

**METODY ILOŚCIOWE
W BADANIACH EKONOMICZNYCH**

QUANTITATIVE METHODS
IN ECONOMICS

Vol. XV, No. 4

Warsaw University of Life Sciences – SGGW
Faculty of Applied Informatics and Mathematics
Department of Econometrics and Statistics

**METODY ILOŚCIOWE
W BADANIACH EKONOMICZNYCH**

**QUANTITATIVE METHODS
IN ECONOMICS**

Volume XV, No. 4

Warsaw 2014

EDITORIAL BOARD

Editor-in-Chief: Bolesław Borkowski

Deputy Editor-in-Chief: Hanna Dudek

Theme Editors:

Econometrics: Bolesław Borkowski

Multidimensional Data Analysis: Wiesław Szczesny

Mathematical Economy: Zbigniew Binderman

Analysis of Labour Market: Joanna Landmesser

Financial Engineering: Grzegorz Koszela

Statistical Editor: Wojciech Zieliński

Technical Editors: Jolanta Kotlarska, Elzbieta Saganowska

Language Editor: Agata Kropiwek

Native Speaker: Yochanan Shachmurove

Editorial Assistant: Monika Krawiec

SCIENTIFIC BOARD

Peter Friedrich (University of Tartu, Estonia)

Paolo Gajo (University of Florence, Italy)

Vasile Glavan (Moldova State University, Moldova)

Yuriy Kondratenko (Black Sea State University, Ukraine)

Vassilis Kostoglou (Alexander Technological Educational Institute of Thessaloniki, Greece)

Robert Kragler (University of Applied Sciences, Weingarten, Germany)

Karol Kukuła (University of Agriculture in Krakow)

Alexander N. Prokopenya (Brest State Technical University, Belarus)

Yochanan Shachmurove (The City College of The City University of New York, USA)

Mirbulat B. Sikhov (al-Farabi Kazakh National University, Kazakhstan)

Ewa Syczewska (Warsaw School of Economics, Poland)

Andrzej Wiatrak (University of Warsaw, Poland)

Dorota Witkowska (University of Lodz, Poland)

ISSN 2082 – 792X

© Copyright by Department of Econometrics and Statistics WULS – SGGW
(Katedra Ekonometrii i Statystyki SGGW)

Warsaw 2014, Volume XV, No. 4

The original version is the paper version

Journal homepage: qme.sggw.pl

Published by Warsaw University of Life Sciences Press

SPIS TREŚCI

Stanisław Gędek – Kształtowanie się kursu liry tureckiej wobec podstawowych walut światowych	7
Waldemar Karpa – Alternatywna metoda kalkulacji prognozy efektywności kosztowej w analizie farmakoeconomicznej na przykładzie grupy chorób niedokrwiennej serca	17
Andrzej Karpio, Dorota Żebrowska – Suchodolska – Efektywność inwestycyjna funduszy emerytalnych na tle FIO stabilnego wzrostu	26
Paweł Kliber – Estymacja miary martyngałowej na podstawie cen opcji z Giełdy Papierów Wartościowych w Warszawie	37
Krzysztof Kompa – Zastosowanie mierników taksonomicznych do oceny efektywności finansowej giełd europejskich w latach 2002-2011	52
Michał Kościółek – Wykorzystanie analizy wielokryterialnej do badania potencjału gospodarczego województwa podkarpackiego	62
Sławomir Lisek – Wrażliwość miary syntetycznej na wielkości krytyczne wskaźników służących do jej budowy	71
Anna Murawska – Ocena poziomu życia w krajach Unii Europejskiej (UE-28) w aspekcie zrównoważonego rozwoju za pomocą wielowymiarowej analizy porównawczej	80
Anna M. Olszewska, Alicja E. Gudanowska – Wykorzystanie wybranych metod porządkowania obiektów do klasyfikacji województw pod kątem ich potencjału innowacyjnego	91
Marcin Olszewski, Tomasz Ząbkowski – Zastosowanie technik eksploracji tekstu do analizy opinii konsumenckich	101
Dorota Pekasiewicz – Wybrane testy statystyczne dla wartości nietypowych i ich zastosowanie w analizach ekonometrycznych	111
Elżbieta Roszko-Wójtowicz – Ocena potencjału państw Unii Europejskiej do generowania innowacji z zastosowaniem analizy skupień	121
Robert Rusielik – Efektywność i produktywność rolnictwa w Polsce - analiza z wykorzystaniem indeksów TFP Hicksa-Moorsteena	137
Agnieszka Sompolska-Rzechuła – Wykorzystanie miar pozycyjnych w ocenie poziomu rozwoju społecznego w Polsce	149

