wartości przyjmuje parametr  $r^*$  tym to prawdopodobieństwo jest większe, tzn. dla  $r^*=0.5$  wynosi 0.96 natomiast dla  $r^*=0.3$  ponad 0.99. Wyniki te wskazują, że przyjęcie stosunkowo dużej wartości dla mediany rozkładu a priori efektywności powoduje, iż rola zmiennych  $w_{ij}$  wyjaśniających poziom efektywności będzie mniejsza.

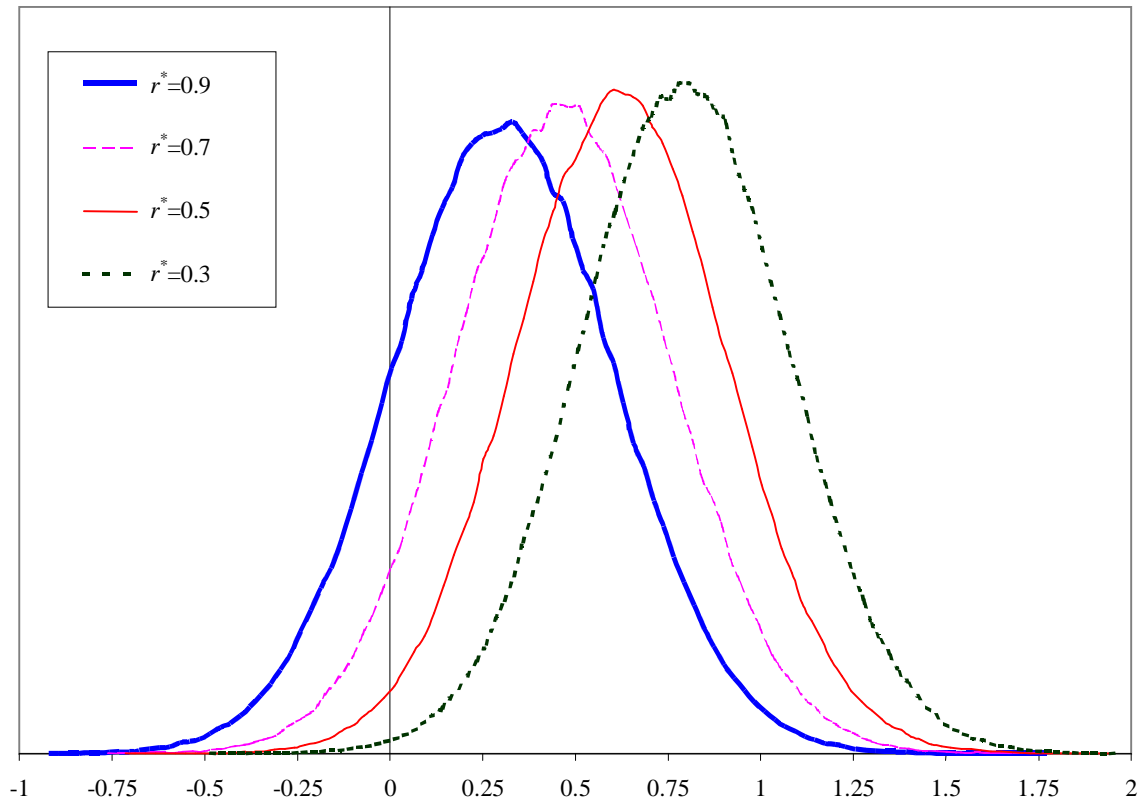
Rysunek 24. Brzegowe gęstości a posteriori składowych wektora  $\gamma$ - VED ( $m=4, r^*=0.7$ ).



Rysunek 25. Brzegowe gęstości a posteriori składowych wektora  $\gamma$ - VED ( $m=4, r^*=0.9$ ).



Rysunek 26. Brzegowe gęstości a posteriori dla  $\gamma_2$  w zależności od  $r^*$  - VED ( $m=4$ ).



Powyższy test H.P.D. należy traktować jako procedurę wstępną, gdyż wykorzystuje jedynie rozkład a posteriori wybranej składowej wektora  $\gamma$ , tj. rozkład brzegowy względem pozostałych składowych, a nie ich łączny rozkład. Warto przypomnieć, że test H.P.D. nie jest „standardowym” testem bayesowskim wykorzystującym iloraz szans a posteriori (*posterior odds ratio*, zob. Zellner [1971], Osiewalski i Pipień [1999]), gdyż nie wprowadza prawdopodobieństwa a priori prawdziwości restrykcji i nie pozwala na obliczenie prawdopodobieństw a posteriori konkurencyjnych modeli (pełnego i po redukcji). Stosowanie ilorazu szans a posteriori jest trudne w realizacji z uwagi na przyjętą metodę numerycznej analizy rozkładów a posteriori, tj. algorytm Gibbsa.

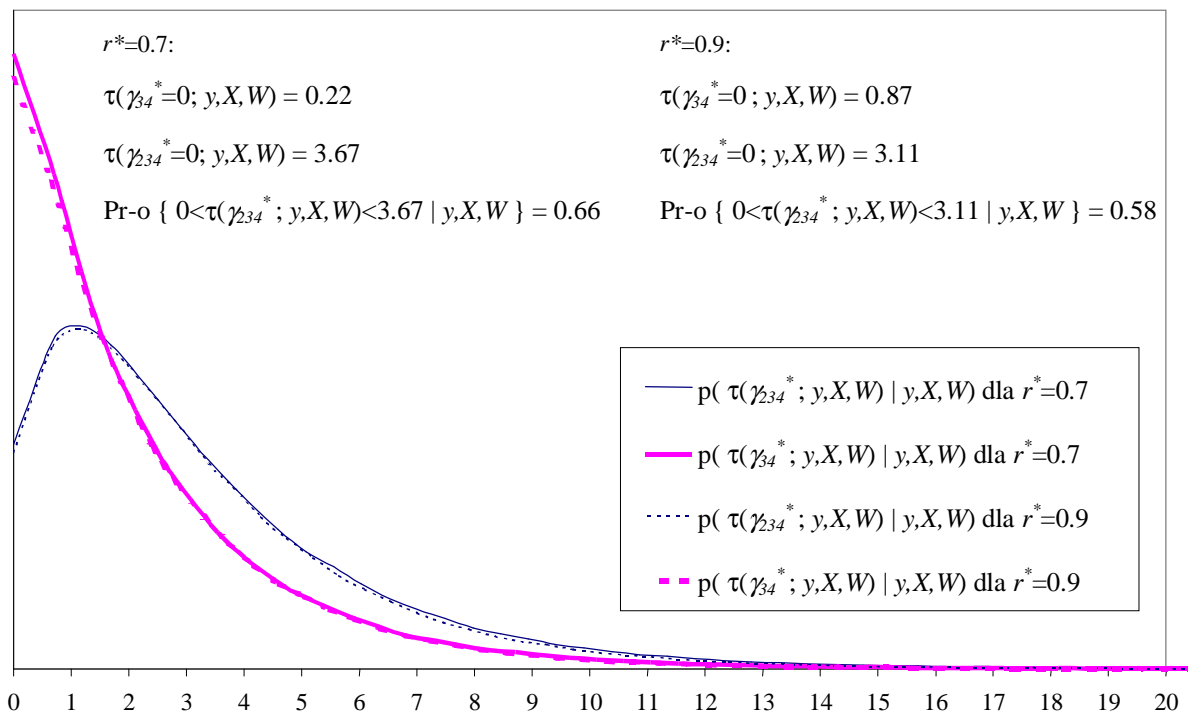
W świetle zaprezentowanych wstępnych wyników testu H.P.D. pojawia się problem wyboru stopnia redukcji modelu VED z  $m=4$ . Możliwa jest pełna redukcja do przypadku CED albo częściowa poprzez usunięcia tylko zmiennych  $w_{i3}$  i  $w_{i4}$ . W celu uzyskania rozstrzygnięcia tego problemu wykorzystano bayesowski odpowiednik testu F, omówiony w podrozdziale 4.4. Hipotezy, które odpowiadają obu redukcjom mają postać:

- $H_0: \gamma_2 = \gamma_3 = \gamma_4 = 0$ , niech zatem  $\gamma_{234}^* = [\gamma_2 \ \gamma_3 \ \gamma_4]'$ , bądź
- $H_0: \gamma_3 = \gamma_4 = 0$ , więc  $\gamma_{34}^* = [\gamma_3 \ \gamma_4]'$

Rezultaty otrzymane za pomocą tego testu prezentuje Rysunek 27, który przedstawia gęstość a posteriori zmiennych losowych  $\tau(\gamma_{234}^*; y, X, W)$  i  $\tau(\gamma_{34}^*; y, X, W)$ . Przeprowadzone badania wskazują, iż dla omawianych wartości  $r^*$  rozkład a posteriori tych zmiennych losowych, unormowanych ze względu na parametr położenia i skali, nie zależy od  $r^*$ . Ten wniosek oparty jest na spostrzeżeniu, że wykresy rozkładów a posteriori odpowiadających różnym  $r^*$  prawie się pokrywają. Wartość funkcji gęstości zmiennej  $\tau$  w punktach  $\gamma_{234}^*=0$  i  $\gamma_{34}^*=0$  wynosi odpowiednio 3.67 i 0.22, gdy  $r^*=0.7$  i 3.12 oraz 0.87 dla  $r^*=0.9$ . Przyjmując  $r^*=0.7$  ( $r^*=0.9$ ) pierwszy punkt leży w przedziałach H.P.D z prawdopodobieństwem przynajmniej 0.66 (0.58), natomiast drugi leży dokładnie w przedziałach o największej gęstości a posteriori z prawdopodobieństwem nie mniejszym niż 0.15. Oznacza to, że  $\gamma_{34}^*=0$  należy do każdego praktycznie sensownego przedziału H.P.D; analogicznie choć nieco słabszą konkluzję można wysunąć dla  $\gamma_{234}^*=0$ . Natomiast dla innych wartości parametru rozkładu efektywności uzyskano następujące wyniki:

	$r^*=0.5$	$r^*=0.3$
$\tau(\gamma_{34}^*=0; y, X, W)$	5.17	8.00
$\tau(\gamma_{234}^*=0; y, X, W)$	0.002	0.27
$\text{Pr-o } \{0 < \tau(\gamma_{234}^*; y, X, W) < \tau(\gamma_{234}^*=0; y, X, W) \mid y, X, W\}$	0.75	0.93
$\text{Pr-o } \{0 < \tau(\gamma_{34}^*; y, X, W) < \tau(\gamma_{34}^*=0; y, X, W) \mid y, X, W\}$	0.001	0.10

Rysunek 27. Gęstości a posteriori zmiennej  $\tau(\gamma^*; y, X, W)$  dla  $\gamma_{234}^*=[\gamma_2 \ \gamma_3 \ \gamma_4]$  i  $\gamma_{34}^*=[\gamma_3 \ \gamma_4]$  - VED ( $m=4$ ).



Prezentowane wyniki wskazują, iż z całą pewnością należy usunąć zmienne  $w_{i3}$  i  $w_{i4}$ , a warto ewentualnie pozostawić w modelu zmienną  $w_{i3}$ , aby ewentualnie zbadać jej istotność w model po redukcji.

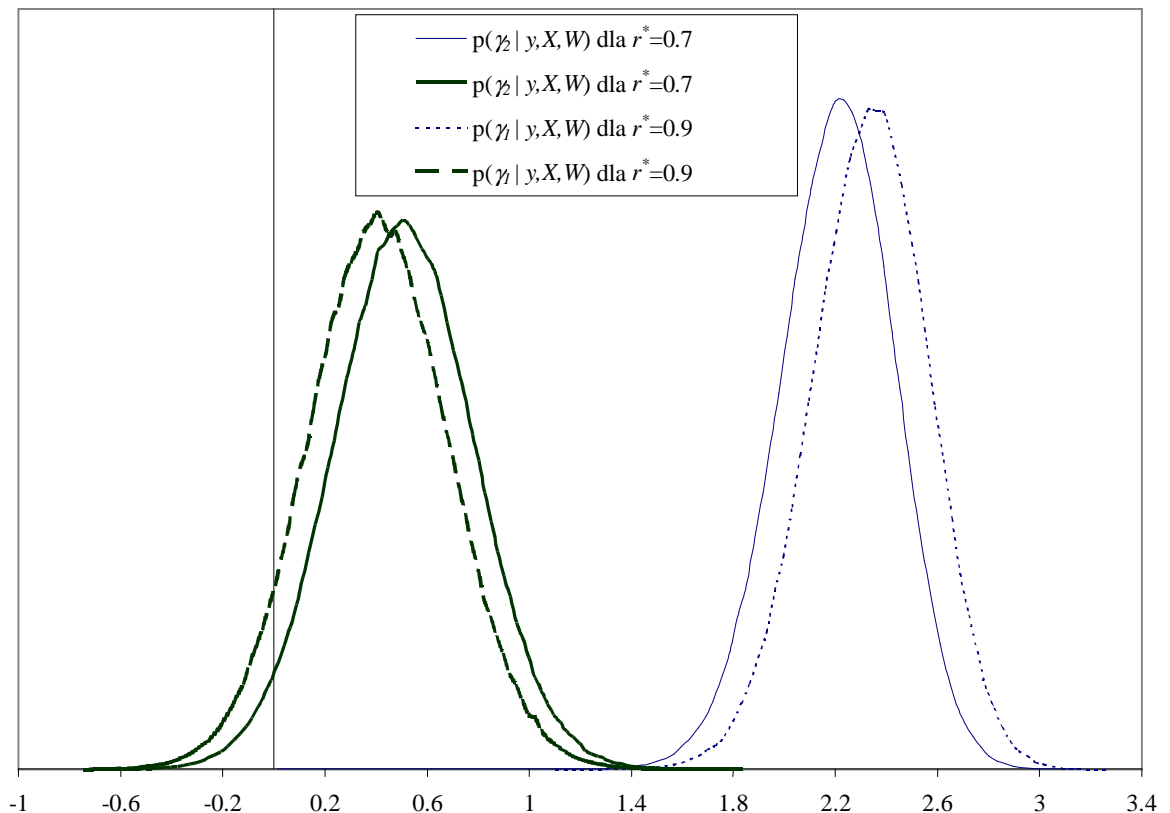
Zatem kolejnym krokiem była redukcja modelu VED z  $m=4$  o zmienne  $w_{i3}$  i  $w_{i4}$ . Wyniki, które w ten sposób uzyskano dla modelu warunkowego opisujące charakteru produkcji, kosztu i poziomu efektywności były identyczne jak w przypadku modeli VED z  $m=4$ , zarówno dla stałej  $r^*$  równej 0.7, jak i 0.9. Jedyne istotne różnice dotyczyły rozkładów a posteriori składowych wektora  $\gamma$ , których główne charakterystyki przedstawione są poniżej:

Tabela 9. Wartości oczekiwane i odchylenia standardowe a posteriori dla  $\gamma$ - VED ( $m=2$ ,  $r^*=0.7$  i  $r^*=0.9$ ).

	$r^*=0.7$		$r^*=0.9$	
	$E(\cdot y,X,W)$	$D(\cdot y,X,W)$	$E(\cdot y,X,W)$	$D(\cdot y,X,W)$
$\gamma_1$	2.205	0.224	2.332	0.228
$\gamma_2$	0.509	0.275	0.409	0.276

Redukcja modelu spowodowała, że rola zmiennej  $w_{i2}$  (specjalizacji na depozytowe i kredytowe) ze statystycznego punktu widzenia istotnie wzrosła. Wyniki te potwierdzają prawidłowość, że parametr  $r^*$  ma istotny wpływ na wartość oczekiwaną a posteriori rozkładu dla  $\gamma_2$  (wraz im  $r^*$  większe, tym wartość tej charakterystyki mniejsza). Pełną wiedzę a posteriori o wektorze  $\gamma$  przedstawia rysunek poniżej:

Rysunek 28. Brzegowe gęstości a posteriori parametrów  $\gamma_1$  i  $\gamma_2$  - VED ( $m=2$ ).



Zastosowany test H.P.D. wskazuje, iż dla  $r^*=0.9$  ( $r^*=0.7$ ) punkt  $\gamma_2^*=0$  należy do przedziału o największej gęstości a posteriori, jeżeli przyjmie się prawdopodobieństwo na poziomie nie niższym niż 0.87 (0.93). Ponieważ najczęściej przyjmowane poziomy prawdopodobieństwa a posteriori odpowiadają klasycznym poziomom ufności: 0.1 lub 0.05 (ewentualnie 0.01), więc w świetle prezentowanych wyników nie można jednoznacznie wnioskować o pozostawieniu bądź usunięciu zmiennej  $w_{i2}$ . Redukcja modelu VED z  $m=4$  przyczyniła się do znacznego zwiększenia precyzji szacunku parametru  $\gamma_2$  (zwłaszcza dla  $r^*=0.9$ ). Jednakże niewykluczona jest redukcja modelu VED z  $m=2$  do modelu CED, aczkolwiek autor skłania się ku pozostawieniu zmiennej  $w_{i2}$  w modelu. Dodatniość parametru  $\gamma_2$  stojącego przy tej zmiennej dychotomicznej oznacza, że oddziały kredytów ( $w_{i2}=1$ ) wydają się systematycznie bardziej efektywniejsze, tj. mają średnio niższy poziom nieuzasadnionych kosztów zmiennych.