Jacek Strojny – Relacja między eksportem rolnym i eksportem ogółem krajów UE	159
Ewa M. Syczewska – Przyczynowość w sensie Grangera – wybrane metody	169
Maria Szmuksta- Zawadzka, Jan Zawadzki – Zastosowanie wybranych modeli adaptacyjnych w prognozowaniu brakujących danych w szeregach ze złożoną sezonowością dla luk niesystematycznych	181
Andrzej Szuwarzyński – Wykorzystanie Data Envelopment Analysis do porównania polityki wiekowej w krajach Unii Europejskiej	195
Tomasz Śpiewak – Wykorzystanie wskaźników środków pieniężnych w analizie finansowej przedsiębiorstwa	205
Agnieszka Tłuczak – Zastosowanie modelu potencjału w analizie zróżnicowania regionalnego rolnictwa w krajach Unii Europejskiej	218
Maria Parlińska, Łukasz Pietrych, Iryna Petrovska – Ocena zróżnicowania produkcji mlecznej w Ukrainie z zastosowaniem wielowymiarowych metod statystycznych	229
Małgorzata Wrzosek – Model dynamiki rynku pracy z heterogeniczną siłą roboczą	236
Adam Zaremba, Przemysław Konieczka – Międzyrynkowe premie za wartość, wielkość i momentum na giełdach akcji	246

KSZTAŁTOWANIE SIĘ KURSU LIRY TURECKIEJ WOBEC PODSTAWOWYCH WALUT ŚWIATOWYCH

Stanisław Gędek

Katedra Ekonomii, Politechnika Rzeszowska
e-mail: gedeks@prz.edu.pl

Streszczenie: Zmienność kursów walutowych jest przedmiotem intensywnych badań, jednakże badania wzajemnego wpływu kursów walutowych podejmowane są bardzo rzadko. Gospodarka turecka rozwija się dynamicznie. Kursy liry tureckiej wykazują jednak znaczną zmienność, mogą więc być dobrym laboratorium służącym badaniu oddziaływania światowego rynku walutowego na kształtowanie się tych kursów. Przedmiotem opracowania jest weryfikacja hipotezy, iż kurs liry tureckiej do walut światowych jest determinowany przez kurs tej waluty do euro i dolara.

Słowa kluczowe: kursy walutowe, lira turecka, kointegracja, funkcja reakcji na impuls

WPROWADZENIE

Rynek walutowy jest największym i najszybciej rosnącym światowym rynkiem. Dzielne obroty na tym rynku osiągały w roku 2013 poziom 5,3 biliona dolarów, wobec około 4 bilionów dolarów w 2010 i jedynie około 1,2 biliona dolarów w roku 2001¹. Dla porównania produkt krajowy brutto Stanów Zjednoczonych w roku 2010 wynosił 16,8 biliona dolarów². Rynek ten funkcjonuje praktycznie 24 godziny na dobę, przez 7 dni w tygodniu.

Wprowadzenie systemu płynnych kursów walutowych w pierwszej połowie lat 70 ubiegłego stulecia spowodowało znaczne zwiększenie zakresu wahań w stosunku do występujących w okresie obowiązywania systemu z Bretton Woods.

¹ *Triennial Central Bank Survey. Foreign exchange turnover in April 2013: preliminary global results.* Monetary and Economic Department, Bank for International Settlements 2013.

² *International Monetary Fund: World Economic Outlook Database, April 2014.* <http://www.imf.org/external/pubs/ft/weo/2014/01/weodata/weorept.aspx> [dostęp 25.08.2014].

Zdarza się obecnie i to stosunkowo często, że zakres tych wahań przekracza 10% w ciągu dnia i 100% w ciągu roku [por. na przykład Karras i in. 2005].

Zmienność kursów walutowych jest czynnikiem determinującym zachowanie innych zmiennych makroekonomicznych, stąd też problematyka ta jest przedmiotem intensywnych badań. Próby wyjaśnienia przyczyn ich zmienności podejmowane są może rzadziej niż analiza skutków ich zmienności, literatura tego zagadnienia jest jednak dosyć obszerna [por. dla przykładu: Lanne i Vesala 2006, Tarashev 2007]. Problematyka wzajemnego oddziaływania kursów walutowych, jako przyczyna zmienności kursów walutowych podejmowane była bardzo rzadko, jest jednak obecna w literaturze zagadnienia [por. dla przykładu Witkowska 2011].