W celu uzupełnienia prezentowanych wyników przeprowadzono również estymację modelu CED, który uzyskuje się w poprzez usunięcie z modelu zmiennej omawianej zmiennej  $w_{i2}$  informującej o specjalizacji oddziałów. Otrzymane wyniki dotyczące opisujące technologię produkcji i kosztu oraz ranking efektywności nie różnią się istotnie zarówno od tych uzyskanych dla modeli VED z  $m=4$  i VED z  $m=2$ . Podstawowe wyniki dla CED prezentowane są w Załączniku 2, w Tabelach 18-20.

#### **5.4. Estymacja modelu bayesowskiego modelu VED – przypadek dwóch produktów.**

Dotychczas prezentowane wyniki dotyczyły przypadku jednoproduktowej granicznej funkcji kosztu i stanowiły ilustrację empiryczną bayesowskiej estymacji stochastycznych modeli granicznych. Jednakże pogłębiona analiza ekonomiczna procesu produkcji poprzez funkcję kosztów, obejmująca analizę zakresu produkcji bądź specjalizacji oddziałów, wymaga rozważenia przynajmniej dwóch produktów. W dalszej części podjęto się próbę bayesowskiej estymacji dwuproduktowej granicznej funkcji kosztu, rozróżniając:

- kredyty komercyjne (zmienna  $Q_1$ ),
- kredyty detaliczne (zmienna  $Q_2$ )<sup>22</sup>.

---

<sup>22</sup> Szerzej są one omówione w podrozdziale 5.1.



Analiza uwzględniająca większą liczbę produktów (np. dodatkowy podział na kredyty złotowe i walutowe) wykracza poza zakres niniejszej pracy. Prezentowane są jedynie najważniejsze wyniki dotyczące modelu dwuproduktowego, zwłaszcza te dotyczące analizy zakresu produkcji bądź specjalizacji oddziałów.

Dwuproduktowa krótkookresowa graniczna funkcja kosztów zmiennych - przy założeniu, że efektywność jest czystym efektem indywidualnym dla każdej z firm ( $z_{it}=z_i$ ) - przybiera postać:

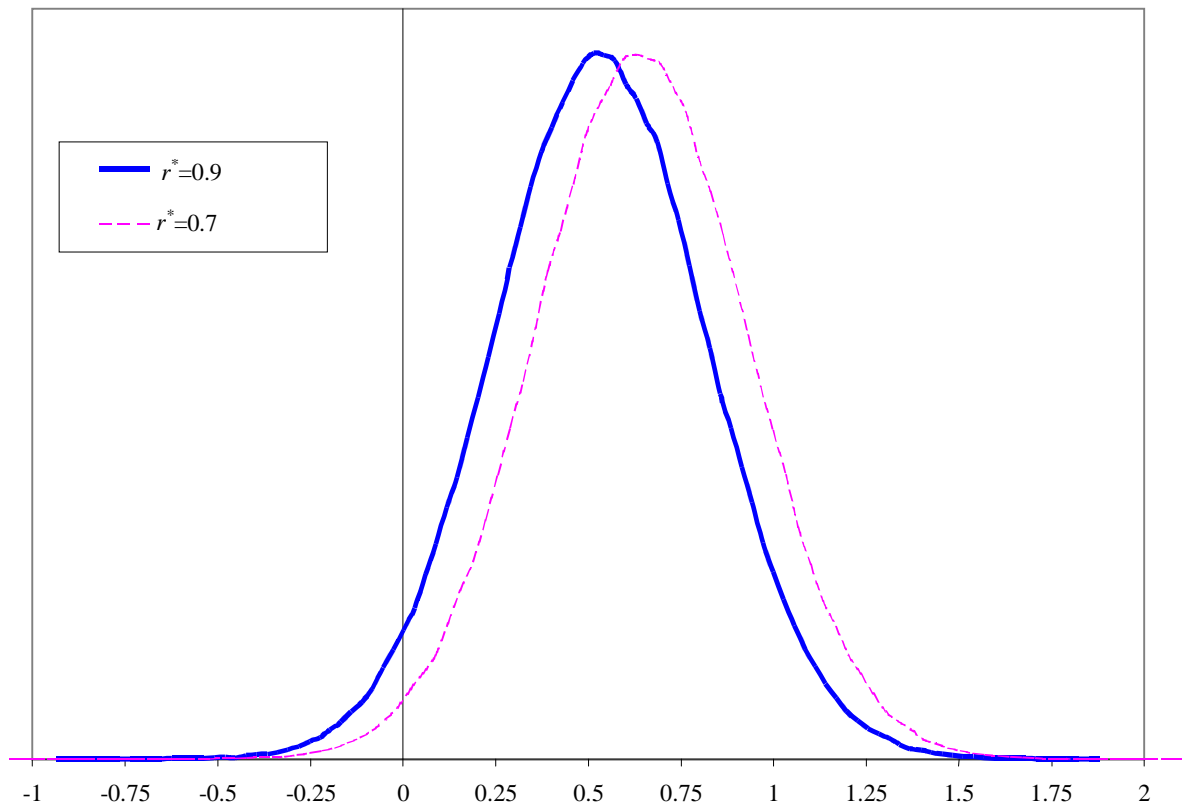
$$\begin{aligned} \ln VC_{it} = & \delta_0 + \delta_1 \ln w_{it,D} + \delta_2 \ln w_{it,L} + \delta_3 \ln Q_{it,1} + \delta_4 \ln K_i + \delta_5 \ln w_{it,D} \ln w_{it,L} \\ & + \delta_6 \ln w_{it,D} \ln Q_{it,1} + \delta_7 \ln w_{it,D} \ln K_i + \delta_8 \ln w_{it,L} \ln Q_{it,1} + \delta_9 \ln w_{it,L} \ln K_i \\ & + \delta_{10} \ln Q_{it,1} \ln K_i + \delta_{11} (\ln w_{it,D})^2 + \delta_{12} (\ln w_{it,L})^2 + \delta_{13} (\ln Q_{it,1})^2 + \delta_{14} (\ln K_i)^2 \\ & + \delta_{15} \ln Q_{it,2} + \delta_{16} \ln w_{it,D} \ln Q_{it,2} + \delta_{17} \ln w_{it,L} \ln Q_{it,2} + \delta_{18} \ln Q_{it,1} \ln Q_{it,2} \\ & + \delta_{19} \ln Q_{it,2} \ln K_i + \delta_{20} (\ln Q_{it,2})^2 + v_{it} + z_i, \end{aligned} \quad (75)$$

Punktem wyjścia estymacji powyższego równania było przyjęcie bayesowskiego modelu VED z  $m=4$ , z którego na drodze redukcji otrzymano model VED z  $m=2$ , w którym jedyną zmienną posiadającą istotny wpływ na zróżnicowanie poziomu efektywności oddziałów jest zmienna wyrażająca ich specjalizację ( $w_{i2}$ ). Wyniki estymacji modelu dwuproduktowego, tj. informujące o charakterze technologii oddziałów, były prawie identyczne, jak te uzyskane dla przypadku z jednym produktem, obszernie przedstawione we wcześniejszym podrozdziale. Jedynie z uwagi na większą liczbę parametrów w równaniu (75) można było spodziewać się większych wartości odchyłeń standardowych dla badanych charakterystyk, co potwierdziły wyniki empiryczne. Prezentowane dalej wyniki uzyskano wyniki w oparciu o  $5 \cdot 10^5$  cykli Gibbsa, a algorytm uzyskał zbieżność po około 100 tysiącach dodatkowych cykli<sup>23</sup>.

Ze statystycznego punktu widzenia wnioskowanie o istotnej roli zmiennej  $w_{i2}$  w modelu VED z  $m=2$ , przy użyciu testu H.P.D. oraz bayesowskiego odpowiednika testu F, jest bardziej uzasadnione niż w przypadku jednoproduktowej funkcji kosztu. Test Lindleya wskazuje, iż punkt  $\gamma_2^*=0$  należy do przedziału o największej gęstości a posteriori dla prawdopodobieństwa a posteriori równych przynajmniej 0.97 ( $r^*=0.7$ ) lub 0.94 ( $r^*=0.9$ ). A zatem przyjęcie restrykcji  $\gamma_2=0$  nie miałyby silnego uzasadnienia w danych. Odpowiednie brzegowe gęstości a posteriori dla parametru  $\gamma_2$  przedstawione są na rysunku poniżej.

<sup>23</sup> W omawianym modelu VED ( $m=2$ ) z dwoma produktami komputer klasy Pentium II 400 Mhz na wykonanie  $6 \cdot 10^5$  cykli Gibbsa potrzebował ponad 14 godziny.

Rysunek 29. Brzegowe gęstości a posteriori dla  $\gamma_2$  w zależności od  $r^*$  - VED ( $m=2, Q_1$  i  $Q_2$ ).



Analizując ranking efektywności oddziałów (Tabela 10) można zauważyć, że uległ on pewnym zmianom w stosunku do modelu z jednym zagregowanym produktem, aczkolwiek są one niewielkie, co potwierdza współczynnik korelacji rang Spearmana, który wynosi odpowiednio 0.83 i 0.84 dla  $r^*=0.7$  oraz 0.9. Jednakże ranking efektywności nie zależy od założeń a priori tj. od stałej  $r^*$ . Badane oddziały charakteryzują się dość wysoką efektywnością, a średnia wartość oczekiwana efektywności wszystkich oddziałów wynosi 0.928 ( $\pm 0.018$ ), gdy tymczasem dla modelu z jednym produktem wynosiła 0.918 ( $\pm 0.017$ ). Nieznaczne podwyższenie oceny efektywności wiąże się z większą liczbą zmiennych, a więc z lepszym dopasowaniem modelu dwuproduktowego do danych. Zamiana miejsc w rankingu w przypadku dwóch najefektywniejszych oddziałów (o numerze 56 i 57) jest nieistotna ze statystycznego punktu widzenia z uwagi na niewielkie różnice w efektywności i w stosunku do precyzji ich szacowania. Duże różnice miejsc w rankingu dotyczą przede wszystkim oddziałów o efektywności bliskiej przeciętnej. Obszerne wyniki porównujące oba modele przedstawia są poniżej.

Tabela 10. Wartości oczekiwane i odchylenia standardowe a posteriori wskaźników efektywności modelu VED ( $m=2, r^*=0.7$ ).

$i$	Model z $Q_1$ i $Q_2$			Model z $Q = Q_1 + Q_2$			Różnice w rankingu
	$E(r_i/y, X)$	$D(r_i/y, X)$	Ranking	$E(r_i/y, X)$	$D(r_i/y, X)$	Ranking	
1	0.886	0.031	45	0.905	0.034	37	8
2	0.836	0.028	56	0.901	0.029	39	17
3	0.936	0.043	32	0.876	0.051	46	-14
4	0.877	0.019	49	0.900	0.020	40	9
5	0.860	0.020	51	0.893	0.021	42	9
6	0.957	0.020	22	0.929	0.019	30	-8
7	0.991	0.008	3	0.984	0.014	5	-2
8	0.835	0.017	57	0.845	0.017	51	6
9	0.980	0.016	12	0.965	0.023	11	1
10	0.918	0.017	38	0.898	0.019	41	-3
11	0.875	0.018	50	0.853	0.017	49	1
12	0.955	0.025	23	0.889	0.024	44	-21
13	0.883	0.017	47	0.835	0.014	55	-8
14	0.894	0.050	44	0.841	0.033	53	-9
15	0.842	0.014	55	0.841	0.014	52	3
16	0.965	0.017	19	0.961	0.016	15	4
17	0.981	0.013	10	0.977	0.015	7	3
18	0.980	0.017	11	0.977	0.020	8	3
19	0.859	0.016	52	0.836	0.016	54	-2
20	0.879	0.019	48	0.853	0.015	50	-2
21	0.983	0.012	8	0.964	0.016	12	-4
22	0.897	0.014	43	0.901	0.015	38	5
23	0.974	0.021	14	0.920	0.027	32	-18
24	0.926	0.014	36	0.887	0.014	45	-9
25	0.905	0.015	41	0.891	0.014	43	-2
26	0.958	0.023	20	0.983	0.012	6	14
27	0.844	0.013	54	0.831	0.014	57	-3
28	0.915	0.021	39	0.930	0.024	29	10
29	0.911	0.017	40	0.936	0.016	27	13
30	0.902	0.020	42	0.855	0.014	48	-6
31	0.972	0.014	15	0.953	0.015	20	-5
32	0.938	0.017	30	0.910	0.016	34	-4
33	0.945	0.018	26	0.973	0.013	9	17
34	0.937	0.016	31	0.963	0.014	14	17
35	0.949	0.015	25	0.964	0.015	13	12
36	0.935	0.016	33	0.957	0.015	18	15
37	0.850	0.013	53	0.833	0.013	56	-3
38	0.954	0.014	24	0.944	0.015	24	0
39	0.966	0.016	17	0.946	0.015	23	-6
40	0.940	0.018	28	0.908	0.014	36	-8
41	0.933	0.024	34	0.910	0.014	33	1
42	0.939	0.019	29	0.971	0.017	10	19
43	0.965	0.017	18	0.943	0.014	25	-7
44	0.988	0.009	5	0.961	0.014	16	-11
45	0.985	0.011	6	0.992	0.007	3	3
46	0.982	0.013	9	0.947	0.014	22	-13
47	0.978	0.013	13	0.940	0.014	26	-13
48	0.942	0.014	27	0.909	0.013	35	-8
49	0.924	0.019	37	0.957	0.016	19	18

<i>i</i>	Model z $Q_1$ i $Q_2$			Model z $Q = Q_1 + Q_2$			Różnice w rankingu
	$E(r_i / y, X)$	$D(r_i / y, X)$	Ranking	$E(r_i / y, X)$	$D(r_i / y, X)$	Ranking	
50	0.983	0.011	7	0.961	0.013	17	-10
51	0.933	0.018	35	0.920	0.016	31	4
52	0.958	0.014	21	0.948	0.013	21	0
53	0.885	0.017	46	0.870	0.015	47	-1
54	0.972	0.018	16	0.931	0.014	28	-12
55	0.988	0.010	4	0.986	0.012	4	0
<b>56</b>	<b>0.993</b>	<b>0.007</b>	<b>2</b>	<b>0.996</b>	<b>0.004</b>	<b>1</b>	<b>1</b>
<b>57</b>	<b>0.994</b>	<b>0.006</b>	<b>1</b>	<b>0.995</b>	<b>0.005</b>	<b>2</b>	<b>-1</b>
<b>58</b>	<b>0.698</b>	<b>0.023</b>	<b>58</b>	<b>0.769</b>	<b>0.016</b>	<b>58</b>	<b>0</b>

Wartości oczekiwane i odchylenia standardowe a posteriori parametrów granicznej funkcji kosztów danej równaniem (75) przedstawia Tabela 11. Oczywiście, z uwagi na inną specyfikację zmiennych, różnią się one od tych uzyskanych dla modelu z jednym produktem, lecz nie zależą od różnych wartości stałej  $r^*$ . Także średnie wartości oczekiwane elastyczności kosztu względem cen czynników produkcji i nakładów czynnika stałego i odpowiednie średnie odchylenia standardowe a posteriori dla obu modeli są prawie identyczne. Rozważany model pozwala m.in. na określenie wpływu zmiany wielkości poszczególnych produktów na zmianę kosztu zmiennego. Spośród obu produktów wielkość kredytów komercyjnych ma większy wpływ na poziom kosztów zmiennych, tj. ich wyłączny wzrost o 1% (wartościowo o około 1 063 tys. zł średnio w oddziale) powoduje wzrost kosztu zmiennego obserwowanego w oddziałach przeciętnie o 0.75% (35 tys. zł) z błędem  $\pm 0.023\%$ . Natomiast jednoprocenowy wzrost wielkości kredytów detalicznych, czyli o około 113 tys. zł, powoduje wzrost tych kosztów średnio tylko o 0.12% ( $\pm 0.017\%$ ) – wartościowo 5.6 tys. zł. Natomiast łączny wpływ zmiany wielkości obu produktów równocześnie o 1% powoduje wzrost kosztu zmiennego średnio o 0.87% ( $\pm 0.04\%$ ) *ceteris paribus*. Natomiast w przypadku modelu z jednym produktem średnia wartość oczekiwana elastyczności kosztu względem zagregowanego produktu kształtuje się na zbliżonym poziomie rzędu 0.86, a średnie odchylenie standardowe wynosi 0.025. Rysunek 30 przedstawia gęstości a posteriori elastyczności kosztu względem każdego z produktów i ich agregatu. W świetle tych wyników spodziewać się można, że większość oddziałów charakteryzuje się rosnącym efektem skali produkcji względem czynników zmiennych. Jego miara, tzw. zagregowany współczynnik skali produkcji (por. Równanie (11)) jako odwrotność sumy elastyczności względem produktów informuje, że zmiana kosztu zmiennego o 1% spowodowana proporcjonalnym wzrostem nakładów czynników zmiennych, tj. liczby pracowników oraz wartości depozytów i

innych pozyskanych środków pieniężnych przy ustalonej wielkości nakładów czynnika stałego (powierzchni biurowej), powoduje w przypadku przeciętnego oddziały proporcjonalny wzrost każdego z produktów (kredytów komercyjnych i detalicznych) średnio o 1.15%.

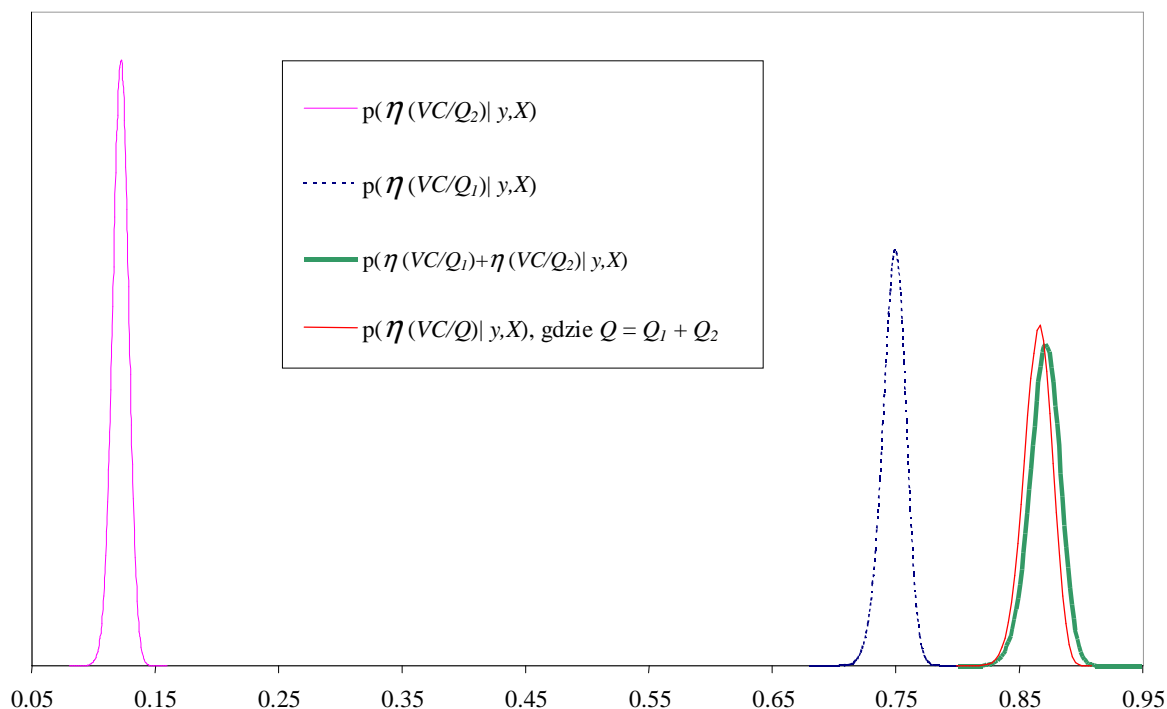
Tabela 11. Wartości oczekiwane i odchylenia standardowe a posteriori parametrów modelu - VED ( $m=2$ ,  $r^*=0.7$  i  $0.9$ ,  $Q_1$  i  $Q_2$ ).

Parametr	Zmienna	$r^*=0.7$		$r^*=0.9$	
		$E(\cdot y,X)$	$D(\cdot y,X)$	$E(\cdot y,X)$	$D(\cdot y,X)$
$\delta_0$	Stała	6.115	3.531	6.193	3.509
$\delta_1$	$\ln w_D$	0.329	0.590	0.344	0.589
$\delta_2$	$\ln w_L$	-0.974	0.718	-0.968	0.721
$\delta_3$	$\ln Q_1$	0.666	0.212	0.651	0.208
$\delta_4$	$\ln K$	0.038	0.187	0.042	0.183
$\delta_5$	$\ln w_D \ln w_L$	-0.016	0.057	-0.016	0.057
$\delta_6$	$\ln w_D \ln Q_1$	0.010	0.027	0.011	0.027
$\delta_7$	$\ln w_D \ln K$	0.080	0.023	0.080	0.023
$\delta_8$	$\ln w_L \ln Q_1$	-0.050	0.015	-0.051	0.015
$\delta_9$	$\ln w_L \ln K$	-0.004	0.013	-0.004	0.013
$\delta_{10}$	$\ln Q_1 \ln K$	-0.005	0.009	-0.006	0.009
$\delta_{11}$	$(\ln w_L)^2$	-0.033	0.044	-0.034	0.044
$\delta_{12}$	$(\ln w_D)^2$	0.145	0.048	0.145	0.048
$\delta_{13}$	$(\ln Q_1)^2$	0.032	0.007	0.032	0.007
$\delta_{14}$	$(\ln K)^2$	0.034	0.008	0.034	0.007
$\delta_{15}$	$\ln Q_2$	-0.453	0.183	-0.454	0.183
$\delta_{16}$	$\ln w_D \ln Q_2$	0.001	0.024	0.000	0.024
$\delta_{17}$	$\ln w_L \ln Q_2$	-0.010	0.014	-0.009	0.014
$\delta_{18}$	$\ln Q_1 \ln Q_2$	-0.042	0.009	-0.042	0.009
$\delta_{19}$	$\ln K \ln Q_2$	-0.035	0.010	-0.034	0.010
$\delta_{20}$	$(\ln Q_2)^2$	0.053	0.005	0.052	0.005
$D^2(v_{i,t}) = \sigma_v^2$		0.000421	0.000054	0.000425	0.000055

Tabela 12. Średnia wartości oczekiwane i odchylenia standardowe a posteriori elastyczności - VED ( $m=2$ ,  $r^*=0.7$  i  $0.9$ ,  $Q_1$  i  $Q_2$ ).

	$\eta(VC/w_D)$	$\eta(VC/w_L)$	$\eta(VC/Q_1)$	$\eta(VC/K)$	$\eta(VC/Q_2)$	$\eta(VC/Q_1) + \eta(VC/Q_2)$	$1 - \eta(VC/w_L) - \eta(VC/w_D)$
$r^*=0.7$							
Srednia wartość oczekiwana	0.808	0.178	0.748	0.043	0.122	0.870	0.014
Srednie odchylenie standardowe	0.045	0.031	0.023	0.022	0.017	0.040	0.076
$r^*=0.9$							
Srednia wartość oczekiwana	0.808	0.178	0.750	0.043	0.123	0.873	0.015
Srednie odchylenie standardowe	0.044	0.031	0.023	0.021	0.017	0.040	0.076

Rysunek 30. Rozkład a posteriori elastyczności względem produktów  $Q_1$  i  $Q_2$  oraz  $Q$  - VED ( $m=2, r^*=0.7, Q_1$  i  $Q_2$ ) i VED ( $m=2, r^*=0.7, Q$ )



Prezentowane powyżej informacje o charakterze procesu produkcji mogą być w praktyce wykorzystane m.in. w analizie bieżącej działalności oddziałów, mającej na celu obniżkę kosztów działalności i określenie głównych źródeł ich powstawania. Natomiast analiza specjalizacji lub badania ewentualnych korzyści zakresu pełnego produkcji wydaje się być pomocna w tworzeniu polityki rozwoju placówek badanego banku. Tabela 13, prezentując m.in. strukturę produkcji, pozwala na analizę wpływu zmiany wartości kredytów komercyjnych i detalicznych na koszt zmienny w wybranych oddziałach. Szczególną uwagę zwracają trzy oddziały o numerach 7, 41 i 43, dla których wartości oczekiwane a posteriori elastyczności kosztu względem kredytów detalicznych są ujemne oraz te, dla których co prawda są dodatnie (bliskie zero), lecz charakteryzują się dużym odchyleniem standardowym a posteriori. Należy w tym momencie przypomnieć, że prezentowane wyniki uzyskano poprzez narzucenie restrykcji w sposób identyczny jak w modelu z jednym produktem (agregatem), tj. dla średnich elastyczności względem czynników i produktów dla wszystkich oddziałów. Próba narzucenia warunku dodatnich elastyczności dla każdego oddziału spowodowała drastyczny spadek efektywności losowania Gibbsa. Spowodowane było to tym, iż rozkład a posteriori dla elastyczności względem kredytów detalicznych ( $Q_2$ ) o wartości oczekiwanej  $-0.121$  i małym odchyleniem standardowym  $0.015$ , uzyskany bez narzucania

restrykcji dla oddziału o numerze 7 wskazuje, iż dla tego oddziału była ona nieuzasadniona w świetle danych. Ujemne oceny dla elastyczności względem kredytów detalicznych obserwuje się wśród oddziałów, które specjalizują się w udzielaniu kredytów komercyjnych ( $Q_1$ ), np. w przypadku omawianego oddziału o numerze 7 kredyty detaliczne stanowią tylko 0.5% sumy kredytów ogółem i ewentualnej nadwyżki depozytów nad kredyty. Natomiast specjalizacja oddziałów (np. o numerze 14 i 26) w udzielaniu kredytów detalicznych ma odzwierciedlenie w wysokiej wartości oceny tych elastyczności. Należy wyjaśnić, iż posiadanie przez oddział o numerze 14 tak dużego wolumenu kredytów detalicznych jest efektem współpracy z wyspecjalizowaną firmą pośredniczącą, która zajmuje się techniczną stroną przedsięwzięcia, tj. obsługą klientów detalicznych – potencjalnych kredytobiorców (przyjmowanie wniosków kredytowych, badanie wiarygodności i wypłacalności klientów itp.).

Dodatkowo w celu pełniej prezentacji omawianych charakterystyk przedstawiono wartości oczekiwane i wielkość odchyłeń standardowych a posteriori dla elastyczności traktowanych jako funkcje tylko jednej zmiennej przyjmując wartości pozostałych zmiennych na poziomie średnich geometrycznych z próby. Rysunek 33 potwierdza uzyskane wcześniej wyniki dotyczące ujemnych wartości oczekiwanych elastyczności względem kredytów detalicznych w przypadku oddziałów specjalizujących się w udzielaniu kredytów klientom komercyjnym.

Tabela 13. Wartości oczekiwane, odchylenia standardowe elastyczności kosztu i zagregowany współczynnik skali produkcji dla poszczególnych oddziałów względem produktów:  $SL(Q_1, Q_2) = (\eta(Q_1)_i + \eta(Q_2)_i)^{-1}$  - VED ( $m=2$ ,  $r^*=0.7$ ,  $Q_1$  i  $Q_2$ ).

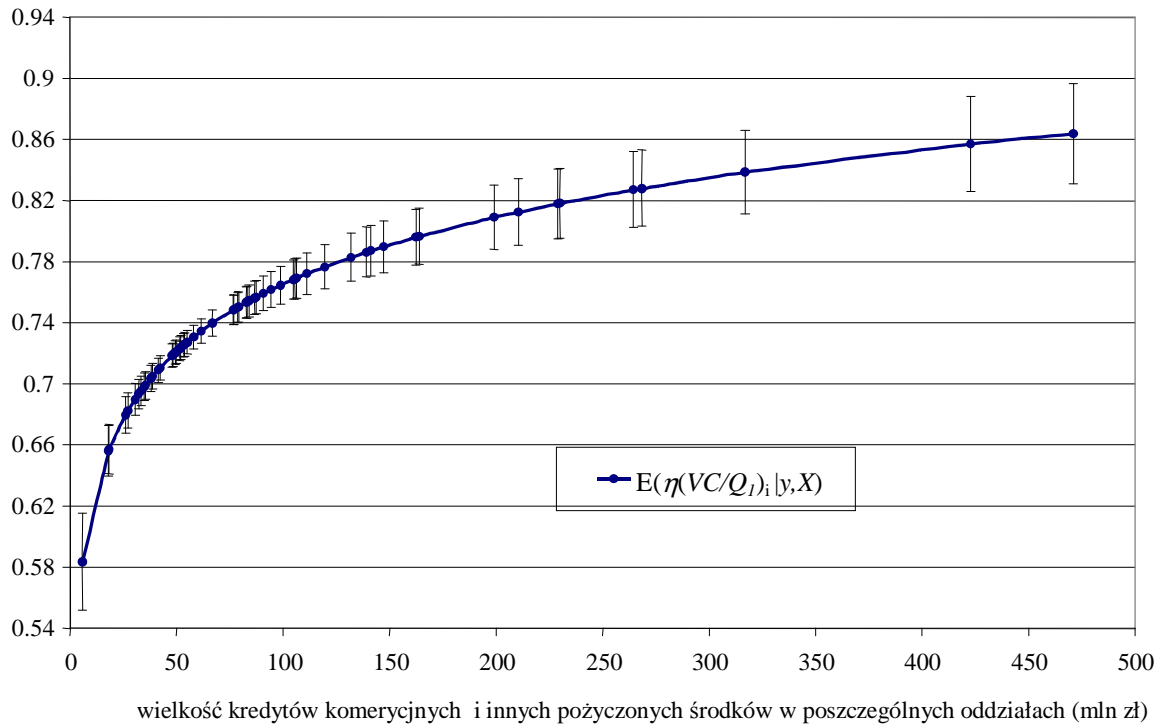
$i$	$E(\eta(Q_1)_i \cdot)$	$D(\eta(Q_1)_i \cdot)$	$E(\eta(Q_2)_i \cdot)$	$D(\eta(Q_2)_i \cdot)$	$E(\eta(Q_1)_i + \eta(Q_2)_i \cdot)$	$D(\eta(Q_1)_i + \eta(Q_2)_i \cdot)$	$SL_0^{24}$	$Q_2/(Q_1+Q_2)$
1	0.834	0.032	0.086	0.020	0.919	0.052	1.087	3.0%
2	0.864	0.030	<b>0.007</b>	<b>0.019</b>	0.871	0.049	1.148	<b>1.3%</b>
3	0.800	0.030	0.223	0.019	1.023	0.049	<b>0.978</b>	6.6%
4	0.803	0.031	0.078	0.019	0.881	0.050	1.135	3.9%
5	0.841	0.030	<b>0.001</b>	<b>0.015</b>	0.841	0.045	1.188	<b>1.7%</b>
6	0.764	0.030	0.150	0.014	0.914	0.045	1.094	8.0%
7	<b>0.881</b>	<b>0.030</b>	<b>-0.121</b>	<b>0.015</b>	<b>0.760</b>	<b>0.045</b>	<b>1.316</b>	<b>0.5%</b>
8	0.789	0.030	0.097	0.015	0.886	0.045	1.129	4.9%
9	0.773	0.032	0.108	0.027	0.881	0.059	1.135	7.1%
10	0.756	0.032	0.127	0.027	0.883	0.059	1.133	9.1%
11	0.769	0.032	0.157	0.027	0.926	0.060	1.080	7.9%
12	0.750	0.032	0.162	0.026	0.912	0.058	1.096	11.6%
13	0.724	0.022	0.195	0.013	0.919	0.035	1.088	15.6%
14	0.633	0.022	<b>0.449</b>	0.013	1.082	0.034	<b>0.924</b>	<b>62.6%</b>
15	0.775	0.022	0.088	0.013	0.864	0.034	1.159	5.9%

<sup>24</sup> Przyjęcie, że elastyczności  $\eta(Q_1)_i$  i  $\eta(Q_2)_i$  posiadają rozkład zbliżony do normalnego powoduje, że momenty zmiennej losowej  $(\eta(Q_1)_i + \eta(Q_2)_i)^{-1}$  nie istnieją, w tym wartość oczekiwana  $E((\eta(Q_1)_i + \eta(Q_2)_i)^{-1})$ , zatem za ocenę  $SL(Q_1, Q_2)$  przyjęto wyrażenie  $(E(\eta(Q_1)_i|\cdot) + E(\eta(Q_2)_i|\cdot))^{-1}$

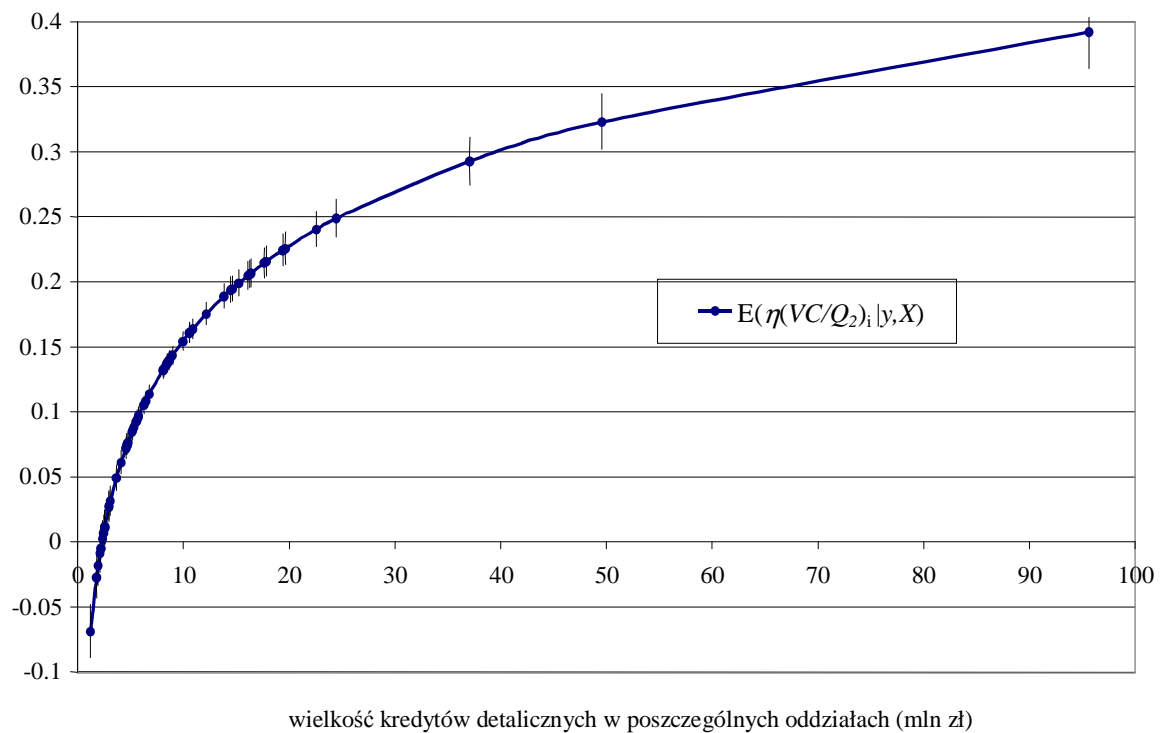
$i$	$E(\eta(Q_1)_i  \cdot)$	$D(\eta(Q_1)_i  \cdot)$	$E(\eta(Q_2)_i  \cdot)$	$D(\eta(Q_2)_i  \cdot)$	$E(\eta(Q_1)_i + \eta(Q_2)_i  \cdot)$	$D(\eta(Q_1)_i + \eta(Q_2)_i  \cdot)$	SL0 <sup>24</sup>	$Q_2/(Q_1+Q_2)$
16	0.803	0.021	<b>0.011</b>	<b>0.012</b>	0.814	0.034	1.229	<b>3.1%</b>
17	0.785	0.025	0.053	0.015	0.838	0.041	1.193	4.4%
18	0.725	0.026	0.133	0.014	0.858	0.040	1.166	13.3%
19	0.759	0.026	0.157	0.015	0.916	0.041	1.092	8.7%
20	0.723	0.026	0.171	0.015	0.894	0.041	1.119	15.1%
21	0.758	0.019	0.116	0.013	0.875	0.032	1.144	8.1%
22	0.783	0.020	0.067	0.014	0.850	0.033	1.176	4.8%
23	0.746	0.020	0.219	0.015	0.965	0.035	1.036	11.5%
24	0.723	0.019	0.166	0.015	0.889	0.035	1.125	14.5%
25	0.751	0.032	0.146	0.020	0.897	0.052	1.115	8.1%
26	0.628	0.031	0.315	0.019	0.942	0.050	1.060	52.2%
27	0.747	0.033	0.106	0.020	0.853	0.053	1.172	8.6%
28	0.765	0.030	0.154	0.021	0.918	0.050	1.088	6.9%
29	0.777	0.018	0.094	0.011	0.871	0.029	1.148	4.7%
30	0.655	0.019	0.321	0.011	0.976	0.030	1.025	42.7%
31	0.753	0.020	0.094	0.012	0.847	0.032	1.181	7.9%
32	0.693	0.019	0.172	0.011	0.865	0.031	1.156	21.4%
33	0.807	0.020	<b>0.015</b>	<b>0.016</b>	0.822	0.036	1.217	<b>2.4%</b>
34	0.792	0.019	0.033	0.016	0.825	0.034	1.212	3.1%
35	0.779	0.020	0.066	0.016	0.845	0.036	1.183	4.6%
36	0.805	0.019	<b>0.014</b>	<b>0.016</b>	0.820	0.035	1.221	<b>2.8%</b>
37	0.750	0.017	0.072	0.015	0.822	0.032	1.217	7.2%
38	0.745	0.016	0.136	0.015	0.881	0.032	1.135	9.7%
39	0.767	0.017	0.032	0.015	0.799	0.032	1.252	5.1%
40	0.725	0.017	0.089	0.015	0.814	0.031	1.229	11.5%
41	<b>0.786</b>	<b>0.017</b>	<b>-0.014</b>	<b>0.012</b>	<b>0.773</b>	<b>0.029</b>	<b>1.295</b>	<b>3.2%</b>
42	0.768	0.018	0.101	0.013	0.869	0.031	1.151	5.7%
43	<b>0.776</b>	<b>0.017</b>	<b>-0.002</b>	<b>0.012</b>	<b>0.773</b>	<b>0.030</b>	<b>1.292</b>	<b>3.6%</b>
44	0.708	0.017	0.154	0.011	0.862	0.029	1.160	16.4%
45	0.755	0.020	0.079	0.018	0.834	0.039	1.199	5.8%
46	0.679	0.020	0.256	0.018	0.934	0.038	1.070	29.7%
47	0.692	0.022	0.189	0.019	0.881	0.040	1.135	20.8%
48	0.719	0.025	0.107	0.020	0.825	0.045	1.211	12.9%
49	0.769	0.014	0.092	0.015	0.862	0.029	1.161	5.8%
50	0.691	0.016	0.205	0.016	0.896	0.033	1.116	21.1%
51	0.719	0.014	0.162	0.015	0.882	0.029	1.135	13.1%
52	0.750	0.015	0.062	0.014	0.812	0.029	1.232	6.7%
53	0.692	0.028	0.235	0.030	0.927	0.058	1.079	23.7%
54	0.733	0.029	0.035	0.032	0.769	0.061	1.302	7.1%
55	0.708	0.031	0.185	0.033	0.893	0.064	1.120	16.6%
56	0.646	0.031	0.176	0.033	0.822	0.063	1.217	32.2%
57	0.710	0.016	0.093	0.011	0.804	0.027	1.245	10.5%
58	0.598	0.016	0.198	0.011	0.795	0.026	1.256	47.4%



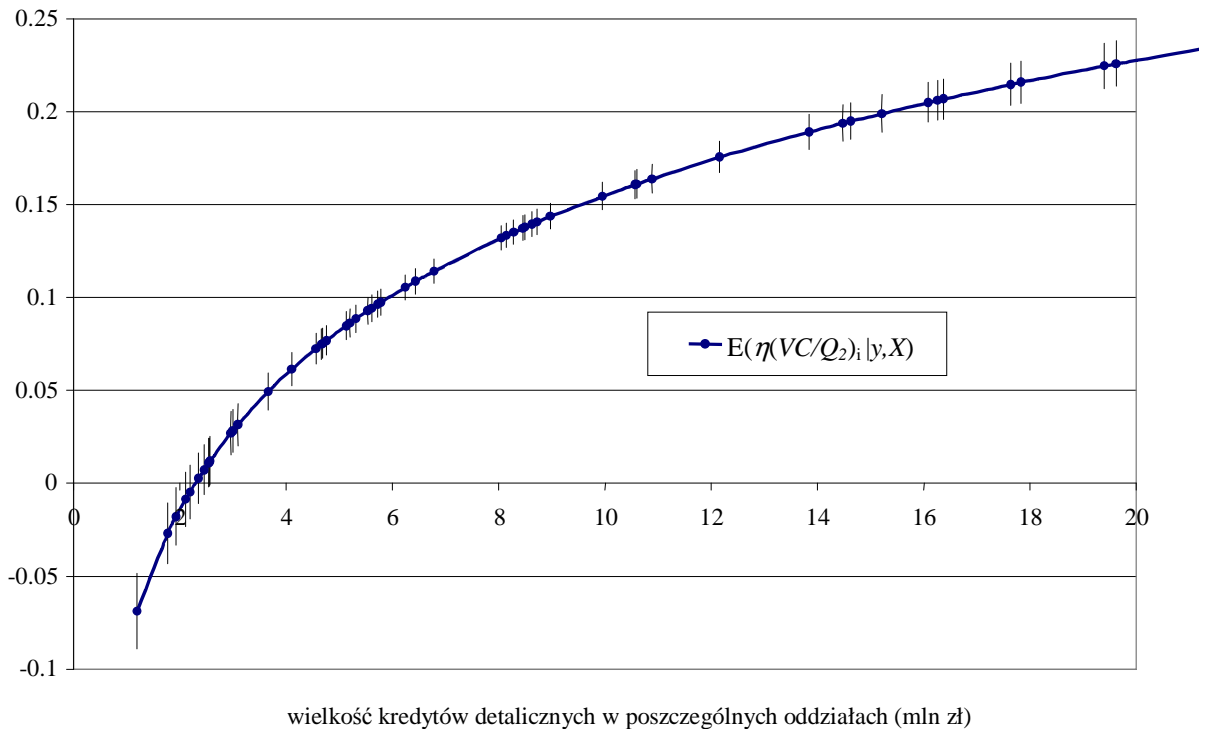
Rysunek 31. Wartości oczekiwane i odchylenia standardowe a posteriori elastyczności kosztu względem produktu  $Q_1$  dla poszczególnych oddziałów (przy ustalonych wartościach pozostałych zmiennych egzogenicznych na poziomie średniej geometrycznej z próby) - VED ( $m=2, r^*=0.7, Q_1$  i  $Q_2$ ).



Rysunek 32. Wartości oczekiwane i odchylenia standardowe a posteriori elastyczności kosztu względem produktu  $Q_2$  dla poszczególnych oddziałów (przy ustalonych wartościach pozostałych zmiennych egzogenicznych na poziomie średnich geometrycznych z próby) - VED ( $m=2, r^*=0.7, Q_1$  i  $Q_2$ ).



Rysunek 33. Wartości oczekiwane i odchylenia standardowe a posteriori elastyczności kosztu względem produktu  $Q_2$  dla poszczególnych oddziałów, dla których  $Q_2 < 20$  mln przy ustalonych wartościach pozostałych zmiennych egzogenicznych na poziomie średnich geometrycznych z próby) - VED ( $m=2$ ,  $r^*=0.7$ ,  $Q_1$  i  $Q_2$ ).

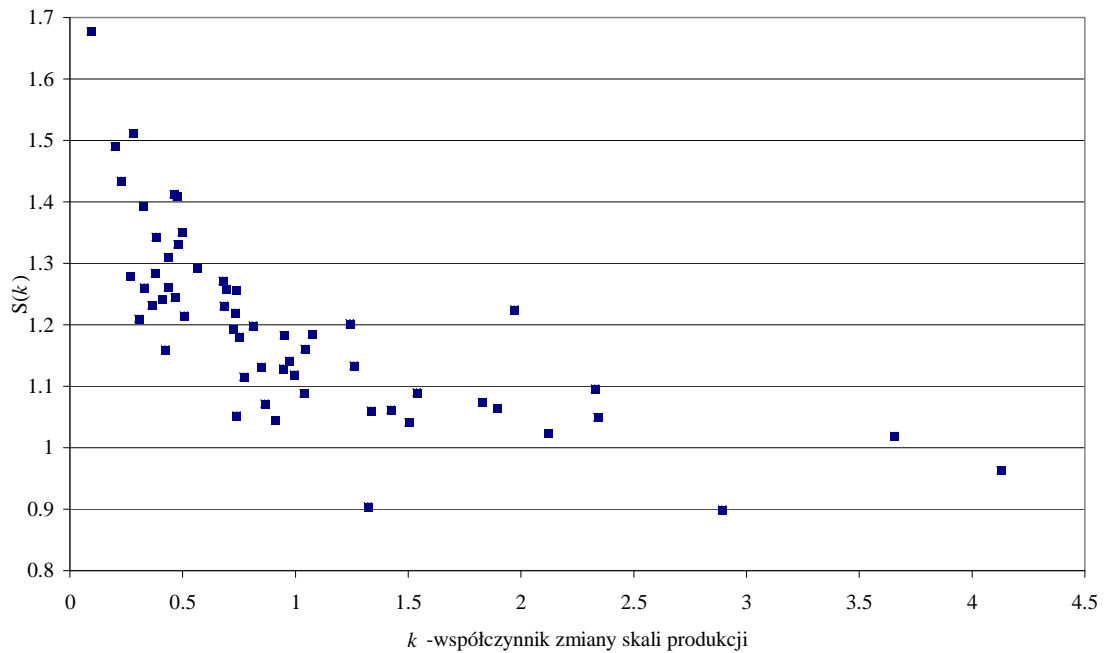


Prezentowane w Tabeli 13 wyniki wskazują, iż prawie wszystkie oddziały (z wyjątkiem dwóch o numerach 3 i 14) wykazują rosnący efekt skali względem czynników zmiennych, mierzony poprzez wartości zagregowanego współczynnika skali produkcji -  $SL(Q_1, Q_2)$ , zdefiniowany w podrozdziale 1.3. Interpretując te wyniki można stwierdzić, iż np. w przypadku drugiego pod względem wielkości kredytów ogółem oddziału zmiana kosztu zmiennego o 1% (około 126 tys. zł) spowodowana proporcjonalnym wzrostem czynników zmiennych powoduje wzrostem wartości udzielonych kredytów (obu typów) o około 1.15% (łącznie o 4 945 tys. zł).

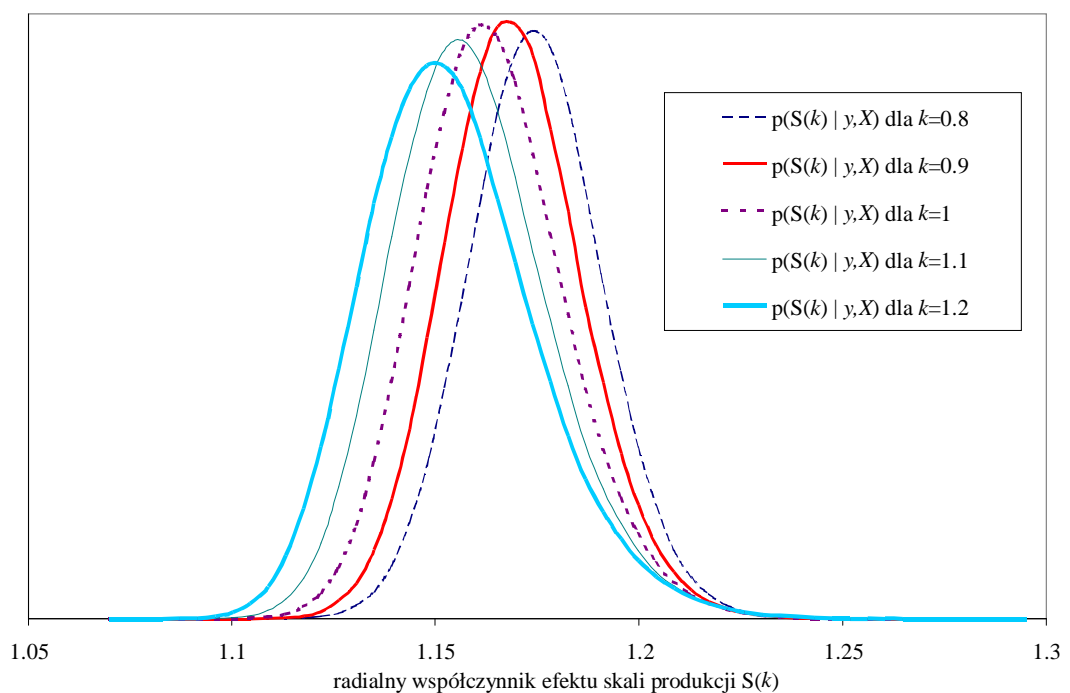
Inną propozycją określenia korzyści skali produkcji jest jego pomiar poprzez wykorzystanie tzw. radialnego współczynnika efektu skali produkcji  $S(k)$ , gdzie  $k$  jest współczynnikiem proporcjonalnej zmiany skali produkcji (omówiony w podrozdziale 1.3, na stronie 26). Rysunek 34 przedstawia wielkości korzyści skali produkcji obliczone dla poszczególnych oddziałów, przy czym zakres wartości parametru  $k$  dla każdego oddziału określono poprzez odniesienie wartości kredytów ogółem poszczególnych oddziałów do średniej z próby. Jedyne trzy oddziały (o numerach 1, 3 i 14) wykazują malejący efekt skali produkcji (względem czynników zmiennych). Pozostałe oddziały wykazują rosnący efekt skali, np. w przypadku wcześniej omawianego oddziału o numerze 2, który posiada około 3.6-krotnie większy wolumen kredytów, wzrost kosztu zmiennego o 1% procent spowoduje

równoczesny wzrost każdego z kredytów o 1.02%. W celu pełnej prezentacji wiedzy o pomiarze radialnego efektu skali produkcji, Rysunek 35 przedstawia rozkłady a posteriori dla różnych wartości współczynnika zmiany skali produkcji -  $k$ . Wysmukłe i symetryczne kształty gęstości świadczą o stosunkowo wysokiej precyzji uzyskanych wyników i przemawiają za hipotezą o rosnących korzyściach skali produkcji.

Rysunek 34. Radialny współczynnik efektu skali produkcji poszczególnych oddziałów jako funkcja  $k$  (oddziały posortowane wg średniej geometrycznej dla  $Q_1$  i  $Q_2$ ) - model VED ( $m=2, r^*=0.7, Q_1$  i  $Q_2$ ).



Rysunek 35. Rozkład a posteriori radialnego współczynnika efektu skali produkcji ( $S(k)$ ) jako funkcji  $k$  przy ustalonych wartościach zmiennych na poziomie średnich geometrycznych z próby - VED ( $m=2, r^*=0.7, Q_1, Q_2$ ).



Uzyskane wyniki empiryczne pomiaru korzyści skali produkcji oddziałów z wykorzystaniem obu mierników  $SL(Q_1, Q_2)$  i  $S(k)$  w przeważającej większości świadczą o rosnącym efekcie skali, choć w przypadku niektórych oddziałów podają one istotnie różne oceny stopnia tych korzyści. Oddziały powinny być zainteresowane zwiększaniem wolumenu kredytów aż do momentu, gdy osiągnęłyby stały, a potem malejący efekt skali produkcji. Pomiar poprzez zagregowany współczynnik skali produkcji wydaje się dawać lepsze wyniki, ponieważ rozpiętość uzyskanych wyników jest dużo mniejsza niż w przypadku radialnego współczynnika skali, dającego trudno interpretowalne wyniki, np. 1.5.

Prezentowane wyniki dotyczące pomiaru efektu skali produkcji potwierdziły rezultaty wcześniejszych badań dla jednoproduktowej funkcji produkcji. Rozróżnienie dwóch produktów pozwala na głębszą analizę kosztów poprzez określenie ewentualnych korzyści zakresu produkcji (pełnego asortymentu - *economies of scope*) bądź specjalizacji. Przyjmuje się, że specjalizacja oddziału ma miejsce, gdy udział jednego z kredytów w ogólnym wolumenie jest znacząco wyższy od przeciętnej obserwowanej w próbie. W tym celu dla przeciętnego oddziału oraz każdego z osobna zbadano warunek wystarczający na istnienie korzyści zakresu produkcji – wzór (17) - i dokonano pomiaru ewentualnych korzyści. W pracach empirycznych prezentowanych w fachowych czasopismach ten warunek bada się zwykle dla hipotetycznego oddziału, dla którego  $w_D = w_L = Q_1 = Q_2 = K = 1$  (por. wzór (18)). Badacze stosujący te uproszczenie starają się tak wyskalować dane, aby średnie wartości tych zmiennych z próby były w przybliżeniu równe jedności. Badania prezentowane dalej pokazują, iż w świetle posiadanych danych oba wzory (17) i (18) - prowadzą do całkiem odmiennych konkluzji. Należy podkreślić w tym momencie, iż podejście bayesowskie, w odróżnieniu od klasycznego, pozwala na wyznaczenie prawdopodobieństwa a posteriori prawdziwości któregoś z dwóch omawianych przypadków – korzyści pełnego zakresu produkcji bądź specjalizacji.

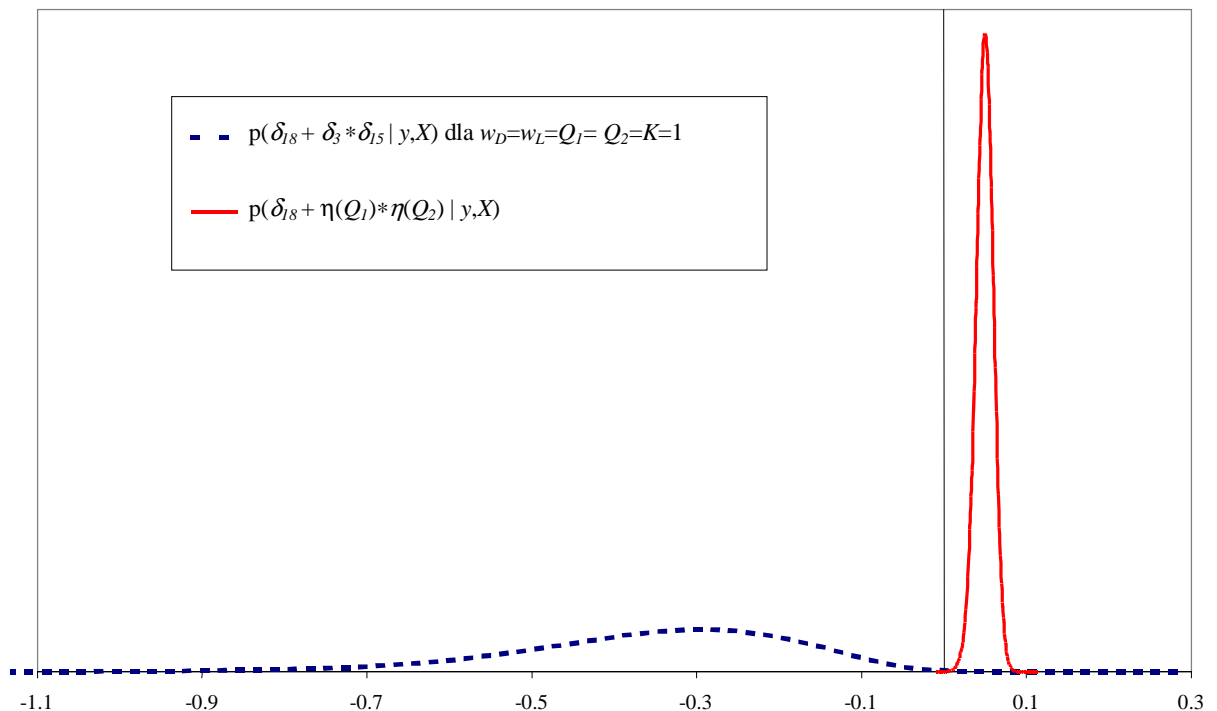
Gęstości a posteriori dla wielkości występującej w warunku wystarczającym na istnienie korzyści zakresu produkcji dla przeciętnego oddziału prezentuje Rysunek 36. Test oparty na przedziale o największej gęstości a posteriori wskazuje, iż prawdopodobieństwo a posteriori prawdziwości warunku wynosi dla przeciętnego oddziału 3/500 000. Prawie cała masa prawdopodobieństwa a posteriori jest skupiona na prawo od zera, na stosunkowo krótkim odcinku, co oznacza dużą precyzję szacunku. Wyniki dla warunku przybliżonego – wzór (18) - są całkowicie odmiennie, ponieważ dotyczy on hipotetycznego oddziału charakteryzującego się jednostkowymi cenami czynników zmiennych, jednostkowym

nakładem czynnika stałego i jednostkowymi wielkościami produktów; w rzeczywistości wartości jednostkowe odbiegają od średnich z próby. W drugim przypadku wykres gęstości a posteriori jest stosunkowo płaski, może trochę asymetryczny, ale przede wszystkim na tyle oddalony na lewo od zera, że prawdopodobieństwo a posteriori prawdziwości tego warunku wynosi aż 0.998. Rezultaty te po raz kolejny świadczą, iż w przypadku funkcji translogarytmicznej niedopuszczalne jest obliczanie wybranych charakterystyk dla punktów spoza zbioru danych (bardzo odległych od przeciętnych wartości z próby) oraz stosowanie grubych uproszczeń, które prowadzą do krańcowo różnych konkluzji. Wyniki uzyskane dla warunku wystarczającego nie rozstrzygają jednak o braku korzyści z pełnego asortymentu produkcji zarówno w przypadku przeciętnego oddziału jak i pozostałych, zwłaszcza, że gęstość a posteriori dla miernika ewentualnych korzyści ( $SCP$  – wzór (20); Rysunek 37) wskazuje zdecydowanie na istnienie tych korzyści. Prawdopodobieństwo a posteriori, że  $SCP < 0$  wynosi tylko 67/500 000. Dodatkowo zbadano prawdziwość tego warunku także dla poszczególnych oddziałów, co przedstawia Tabela 14, i dokonano pomiaru ewentualnych korzyści – możliwość redukcji kosztu z tytułu produkcji obu produktów jednocześnie – poprzez miernik  $SCP$ . Wyniki dla oddziałów z końca tabeli (zwłaszcza o numerach 7, 41, i 43, które charakteryzują się ujemnymi elastycznościami kosztu względem kredytów detalicznych, z całą pewnością wskazują na występowanie korzyści pełnego asortymentu. W badanym okresie oddziały te wykazują specjalizację, ponieważ w ich portfelu kredytowym zdecydowanie przeważają kredyty udzielone klientom komercyjnym, które stanowią od 93% do 99.5% wolumenu wszystkich kredytów. Dokonano także pomiaru korzyści zakresu produkcji ( $SCP_i$ ), np. w przypadku oddziału o numerze 5, przy obserwowanych cenach czynników i nakładzie czynnika stałego, koszt równoczesnego zwiększenia wolumenu obu kredytów o różnice między wielkościami przeciętnymi a minimalnymi obserwowanymi w próbie, byłby niższy o 17.7% ( $\pm 0.032\%$ ) od kosztu jakie poniosłyby dwa identyczne oddziały, z których każdy zamierzałby specjalizować się, poprzez zwiększenie produkcji tylko jednego z produktów o wspomnianą wielkość. Wyniki te wskazują, że oddziały najsilniej specjalizujące się w udzielaniu kredytów komercyjnych mogłyby obniżyć koszty produkcji dodatkowego wolumenu kredytów ogółem, jeżeli odstąpiłyby od dalszego zwiększania kredytów komercyjnych na rzecz detalicznych. Natomiast z drugiej strony obserwuje się, iż dla oddziałów o numerach 14, 30 i 26, które posiadają największy udział kredytów detalicznych w kredytach ogółem, warunek wystarczający na istnienie korzyści pełnego zakresu produkcji jest silnie niespełniony. Ujemne wartości miernika  $SCP_i$  świadczą o ewentualnych korzyściach jakie wynikałyby ze specjalizacji, jednakże trudno powiedzieć czy

w udzielaniu kredytów detalicznych czy komercyjnych. W przypadku oddziałów znajdujących się w środku omawianej tabeli trudno jednoznacznie rozstrzygać na podstawie miernika  $SCP_i$  o istnieniu korzyści zakresu produkcji czy specjalizacji, jeżeli sam warunek wystarczający nie jest spełniony.

Przedstawione wyniki, zwłaszcza te dotyczące pomiaru ewentualnych korzyści z pełnego asortymentu, należy traktować jako wstępne. Zastosowana bayesowska technika estymacji granicznych, stochastycznych modeli kosztu zaowocowała wynikami empirycznymi, których uzyskanie na gruncie klasycznym jest trudne albo wręcz niemożliwe. Istnieje potrzeba zdefiniowania dla translogarytmicznej funkcji kosztu lepszej niż  $SCP$  i  $SC$  - por. wzory (19) i (20) - miary korzyści pełnego zakresu zarówno dla dwóch produktów jak i większej ich liczby. Wyróżnienie większej liczby produktów umożliwiłoby przeprowadzenie pogłębionych i specjalistycznych badań, które na gruncie klasycznym są podejmowane już od ponad 15 lat, lecz nie wydają się przynosić zadawalających rezultatów.

Rysunek 36. Rozkład a posteriori warunku wystarczającego na istnienie korzyści zakresu produkcji (*economies of scope*) (przy ustalonych wartościach zmiennych na poziomie średnich arytmetycznych z próby) - model VED ( $m=2, r^*=0.7, Q_1$  i  $Q_2$ ).



Rysunek 37. Rozkład a posteriori miary korzyści zakresu produkcji ( $SCP > 0$ ) (przy wartościach zmiennych objaśniających ustalonych na poziomie średnich arytmetycznych z próby) - model VED ( $m=2, r^*=0.7, Q_1$  i  $Q_2$ ).

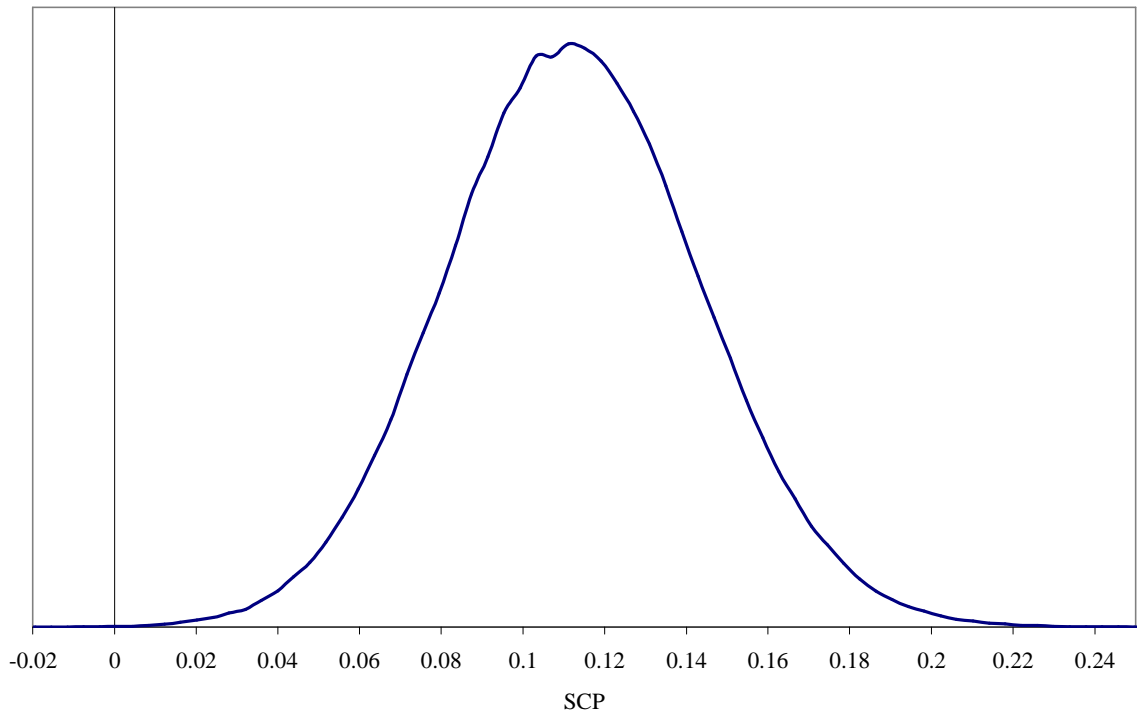


Tabela 14. Wartości oczekiwane i odchylenia standardowe a posteriori dla wielkości  $\delta_{18} + \eta(Q_1)_i \cdot \eta(Q_2)_i$  i dla miary korzyści zakresu produkcji SCP – model VED ( $m=2, r^*=0.7, Q_1$  i  $Q_2$ ).

$i$	$E(\delta_{18} + \eta(Q_1)_i \cdot \eta(Q_2)_i   \cdot)$	$D(\delta_{18} + \eta(Q_1)_i \cdot \eta(Q_2)_i   \cdot)$	$E(SCP_i   \cdot)$	$D(SCP_i   \cdot)$	$Q_2 / (Q_1 + Q_2)$
14	0.242	0.031	-0.367	0.042	62.60%
30	0.168	0.018	-0.115	0.024	42.70%
26	0.155	0.016	-0.074	0.032	52.20%
3	0.136	0.029	-0.118	0.059	6.60%
46	0.131	0.014	0.073	0.029	29.70%
23	0.121	0.02	-0.048	0.033	11.50%
53	0.12	0.015	0.188	0.033	23.70%
13	0.099	0.015	0.041	0.037	15.60%
50	0.099	0.011	0.19	0.03	21.10%
55	0.089	0.012	0.31	0.029	16.60%
47	0.088	0.008	0.194	0.026	20.80%
20	0.081	0.014	0.101	0.043	15.10%
12	0.079	0.02	0.09	0.068	11.60%
11	0.078	0.016	0.048	0.029	7.90%
24	0.078	0.013	0.117	0.043	14.50%
19	0.077	0.015	0.171	0.054	8.70%
32	0.077	0.012	0.059	0.03	21.40%
58	0.076	0.007	5.376	1.321	47.40%
28	0.075	0.018	0.066	0.031	6.90%
51	0.074	0.012	0.244	0.027	13.10%
6	0.072	0.017	0.089	0.045	8.00%
56	0.071	0.007	0.604	0.071	32.20%
25	0.068	0.013	0.086	0.023	8.10%
44	0.067	0.009	0.226	0.033	16.40%
38	0.059	0.013	0.172	0.026	9.70%
18	0.054	0.014	0.195	0.062	13.30%

$i$	$E(\delta_{18} + \eta(Q_1)_i * \eta(Q_2)_i   \cdot)$	$D(\delta_{18} + \eta(Q_1)_i * \eta(Q_2)_i   \cdot)$	$E(SCP_i   \cdot)$	$D(SCP_i   \cdot)$	$Q_2 / (Q_1 + Q_2)$
10	0.053	0.015	0.158	0.052	9.10%
21	0.046	0.013	0.129	0.034	8.10%
9	0.041	0.017	0.178	0.058	7.10%
27	0.037	0.01	0.182	0.033	8.60%
42	0.035	0.015	0.209	0.019	5.70%
8	0.034	0.015	0.141	0.036	4.90%
48	0.034	0.008	0.326	0.036	12.90%
29	0.031	0.013	0.139	0.02	4.70%
1	0.029	0.023	0.175	0.065	3.00%
49	0.029	0.015	0.269	0.018	5.80%
31	0.028	0.01	0.198	0.031	7.90%
15	0.026	0.012	0.167	0.034	5.90%
57	0.024	0.011	0.622	0.031	10.50%
40	0.022	0.01	0.308	0.045	11.50%
4	0.02	0.016	0.167	0.041	3.90%
45	0.017	0.011	0.259	0.019	5.80%
37	0.012	0.009	0.247	0.028	7.20%
22	0.01	0.012	0.171	0.025	4.80%
35	0.009	0.012	0.184	0.019	4.60%
52	0.004	0.011	0.362	0.021	6.70%
17	-0.001	0.015	0.187	0.034	4.40%
34	-0.016	0.013	0.189	0.015	3.10%
54	-0.016	0.01	0.463	0.03	7.10%
39	-0.018	0.011	0.276	0.023	5.10%
33	-0.03	0.014	0.177	0.014	2.40%
36	-0.031	0.015	0.192	0.013	2.80%
16	-0.033	0.016	0.21	0.033	3.10%
2	-0.036	0.02	0.173	0.033	1.30%
5	-0.042	0.019	0.177	0.032	1.70%
43	-0.044	0.013	0.283	0.017	3.60%
41	-0.053	0.02	0.269	0.015	3.20%
7	-0.149	0.021	0.064	0.005	0.50%



## Podsumowanie.

Głównym celem pracy była bayesowska analiza mikroekonomicznej efektywności kosztowej oddziałów wybranego polskiego banku komercyjnego za pomocą stochastycznych modeli granicznych. Rozdziały 1-5 przedstawiają różne etapy realizacji tego celu oraz celów częściowych: od sformułowania modelu do prezentacji wyników empirycznych.

W rozdziale pierwszym przedstawiono mikroekonomiczne podstawy analizy efektywności kosztowej firm. Omówiono funkcję produkcji i kosztu, kładąc duży nacisk na definicję krótkookresowej funkcji kosztu zmiennego. W kontekście problemu minimalizacji kosztu zdefiniowano podstawowe kategorie efektywności: techniczną, alokacyjną i kosztową, prezentując dekompozycję efektywności kosztowej na dwie pierwsze składowe. Wprowadzono najprostszy jednorównaniowy stochastyczny model granicznej krótkookresowej funkcji kosztu zmiennego, zwracając uwagę na rolę złożonego, niesymetrycznego składnika losowego definiującego tę klasę modeli i determinującego problemy ich estymacji. Zaproponowano i omówiono własności funkcji translogarytmicznej jako podstawowej formy funkcyjnej dla krótkookresowej funkcji kosztu zmiennego. Następnie przedstawiono podstawowe charakterystyki wieloproduktowego procesu produkcji, tj. elastyczności oraz różne definicje miar efektu skali produkcji. Dodatkowo przedstawiono definicję korzyści pełnego zakresu produkcji, formalny warunek wystarczający na jego istnienie oraz propozycje pomiaru ewentualnych korzyści.

W rozdziale drugim przedstawiono mikroekonomiczny model instytucji finansowych zaproponowany przez Sealeya i Lindleya, zwracając szczególną uwagę na definicję czynników produkcji i produktów w kontekście problemu maksymalizacji zysku. W dalszej części tego rozdziału dokonano przeglądu literatury światowej, podejmującej analizę efektywności kosztowej banków komercyjnych; prezentowane są dwa różne sposoby określenia produktów i czynników produkcji, tzn. podejście produkcyjne i intermediacyjne. W oparciu o to drugie podejście określono produkty i czynniki produkcji dla oddziałów badanego banku. Za czynniki produkcji przyjęto:

- nieruchomości (a także inwestycje w nieruchomości obce) - czynnik stały,
- inne środki trwałe (m.in. sprzęt komputerowy) oraz wartości niematerialne i prawne (m.in. oprogramowanie),

- pracowników oddziału,
- depozyty (złotowe i walutowe) i środki pieniężne pozyskane przez oddział z centrali i innych oddziałów.

Produkcję oddziałów bankowych wyrażono przez łączną wartość różnego rodzaju kredytów złotych i walutowych (komercyjnych i detalicznych) oraz - w przypadku oddziałów depozytowych - nadwyżki między wartością przyjętych lokat a wartością udzielonych kredytów. Natomiast objaśniane przez model całkowite koszty zmienne stanowią sumę kosztów związanych z zaangażowaniem każdego z czynników zmiennych, czyli wynagrodzeń wraz z narzutami, kosztów rzeczowych (m.in. zużycie tonera i papieru do drukarek, dyskietek, materiałów biurowych) oraz kosztów kapitału finansowego (odsetek od depozytów, a także prowizji płaconych od zaciągniętych przez dany oddział kredytów i pożyczek). Dodatkowo, w celu dokonania analizy potencjalnych korzyści zakresu produkcji, rozróżniono dwa rodzaje kredytów, tj. kredyty udzielone klientom detalicznym i kredyty udzielone klientom komercyjnym. Natomiast z uwagi na charakter jednego z czynników produkcji, innych środków trwałych i wartości niematerialnych i prawnych, przyjęto iż trudno mierzalna i nieobserwowana cena tego czynnika zmiennego jest stała zarówno po czasie, jak i po oddziałach, co w konsekwencji zmniejszyło liczbę parametrów rozważanej funkcji kosztu.

W rozdziale trzecim przedstawiono i poddano krytyce podstawowe, klasyczne techniki estymacji modeli granicznych i pomiaru efektywności, przede wszystkim skorygowaną metodę najmniejszych kwadratów (MKN) i metodę największej wiarygodności (MNV) w przypadku danych przekrojowych oraz model efektów stałych i estymator wewnątrzgrupowy w przypadku danych przekrojowo-czasowych. Zwrócono uwagę, iż metody klasyczne charakteryzującą się prostotą obliczeń, ale często nie wykorzystują założenia o postaci rozkładu dla zmiennej reprezentującej nieefektywność, wymagając natomiast innych, dość silnych założeń. Dodatkowo ich własności stochastyczne w małej próbie nie są znane, a ich stosowanie może prowadzić do wyników trudno interpretowalnych z ekonomicznego punktu widzenia.

Część pierwsza rozdziału czwartego przedstawia zarys bayesowskich reguł wnioskowania statystycznego. Zasadnicza część tego rozdziału jest poświęcona konstrukcji bayesowskiego stochastycznego modelu granicznego o wspólnym rozkładzie efektywności dla losowych efektów indywidualnych (CED) oraz jego uogólnienia, tj. modelu o zmiennym rozkładzie efektywności (VED), Modele te wprowadzili Koop, Osiewalski i Steel [1997]. Ze

względu na posiadanie danych kwartalnych z jednego tylko roku przyjęto w obu modelach, że efektywność jest czystym efektem indywidualnym, czyli że  $z_{it}=z_i$  ( $t=1,\dots,T$ ), gdzie zmienna  $z_i$  reprezentuje potencjalną nieefektywność  $i$ -tego oddziału. Założenie stałości nieefektywności po czasie umożliwiło bardziej precyzyjny szacunek przeciętnej efektywności każdej firmy. Miarą efektywności kosztowej jest  $r_i=\exp(-z_i)$ .

Z uwagi na fakt, że w każdym z modeli łączna funkcja gęstości określona na przestrzeni danych i wielkości nieobserwowalnych ma skomplikowaną postać, a tradycyjne kwadratury i metody Monte Carlo z niezależnymi losowaniami nie stanowią efektywnych narzędzi przybliżania odpowiednich całek, więc wykorzystano próbkowanie Gibbsa jako technikę wyznaczania rozkładów a posteriori i ich charakterystyk, zarówno dla nieznanych parametrów granicznej funkcji kosztów, jak i dla wskaźników efektywności. Dostosowano do potrzeb tej pracy schemat algorytmu Gibbsa dla modeli CED i VED, który opracowali Koop, Osiewalski i Steel [1997].

W rozdziale piątym, w pierwszej części dokonano charakterystyki danych źródłowych wykorzystywanych w empirycznej analizie efektywności kosztowej oddziałów. Dodatkowo określono trzy zmienne dychotomiczne służące wyjaśnieniu zróżnicowania efektywności, a odzwierciedlające 1<sup>o</sup> specjalizację oddziałów, 2<sup>o</sup> wielkość oddziału mierzoną wartością produkcji i 3<sup>o</sup> złożoną strukturę organizacyjną, czyli posiadanie przez oddziały filii mających szeroki zakres uprawnień.

Następnie dla danych przekrojowych dokonano przy pomocy skorygowanej metody najmniejszych kwadratów estymacji krótkookresowej granicznej funkcji kosztu zmiennego dla zmodyfikowanej funkcji Cobba-Douglasa dopuszczającej zmienny efekt skali, jako szczególnego przypadku funkcji translogarytmicznej. Uzyskane wyniki bez narzucania jakichkolwiek restrykcji na parametry wskazują, że cena depozytów wydaje się w sposób znaczący wpływać na kształtowanie się kosztu zmiennego, natomiast trudno określić kierunek i siłę wpływu ceny pracy, ponieważ za sprawą silnej współliniowości otrzymane błędy średnie szacunku są znaczne. W prawie wszystkich oddziałach banku zaobserwowano rosnący efekt skali produkcji względem czynników zmiennych. Efektywność oddziałów w badanych czterech kwartałach nie zmienia się w sposób systematyczny, średnia efektywność jest rzędu 0.87 – 0.89, czyli przeciętnie około 11 - 13% kosztu zmiennego poniesionego przez wszystkie oddziały banku było kosztem nadwyżkowym, nie uzasadnionym ani wielkością produkcji ani cenami czynników zmiennych czy wielkością zaangażowania czynnika stałego. Natomiast obserwuje się prawdopodobnie nadmierne zróżnicowanie efektywności wśród

oddziałów, co wynika z uproszczonej metody estymacji funkcji kosztu i pomiaru samej efektywności.

Otrzymane oceny parametrów dla danych przekrojowych zmodyfikowanej funkcji Cobba-Douglasa były podstawą do sformułowania tezy, że w badanym okresie charakter procesu produkcji oddziałów banku nie zmienia się w czasie i wskazują, że do jego opisu należy używać form funkcyjnych dopuszczających przynajmniej zmienność efektu skali. Ewentualne różnice ocen parametrów technologii w badanych kwartałach były spowodowane dużymi błędami średnimi szacunku, które wynikają z faktu estymacji modelu tylko dla danych przekrojowych, a także ze współliniowości. Uzyskane wyniki dały podstawę do budowy modelu dla danych przekrojowo-czasowych, w którym można by przyjąć założenie o stałości technologii i stałości efektywności w omawianym okresie czasu, co pozwoli na precyzyjny szacunek parametrów technologii i przeciętnej efektywności każdego z oddziałów.

W części trzeciej rozdziału piątego przedstawiono estymację stochastycznego modelu granicznego i pomiar efektywności w przypadku danych przekrojowo-czasowych, w oparciu o bayesowski model o zmiennym rozkładzie efektywności (VED). Jako formę funkcyjną kosztu zmiennego przyjęto funkcję translogarytmiczną z jednym zagregowanym produktem, którego wielkość mierzona jest poprzez wartość kredytów i innych środków pożyczonych przez oddziały. Wykorzystując podejście bayesowskie można było uwzględnić poprzez rozkład a priori podstawowe nieliniowe i nierównościowe restrykcje na parametry technologii zgodnie z mikroekonomiczną teorią produkcji i kosztu. W przypadku podejścia klasycznego jest to trudne i nie zawsze możliwe.

Zastosowany algorytm Gibbsa okazał się efektywnym narzędziem uzyskiwania wyników a posteriori; zbieżność osiągnął już po około  $10^5$  cykli wstępnych, a dodatkowe  $5 \cdot 10^5$  cykli wykonano w celu otrzymania zadawalającej aproksymacji obliczanych charakterystyk a posteriori.

Uzyskane wyniki empiryczne dostarczyły wielu cennych i statystycznie istotnych informacji zarówno o charakterze technologii, jak i poziomie efektywności kosztowej oddziałów badanego banku. Charakteryzują się one przeciętnie dość wysoką efektywnością, tj. nadwyżkowy nieuzasadniony koszt zmienny stanowi średnio 8.2% obserwowanego kosztu zmiennego. Dane silnie zmodyfikowały wstępne założenia o efektywności powodując, że mediana rozkładu a posteriori dla miary efektywności wyniosła około 0.93, gdy a priori przyjęto ją na poziomie 0.528 (model VED dla  $r^*=0.7$ ). Zróznicowanie indywidualnych

wskaźników efektywności dla oddziałów jest o wiele mniejsze niż wskazywały wyniki wstępne uzyskane klasycznymi, uproszczonymi metodami dla zmodyfikowanej funkcji Cobba i Douglasa. Podejście bayesowskie pozwoliło wnioskować, iż zarówno struktura organizacyjna oddziału, jak i skala prowadzonej działalności nie mają istotnego wpływu na zróżnicowanie efektywności w poszczególnych grupach oddziałów. Nieznaczne różnice w poziomie efektywności wydają się być spowodowane specjalizacją oddziałów, tj. większą efektywnością charakteryzują się oddziały kredytowe (posiadające nadwyżkę wartości kredytów nad depozytami) niż oddziały depozytowe.

Największy wpływ na kształtowanie się kosztu zmiennego mają wielkość produkcji i cena czynnika finansowego. Wzrost łącznej wartości udzielonych kredytów o 1% powoduje wzrost kosztu zmiennego o około 0.86% ( $\pm 0.025\%$ ) *ceteris paribus*. Natomiast wzrost ceny depozytów i innych pozyskanych przez oddział środków powoduje wzrost kosztu zmiennego o około 0.8% ( $\pm 0.037\%$ ) *ceteris paribus*. Cena pracy ma mniejszy wpływ na poziom kosztu, aczkolwiek też odgrywa istotną statystycznie rolę, tzn. wzrost płacy o 1% powoduje wzrost kosztów zmiennych o 0.18% ( $\pm 0.026\%$ ). Rola nieobserwowanej, stałej dla oddziałów ceny czynnika materiałowego jest znikoma lub żadna. Wśród prawie wszystkich oddziałów obserwuje się rosnący efekt skali produkcji względem czynników zmiennych, zatem w praktyce oddziały mogą zwiększać zysk przede wszystkim zwiększając skalę produkcji. Wyniki a posteriori wskazują, że przyjęcie translogarytmicznej formy funkcyjnej było słuszne, ponieważ zarówno elastyczności względem ceny pracy, nakładów czynnika stałego i wielkości produkcji (także współczynnik efektu skali produkcji) są zmienne po oddziałach, tzn. istotnie zależą od zmiennych specyfikujących krótkookresową funkcję kosztu. Jedynie elastyczności kosztu względem ceny depozytów charakteryzują się na tyle znacznymi odchyleniami standardowymi a posteriori, że trudno byłoby odrzucić hipotezę o jednakowej elastyczności, przyjmowanej a priori w modelu Cobba i Douglasa. Dokonano także analizy procesu produkcji oddziałów banku rozważając dwa produkty: kredyty detaliczne i komercyjne, uzyskując wstępne wyniki dotyczące korzyści pełnego zakresu bądź specjalizacji.

Porównując rezultaty uzyskane najprostszą metodą dla zmodyfikowanej funkcji Cobba i Douglasa i te uzyskane za pomocą bayesowskiego modelu VED można dostrzec, że najważniejsze wyniki empiryczne obu modeli prowadzą do tych samych ogólnych wniosków. Można sobie zatem postawić pytanie: czy warto było zastosować dość skomplikowany bayesowski model efektów losowych, gdy porównywalne informacje o technologii

i efektywności uzyskano stosując uproszczone modele i techniki estymacji? Odpowiedź wydaje się prosta, mianowicie tylko zastosowanie dość wyrafinowanego i ogólnego modelu statystycznego pozwoliło stwierdzić *ex post*, że w przypadku badanych oddziałów uzasadnione jest wykorzystanie zmodyfikowanej funkcji Cobba i Douglasa dopuszczającej jedynie zmienny efekt skali, a szacowanej skorygowaną metodą najmniejszych kwadratów.

Doświadczenie zdobyte w trakcie badań pozwalają na sformułowanie kierunków dalszych pogłębionych analiz zarówno w sferze aplikacyjnej jak i metodologicznej. Analiza empiryczna w oparciu o funkcję kosztu byłaby pełniejsza, gdyby – zamiast agregatu - można byłoby rozważać poszczególne czynniki finansowe, np. depozyty a vista i terminowe. Wielu ekonomistów uzna, iż w horyzoncie czterech kwartałów prace ludzką można uznać za czynnik stały, co spowodowałoby nową specyfikację krótkookresowej funkcji kosztu - i podobnie jak w przypadku kilku czynników finansowych - potrzebę powtórnej estymacji modelu. Dalsze badania mogłyby pójść równocześnie w kierunku poszukiwania nowych mierników korzyści pełnego zakresu bądź specjalizacji.

W pracy dokonując specyfikacji rozkładu *a priori* dla parametrów technologii narzucono jedynie podstawowe warunki regularności ekonomicznej (konieczne dla minimalizacji kosztu). Natomiast rzadko w pracach empirycznych uwzględnia się bądź sprawdza warunek konieczny i wystarczający; Varian [1992] podaje jako konsekwencję tego warunku fakt, że elastyczność względem ceny danego czynnika powinna być niemalejącą funkcją tej ceny. Uzyskane wyniki *a posteriori* wydają się nie potwierdzać tego faktu w przypadku elastyczność względem ceny depozytów. Zatem tematem dalszych pogłębionych analiz może być sprawdzenie warunku koniecznego i wystarczającego na istnienie dodatniego rozwiązania problemu minimalizacji koszt.

Kierunkiem dalszych zainteresowań autora jest zbadanie w jakim stopniu oddziały maksymalizują zysk, czyli analiza efektywności wykorzystująca funkcję zysku. Podjęcie takich badań jest możliwe, gdy oddziały charakteryzują się malejącymi korzyściami skali produkcji (warunek konieczny na istnienie jednoznacznego, dodatniego rozwiązania problemu maksymalizacji zysku).

Obszarem nowych badań mającym aspekt metodologiczny byłaby implementacja bayesowskiego testu wykorzystującego iloraz szans *a posteriori* (*posterior odds ratio*) do obliczania prawdopodobieństw *a posteriori* konkurencyjnych modeli (pełnego i po redukcji; np. VED i CED). Z uwagi na przyjętą metodę numerycznej analizy rozkładów *a posteriori*, tj. algorytm Gibbsa, wykorzystanie powyższego testu wiąże się w ogromnymi trudnościami; w

światowej literaturze przedmiotu poszukuje się uniwersalnego i efektywnego sposobu jego zastosowania.

Prowadzone badania i uzyskane wyniki utwierdzają w przekonaniu, że podjęta tematyka badawcza jest bardzo interesująca z punktu widzenia ekonomii, bankowości i ekonometrii. Może być użyteczna w praktycznej analizie ekonomicznej i zarządzaniu bankiem.

## Wykaz pozycji literaturowych.

- [1] Aigner D., C.A.K. Lovell, P. Schmidt, 1977, „Formulation and Estimation of Stochastic Frontier Production Function Models”, *Journal of Econometrics*, 6.
- [2] Akhaiven J., P.A.V.B. Swamy, S.B. Taubman, R.N. Singamsetti, 1997, „A general method of deriving the inefficiencies of bank from a profit function”, *The Journal of Productivity Analysis*, 8.
- [3] Atkinson S.E., C. Cornwell, 1994a, „Parametric estimation of technical and allocative inefficiency with panel data”, *International Economic Review*, 1.
- [4] Atkinson S.E., C. Cornwell, 1994b, „Estimation of output and input technical efficiency using a flexible functional form and panel data”, *International Economic Review*, 1.
- [5] Bauer P.W., D. Hancock, 1993, „The efficiency of the Federal Reserve in providing check processing services”, *Journal of Banking and Finance*, 17.
- [6] Baumol W., J.C. Panzar, R.D. Willig, 1982, *Contestable Markets and the Theory of Industry Structure*, Harcourt Brace Javanovich, New York.
- [7] Beckers D.E., C.J. Hammond, 1987, „A tractable likelihood function for the normal-gamma stochastic frontier model”, *Economics Letters*, 24.
- [8] Belsley D.A., E. Kuh, R.E. Welsh, 1980, *Regression Diagnostics*, Wiley, New York.
- [9] Berg S.A, F.R. Forsund, L. Hjalmarsson, M. Souminen, 1993, „Banking efficiency in the Nordic countries”, *Journal of Banking and Finance*, 17.
- [10] Berger A. N., 1993, „‘Distribution-free’ estimates of efficiency in the U.S. banking industry and tests of the standard distributional assumptions”, *The Journal of Productivity Analysis*, 4.
- [11] Berger A. N., D. B. Humphrey, 1982, „Measurement and efficiency issues in commercial banking”, w: *Measurement issue in the service sectors* (red, Z.Griliches, University of Chicago Press.
- [12] Berger A. N., D. Hancock, D.B. Humphrey, 1993, „Bank efficiency derived from the profit function”, *Journal of Banking and Finance*, 17.
- [13] Berger A. N., D.B. Humphrey, 1991, „The dominance of inefficiencies over scale and product mix economies in banking”, *Journal of Monetary Economics*, 28.
- [14] Berger A. N., G. Hanweck, D. Humphrey, 1987, „Competitive viability in banking. Scale, scope, and mix economies”, *Journal of Monetary Economics*, 20.
- [15] Berger A. N., L. Mester, 1993, „Inside the black box: What explains differences in the efficiencies of financial institutions?”, *Journal of Banking and Finance*, 21.



- [16] Berger A. N., R. DeYoung, 1997, „Problem loans and cost efficiency in commercial Banks”, *Journal of Banking and Finance*, 21.
- [17] Berger A. N., W.C. Hunter, S.G. Timme, 1993, „The efficiency of financial institutions: A review and preview of research past, present, and future”, *Journal of Banking and Finance*, 17.
- [18] Box G., G. Tiao, 1973, *Bayesian Inference in Statistical Analysis*, Addison-Wesley Publishing Company, Reading.
- [19] van den Broeck J., G. Koop, J. Osiewalski, M.F.J. Steel, 1994, „Stochastic frontier models: A Bayesian perspective”, *Journal of Econometrics*, 61.
- [20] Capiga M., R. Kurzak, 1998, „Miary syntetyczne w ocenie banków”, *Bank*, 10.
- [21] Casella G., E. George, 1992, „Explaining the Gibbs Sampler”, *The American Statistician*, 46.
- [22] Cebenoyan A.S., E.S. Cooperman, C.A. Register, S.C. Hudgins, 1993, „The relative efficiency of stock versus Mutual S&Ls: A stochastic cost frontier approach”, *Journal of Financial Services Research*.
- [23] Christensen L.R., D. Jorgenson, L. Lau, 1973, „Transcendental Logarithmic Production Frontiers”, *The Review of Economics and Statistics*, 55.
- [24] Christensen L.R., W.H. Greene, 1976, „Economics of scale in U.S. electric power generation”, *Journal of Political Economy*, 84.
- [25] Cornwell C., P. Schmidt, R. Sickles, 1990, „Production frontiers with-cross-sectional and time-series variation in efficiency levels”, *Journal of Econometrics*, 46.
- [26] Dietsch M., 1993, „Economies of scale and scope in French commercial banking industry”, *Journal of Productivity Analysis*, 4.
- [27] English M., S. Grosskopf, K. Hayes., S. Yaiswarng, 1993, „Output allocative and technical efficiency of banks”, *Journal of Banking and Finance*, 17.
- [28] Färe R., C.A.K. Lovell, 1978, „Measuring the technical efficiency of production”, *Journal of Economic Theory*, 19.
- [29] Färe R., D. Promont, 1993, „Measuring the efficiency of multiunit banking: An activity analysis approach”, *Journal of Banking and Finance*, 17.
- [30] Farrell M., 1957, „The measurement of productive efficiency”, *Journal of the Royal Statistical Society, Series A*, 120.
- [31] Fernández C., G. Koop, M.F.J. Steel, 1999, A Bayesian analysis of multiple-output production frontiers”, *Journal of Econometrics*, w druku.
- [32] Fernández C., J. Osiewalski, M.F.J. Steel, 1997, „On the use of panel data in stochastic frontier models with improper priors”, *Journal of Econometrics*, 79.
- [33] Ferrier G.D., C.A.K. Lovell, 1990, „Measuring cost efficiency in banking: econometric and linear programming evidence”, *Journal of Econometrics*, 46.
- [34] Frisch R., 1965, *Theory of Production*, Chicago.

- [35] Gabrielsen A., 1975, „On estimating efficient production functions”, *Working Paper No. A-35*, Chr. Michelsen Institute, Department of Humanities and social Sciences, Bergen.
- [36] Gilligan T., M. Smirlock, W. Marshall, 1984, „Scale and scope economies in the multi-product banking firm”, *Journal of Monetary Economics*, 13.
- [37] Grabczan W., 1995, „Analiza i ocena wyników działalności banku”, *Bank i Kredyt*, 11.
- [38] Grabowski R., N. Ragan, R. Rezvanian, 1993, „Organizational forms in banking: An empirical investigation of cost efficiency”, *Journal of Banking and Finance*, 17.
- [39] Greene W.H., 1980, „Maximum likelihood estimation of econometric frontier functions”, *Journal of Econometrics*, 13.
- [40] Greene W.H., 1990, „A gamma-distributed stochastic frontier model”, *Journal of Econometrics*, 46.
- [41] Greene W.H., 1993, „The econometric approach to efficiency analysis” w: Fried H.O., Lovell K.C.A., Schmidt S.S. (red.), *The Measurement Of Productive Efficiency - Techniques And Applications*, Oxford University Press, New York, chap. 2.
- [42] Hassan Y.A., Grabowski R., C. Pasurka, N. Ragan, 1990, „Technical, scale, and allocative efficiencies in U.S. banking: An empirical investigation”, *Review of Economics and Statistics*.
- [43] Hughes J., L.J. Mester, 1993, „A quality and risk-adjusted cost function for banks: Evidence on the „too-big-to-fail” doctrine”, *Journal of Productivity Analysis*, 4.
- [44] Humphrey D., 1985, „Costs and scale economies in bank intermediation”, w: *Handbook of Banking Strategy* (red.: R.C. Aspinwall, R. Eisenbeis), J. Wiley, New York.
- [45] Humphrey D., 1993, „Cost and technical change: Effects from bank deregulation”, *Journal of Productivity Analysis*, 4.
- [46] Jakubczyc J., 1987, *Współliniowość statystyczna*, PWE, Warszawa.
- [47] Jaworski W.L., 1996, *Bankowość*, Poltex, Warszawa.
- [48] Jondrow J., C.A.K. Lovell, I. Materov, P. Schmidt, 1982, „On the estimation of technical inefficiency in the stochastic frontier production function model”, *Journal of Econometrics*, 19.
- [49] Judge G., W. Griffiths, R. Hill, H. Lütkepohl, T. Lee, 1985, *The Theory and Practice of Econometrics (2nd ed.)*, J. Wiley, New York.
- [50] Kaparakis E., S. M. Miller, A.G. Noulas, 1994, „Short-run cost inefficiency of commercial banks: A flexible stochastic frontier approach”, *Journal of Money, Credit, and Banking*, 26.
- [51] Kim M., 1985, „Scale economies in banking: A methodological note”, *Journal of Money, Credit, and Banking*, 1.
- [52] Kim M., 1986, „Banking technology and the existence of a consistent output aggregate”, *Journal of Monetary Economics*, 18.
- [53] Kolari J., Zardokoohi A., 1987, „Bank, cost, structure, and performance”, *Heath & Co.*, Lexington.

- [54] Koop G., J. Osiewalski, M.F.J Steel, 1994, „Bayesian efficiency analysis with a flexible form: The AIM cost function”, *Journal of Business and Economic Statistics*, 12.
- [55] Koop G., J. Osiewalski, M.F.J Steel, 1997, „Hospital efficiency analysis with individual effects: A Bayesian approach”, *Journal of Econometrics*, 76, 77-105.
- [56] Koop G., M.F.J. Steel, J. Osiewalski, 1995, „Posterior analysis of stochastic frontier models using Gibbs sampling”, *Computational Statistics*, 10.
- [57] Koop G, J. Osiewalski, M.F.J. Steel, 1999, „The components of output growth: A stochastic frontier analysis”, *Oxford Bulletin of Economics and Statistics*, 61.
- [58] Koop G, J. Osiewalski, M.F.J. Steel, 2000, „Measuring the sources of output growth in a panel of countries”, *Journal of Business and Economic Statistics*, w druku.
- [59] Koop R., J. Mullahy, 1989, „Moment-based estimation and testing of stochastic frontier models”, Discussion Papers no. 89-100, *Resources For the Future*, Washington.
- [60] Koop R., W.E. Diewert, 1982, „The decomposition of frontier cost function deviations into measures of technical and allocative inefficiency”, *Journal of Econometrics*, 19.
- [61] Kopczewski T., 1999, „Racjonalność zachowań banków komercyjnych w Polsce w latach 1994-97”, maszynopis pracy doktorskiej, Uniwersytet Warszawski.
- [62] Kraft E., D. Tirtiroglu, 1998, „Bank efficiency in Croatia: A stochastic-frontier Analysis”, *Journal of Comparative Economics*, 26.
- [63] Kumbhakar S.C., 1990, „Production frontiers, panel data, and time-varying technical inefficiency”, *Journal of Econometrics*, 46.
- [64] Kumbhakar S.C., 1997, „Modeling allocative inefficiency in translog cost function and cost share equations: an exact relationship”, *Journal of Econometrics*, ??.
- [65] Kumbhakar S.C., S. Ghosh, J.T. McGuckin, 1991, „A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms”, *Journal of Business and Economic Statistics*, 9.
- [66] Lawrence C., 1989, „Banking cost, generalized functional forms, and estimation of economies of scale and scope”, *Journal of Money, Credit, and Banking*, 3.
- [67] Lindley D.V., 1972, „*Bayesian Statistics, a Review*, Society for Industry and Applied Mathematics”, Philadelphia.
- [68] Lovell K.C. A., 1993, „Production frontiers and productive efficiency” w: Fried H.O., Lovell K.C.A., Schmidt S.S. (red.), *The Measurement of Productive Efficiency - Techniques And Applications*, Oxford University Press, New York, chap. 1.
- [69] Marzec J., 1998a, „Modelowanie procesu produkcji banków i badanie ich efektywności kosztowej” *Materialy XXXIV Konferencji Statystyków, Ekonometryków i Matematyków Akademii Ekonomicznych Polski Południowej* (Ustroń, maj 1998), Akademia Ekonomiczna w Katowicach.
- [70] Marzec J., 1998b, „Produkty i czynniki produkcji w badaniach efektywności kosztowej banków”, *Zastosowania rozwiązań informatycznych w bankowości* (materiały konferencyjne pod red. A. Gospodarowicza), Prace Naukowe AE we Wrocławiu nr 797.

- [71] Marzec J., 1999, „Produkty, czynniki produkcji i funkcja kosztów w badaniach efektywności kosztowej banków”, *Ekonomista*, 3.
- [72] Marzec J, Osiewalski J., 1996-97, „Pomiar efektywności kosztowej banków: zarys metodologii”, *Folia Economica Cracoviensia*, 39-40.
- [73] McAllister P.H., D. McManus, 1993, „Resolving the scale efficiency puzzle in banking”, *Journal of Banking and Finance*, 17.
- [74] Meeusen W., J. van den Broeck, 1977, „Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error”, *International Economic Review*, 1977, 8.
- [75] Mester L.J., 1987, „A multiproduct cost study of savings and loans”, *Journal of Finance*, 2.
- [76] Mester L.J., 1993, „Efficiency in the savings and loan industry”, *Journal of Banking and Finance*, 17.
- [77] Molyneux P., Y. Altunbas, E. Gardener, 1996, *Efficiency in European Banking*, J. Wiley & Sons, Chichester.
- [78] Muldur U., Sassenou M., 1993, „Economies of scale and scope in French banking and savings insitutions”, *Journal of Productivity Analysis*, 4.
- [79] Murray J., R. White, 1983, Economies of scale and economies of scope in multiproduct financial insitutions: A study of British Columbia Credit Unions”, *Journal of Finance*, 38.
- [80] Noulas A.G., C.R. Subhash, S.M. Miller, 1990, „Returns to scale and input substitution for large U.S. banks”, *Journal of Money, Credit, and Banking*, 22.
- [81] O’Hagan A., 1994, *Bayesian Inference*, J. Wiley, New York.
- [82] Osiewalski J., 1992, „Uogólnione niescentrowane współczynniki zwiększenia wariancji”, *Zeszyty Naukowe Akademii Ekonomicznej w Krakowie*, 374.
- [83] Osiewalski J., J. Marzec, 1998a, „Bayesian analysis of cost efficiency with an application to bank branches” w: *Global Trends and Changes in European Banking* (E. Miklaszewska, red.), Uniwersytet Jagielloński, Kraków, (s.151-166).
- [84] Osiewalski J., J. Marzec, 1998b, „Nowoczesne metody Monte Carlo w bayesowskiej analizie efektywności kosztowej banków”, *Zastosowania rozwiązań informatycznych w bankowości* (materiały konferencyjne pod red. A. Gospodarowicza), Prace Naukowe AE we Wrocławiu nr 797, s.182-195.
- [85] Osiewalski J., J. Marzec, 1998c, „Bayesowska analiza efektywności kosztowej oddziałów banku: założenia i wyniki”, *Prognozowanie w zarządzaniu firmą* (materiały konferencyjne pod red. M. Cieślak i D. Kwiatkowskiej-Ciotuchy), Prace Naukowe AE we Wrocławiu nr 808; s.24-33.
- [86] Osiewalski J., J. Marzec, M. Pipień, 1999, „Metody Monte Carlo w analizie bayesowskiej (na przykładzie modelu GARCH i granicznej funkcji kosztu)”, *Materiały XXXV Konferencji Statystyków, Ekonometryków i Matematyków Akademii Ekonomicznych Polski Południowej* (Osieczany, marzec1999), Akademia Ekonomiczna w Krakowie.

- [87] Osiewalski J., M.F.J. Steel, 1998, „Numerical tools for the Bayesian analysis of stochastic frontier models”, *Journal of Productivity Analysis*, 9.
- [88] Osiewalski J., M. Pipień, 1999, „Bayesowskie testowanie modeli GARCH i IGARCH”, *Przegląd Statystyczny*, 46.
- [89] Osiewalski J., R. Wróbel-Rotter, 1999, „Estymacja granicznych funkcji produkcji i wskaźników efektywności technicznej na podstawie danych przekrojowych”, *Przegląd Statystyczny*, 1.
- [90] Panzar J.C., R.D. Willig, 1977, „Economies of scale in multi-output production”, *Quarterly Journal of Economics*, 91.
- [91] Panzar J.C., R.D. Willig, 1981, „Economies of scope”, *American Economic Review*, 71.
- [92] Pesek B.P., 1970, „Ban's Supply Function and the Equilibrium Quantity of Money”, *The Canadian Journal of Economics*, 3.
- [93] Pitt M., L.F. Lee, 1981, „The measurement and sources of technical inefficiency in the Indonesian weaving industry”, *Journal of Development Economics*, 9.
- [94] Rangan N., R. Grabowski, N.Y. Aly, C. Pasurka, 1988, „The technical efficiency of US Banks”, *Economics Letters*, 28.
- [95] Richmond J., „Estimating the efficiency of production”, *International Economic Review*, 15.
- [96] Rogowski G., *Analiza i ocena działalności banków z wykorzystaniem metody DEA*, *Bank i Kredyt*, 1996, 9.
- [97] Rogowski G., 1998a, „Analiza efektywności oddziałów banku”, *Bank*, 1.
- [98] Rogowski G., 1998b, „Metodologia analiz efektywności i efektu skali banków”, *Bank i Kredyt*, 11.
- [99] Rogowski G., 1998c, *Metody analizy: oceny banku na potrzeby zarządzania strategicznego*, Wyższa Szkoła Bankowości, Poznań.
- [100] Schaffer S., E. David, 1986, „Economics of superscale and interstate expansion”, *Research Paper*, Federal Reserve Bank of New York.
- [101] Schmidt P., R. Sickles, 1984, „Production frontiers and panel data”, *Journal of Business and Economic Statistics*, 2.
- [102] Sealey C.W., J.T. Lindley, 1977, „Inputs, outputs, and a theory of production and cost at depository financial institutions”, *The Journal of Finance*, 32.
- [103] Shephard R.W., 1981, „Cost and production functions (reprint of first edition)”, *Lecture Notes in Economics and Mathematical System*, Springer-Verlag.
- [104] Sherman H.D., F. Gold, 1985, „Bank branch operating efficiency”, *Journal of Banking and Finance*, 9.
- [105] Stevenson R.E., 1980, „Likelihood functions for generalized stochastic frontier estimation”, *Journal of Econometrics*, 13.
- [106] Tarczyński W., 1998, „Analiza banków notowanych na Giełdzie Papierów Wartościowych”, *Bank i Kredyt*, 4.

- [107] Tierney L., 1994, „Markov chains for exploring posterior distributions”, *Annals of Statistics*, 22.
- [108] Towey R., 1974, „Money Creation and the Theory of Banking Firm”, *Journal of Finance*, 29.
- [109] Varian H.R., 1992, *Microeconomic Analysis*, Norton & Company, Inc., New York.
- [110] Varian H.R., 1995, *Mikroekonomia*, Wydawnictwo Naukowe PWN.
- [111] Winsten C.D., 1957, „Discussion on Mr. Farrell’s Paper”, *Journal of the Royal Statistical Society*, Series A, 120.
- [112] Youn Kim H., 1986, „Economics of scale and economics of scope in multiproduct financial institutions: futher evidence from credit unions”, *Journal of Money, Credit, and Banking*, 2.
- [113] Zardokoohi A., J. Kolari, 1994, „Branch office economies of scale and scope: Evidence from savings banks in Finland”, *Journal of Banking and Finance*, 18.
- [114] Zellner A., 1971, *An Indroduction to Bayesian Inference in Econometrics*, Wiley, New York.

## Załączniki

### Załącznik 1. Model VED ( $m=4, r^*=0.9$ ).

Tabela 15. Wartości oczekiwane i odchylenia standardowe a posteriori parametrów modelu: VED ( $m=4, r^*=0.9$ ).

Parametr	$E(\cdot y, X)$	$D(\cdot y, X)$
$\beta_0$	1.525	3.027
$\beta_1$	0.987	0.454
$\beta_2$	-0.506	0.587
$\beta_3$	0.631	0.222
$\beta_4$	-0.382	0.175
$\beta_5$	-0.014	0.043
$\beta_6$	-0.013	0.025
$\beta_7$	0.053	0.023
$\beta_8$	-0.039	0.015
$\beta_9$	0.001	0.011
$\beta_{10}$	-0.013	0.011
$\beta_{11}$	-0.041	0.036
$\beta_{12}$	0.091	0.040
$\beta_{13}$	0.018	0.008
$\beta_{14}$	0.036	0.009
$D^2(v_{i,t}) = \sigma_v^2$	0.000282	0.000038

Tabela 16. Średnia wartości oczekiwane i odchylenia standardowe a posteriori elastyczności: VED ( $m=4, r^*=0.9$ ).

	$\eta(VC/w_D \cdot 100)$	$\eta(VC/w_L)$	$\eta(VC/Q)$	$\eta(VC/K)$	$1 - \eta(VC/w_L) - \eta(VC/w_D)$
Srednia wartość oczekiwana	0.802	0.179	0.865	0.035	0.019
Srednie odchylenie standardowe	0.037	0.026	0.025	0.022	0.064

Tabela 17. Wartości oczekiwane i odchylenia standardowe a posteriori wskaźników efektywności modelu VED ( $m=4, r^*=0.9$ ).

$i$	$w_{i2}$	$w_{i3}$	$w_{i4}$	$E(\lambda_{iy}, X, W)$	$D(\lambda_{iy}, X, W)$	$E(r_i / y, X)$	$D(r_i / y, X)$
1	1	1	1	0.090	0.046	0.903	0.036
2	0	1	0	0.116	0.035	0.901	0.029
3	0	1	0	0.116	0.035	0.879	0.051
4	0	1	1	0.121	0.057	0.901	0.020
5	0	1	0	0.116	0.035	0.893	0.021
6	0	1	0	0.116	0.035	0.930	0.020
7	0	1	0	0.116	0.035	0.985	0.014
8	0	1	0	0.116	0.035	0.846	0.016
9	0	1	1	0.121	0.057	0.964	0.024
10	0	1	0	0.116	0.035	0.898	0.019
11	0	1	0	0.116	0.035	0.855	0.016
12	0	1	1	0.121	0.057	0.889	0.025
13	0	1	0	0.116	0.035	0.836	0.014
14	1	1	0	0.088	0.032	0.844	0.033
15	0	1	1	0.121	0.057	0.842	0.014
16	0	1	0	0.116	0.035	0.962	0.016
17	1	1	0	0.088	0.032	0.978	0.014
18	0	1	0	0.116	0.035	0.976	0.021
19	0	1	0	0.116	0.035	0.839	0.015

$i$	$w_{i2}$	$w_{i3}$	$w_{i4}$	$E(\lambda_{iY}, X, W)$	$D(\lambda_{iY}, X, W)$	$E(r_i / y, X)$	$D(r_i / y, X)$
20	0	1	0	0.116	0.035	0.854	0.015
21	0	1	0	0.116	0.035	0.967	0.015
22	1	1	0	0.088	0.032	0.903	0.015
23	0	1	0	0.116	0.035	0.922	0.026
24	1	1	1	0.090	0.046	0.888	0.014
25	0	0	0	0.086	0.027	0.894	0.014
26	0	0	0	0.086	0.027	0.984	0.012
27	0	0	0	0.086	0.027	0.833	0.013
28	0	0	1	0.093	0.055	0.934	0.024
29	0	0	0	0.086	0.027	0.938	0.015
30	1	0	0	0.062	0.015	0.858	0.014
31	0	0	0	0.086	0.027	0.955	0.014
32	0	0	0	0.086	0.027	0.910	0.016
33	1	0	0	0.062	0.015	0.976	0.013
34	0	0	0	0.086	0.027	0.966	0.013
35	1	0	0	0.062	0.015	0.967	0.014
36	0	0	0	0.086	0.027	0.961	0.014
37	1	0	0	0.062	0.015	0.835	0.013
38	1	0	0	0.062	0.015	0.946	0.014
39	1	0	0	0.062	0.015	0.948	0.015
40	1	0	0	0.062	0.015	0.909	0.014
41	1	0	0	0.062	0.015	0.912	0.014
42	1	0	0	0.062	0.015	0.973	0.016
43	1	0	0	0.062	0.015	0.945	0.014
44	1	0	0	0.062	0.015	0.963	0.013
45	1	0	0	0.062	0.015	0.993	0.007
46	1	0	0	0.062	0.015	0.949	0.014
47	1	0	0	0.062	0.015	0.943	0.014
48	1	0	0	0.062	0.015	0.911	0.013
49	1	0	0	0.062	0.015	0.959	0.016
50	1	0	0	0.062	0.015	0.963	0.013
51	1	0	0	0.062	0.015	0.921	0.016
52	1	0	0	0.062	0.015	0.951	0.013
53	1	0	0	0.062	0.015	0.871	0.015
54	1	0	0	0.062	0.015	0.933	0.014
55	1	0	0	0.062	0.015	0.987	0.011
<b>56</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0.062</b>	<b>0.015</b>	<b>0.996</b>	<b>0.004</b>
57	1	0	0	0.062	0.015	0.995	0.005
<b>58</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0.062</b>	<b>0.015</b>	<b>0.769</b>	<b>0.016</b>
Srednia dla oddzialow z $w_{i2}=1, w_{i3}=1$ i $w_{i4}=1$						0.896	0.025
Srednia dla oddzialow z $w_{i2}=0, w_{i3}=0$ i $w_{i4}=0$						0.930	0.013
Srednia dla oddzialow z $w_{i2}=1, w_{i3}=0$ i $w_{i4}=0$						0.935	0.013
Srednia dla oddzialow z $w_{i2}=0, w_{i3}=1$ i $w_{i4}=0$						0.903	0.021
Srednia dla oddzialow z $w_{i2}=0, w_{i3}=0$ i $w_{i4}=1$						0.934	0.024
Srednia dla oddzialow z $w_{i2}=1, w_{i3}=1$ i $w_{i4}=0$						0.909	0.021
Srednia dla oddzialow z $w_{i2}=0, w_{i3}=1$ i $w_{i4}=1$						0.899	0.021
Srednia dla oddzialow z $w_{i2}=1$						0.930	0.015
Srednia dla oddzialow z $w_{i2}=0$						0.911	0.019
Srednia dla oddzialow z $w_{i3}=1$						0.902	0.021
Srednia dla oddzialow z $w_{i3}=0$						0.934	0.014
Srednia dla oddzialow z $w_{i4}=1$						0.903	0.022
Srednia dla oddzialow z $w_{i4}=0$						0.923	0.016
Srednia dla oddzialow						0.921	0.017



## Załącznik 2. Model CED ( $r^*=0.7$ ).

Tabela 18. Wartości oczekiwane i odchylenia standardowe a posteriori parametrów modelu - CED ( $r^*=0.7$ ).

Parametr	$E(\cdot y,X)$	$D(\cdot y,X)$
$\beta_0$	1.941	2.935
$\beta_1$	1.056	0.449
$\beta_2$	-0.486	0.580
$\beta_3$	0.590	0.211
$\beta_4$	-0.447	0.172
$\beta_5$	-0.006	0.042
$\beta_6$	-0.022	0.024
$\beta_7$	0.061	0.022
$\beta_8$	-0.041	0.015
$\beta_9$	0.000	0.011
$\beta_{10}$	-0.011	0.011
$\beta_{11}$	-0.046	0.036
$\beta_{12}$	0.091	0.039
$\beta_{13}$	0.020	0.008
$\beta_{14}$	0.038	0.009
$D^2(v_{i,t}) = \sigma_v^2$	0.000273	0.000036

Tabela 19. Średnia wartości oczekiwane i odchylenia standardowe a posteriori elastyczności - CED ( $r^*=0.7$ ).

	$\eta(VC/w_D \cdot 100)$	$\eta(VC/w_L)$	$\eta(VC/Q)$	$\eta(VC/K)$	$1 - \eta(VC/w_L) - \eta(VC/w_D)$
Srednia wartość oczekiwana	0.800	0.181	0.865	0.037	0.019
Srednie odchylenie standardowe	0.036	0.026	0.024	0.024	0.062

Tabela 20. Wartości oczekiwane i odchylenia standardowe a posteriori wskaźników efektywności modelu CED ( $r^*=0.7$ ).

$i$	$w_{i2}$	$w_{i3}$	$w_{i4}$	$E(r_i / y, X)$	$D(r_i / y, X)$
1	1	1	1	0.904	0.034
2	0	1	0	0.903	0.029
3	0	1	0	0.873	0.051
4	0	1	1	0.900	0.020
5	0	1	0	0.891	0.021
6	0	1	0	0.928	0.020
7	0	1	0	0.983	0.015
8	0	1	0	0.844	0.017
9	0	1	1	0.968	0.022
10	0	1	0	0.900	0.019
11	0	1	0	0.850	0.017
12	0	1	1	0.886	0.024
13	0	1	0	0.833	0.014
14	1	1	0	0.833	0.033
15	0	1	1	0.839	0.014
16	0	1	0	0.960	0.017
17	1	1	0	0.973	0.016
18	0	1	0	0.978	0.020
19	0	1	0	0.831	0.016
20	0	1	0	0.850	0.015
21	0	1	0	0.959	0.017
22	1	1	0	0.897	0.015
23	0	1	0	0.920	0.028
24	1	1	1	0.885	0.014
25	0	0	0	0.887	0.014
26	0	0	0	0.981	0.013

$i$	$w_{i2}$	$w_{i3}$	$w_{i4}$	$E(r_i / y, X)$	$D(r_i / y, X)$
27	0	0	0	0.827	0.014
28	0	0	1	0.927	0.025
29	0	0	0	0.933	0.017
30	1	0	0	0.849	0.014
31	0	0	0	0.949	0.015
32	0	0	0	0.911	0.016
33	1	0	0	0.969	0.014
34	0	0	0	0.958	0.014
35	1	0	0	0.958	0.015
36	0	0	0	0.952	0.015
37	1	0	0	0.829	0.013
38	1	0	0	0.937	0.014
39	1	0	0	0.940	0.015
40	1	0	0	0.905	0.014
41	1	0	0	0.904	0.014
42	1	0	0	0.969	0.018
43	1	0	0	0.937	0.014
44	1	0	0	0.956	0.014
45	1	0	0	0.989	0.009
46	1	0	0	0.940	0.014
47	1	0	0	0.934	0.014
48	1	0	0	0.904	0.013
49	1	0	0	0.954	0.018
50	1	0	0	0.955	0.013
51	1	0	0	0.919	0.017
52	1	0	0	0.943	0.013
53	1	0	0	0.867	0.016
54	1	0	0	0.925	0.014
55	1	0	0	0.984	0.013
<b>56</b>	<b>1</b>	<b>0</b>	<b>0</b>	0.996	0.004
57	1	0	0	0.994	0.006
<b>58</b>	<b>1</b>	<b>0</b>	<b>0</b>	0.768	0.016
Średnia dla oddziałów z $w_{i2}=1$				0.908	0.019
Średnia dla oddziałów z $w_{i2}=0$				0.924	0.015
Średnia dla oddziałów z $w_{i3}=1$				0.900	0.021
Średnia dla oddziałów z $w_{i3}=0$				0.928	0.014
Średnia dla oddziałów z $w_{i4}=1$				0.901	0.022
Średnia dla oddziałów z $w_{i4}=0$				0.918	0.016
Średnia dla oddziałów				0.916	0.017