Przedmiotem niniejszego opracowania jest analiza zachowania kursu liry tureckiej (TRY) wobec podstawowych walut światowych. Lira turecka jest waluta należąca do tak zwanej „kruchej piątki” (*fragile five*) walut dużych, rozwijających się gospodarek, o dużych wahaniami kursów³. Przyczyną owej „kruchości” w przypadku liry jest duże zadłużenie przedsiębiorstw tureckich w walutach obcych i znaczne uzależnienie gospodarki tureckiej od eksportu, głównie na rynki Unii Europejskiej⁴, co czyni lirę turecką wrażliwą na oddziaływanie światowego rynku walutowego. Stąd też kursy liry tureckiej są dobrym laboratorium do badania wzajemnego oddziaływania kursów walutowych. Punktem wyjścia w badaniach przeprowadzonych w niniejszej pracy jest hipoteza, że na kurs liry tureckiej do walut światowych wpływa kurs tych walut do euro i dolara. Weryfikacja tej hipotezy jest celem niniejszego opracowania.

DANE EMPIRYCZNE I METODYKA BADAŃ

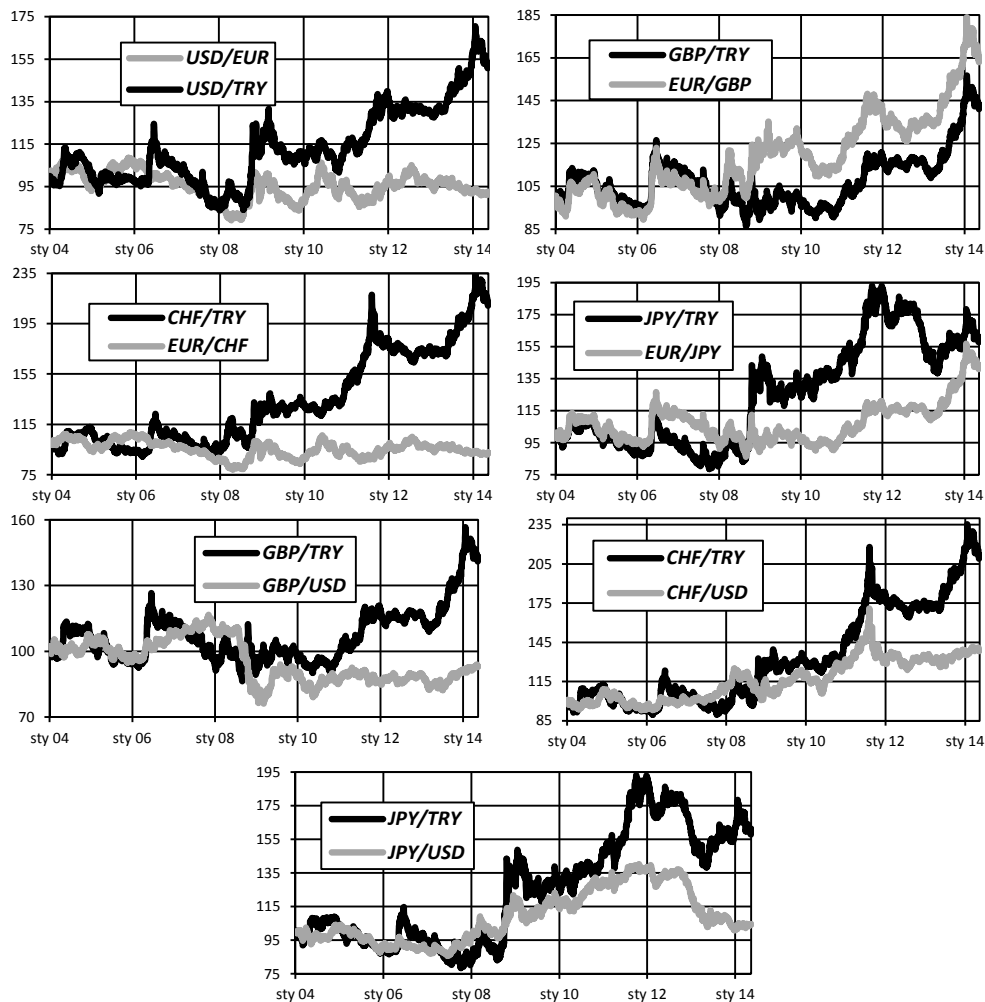
Analiza współzależności pomiędzy kursami liry tureckiej wobec podstawowych walut światowych a kursami tych walut wobec euro i dolara przeprowadzona została w oparciu o podawane przez Reuters⁵ zamknięcia dziennych notowań tych kursów. Szeregi czasowe tych kursów obejmowały przedział 05.01.2004 – 16.05.2014 i liczyły po 2705 obserwacji, czyli 541 tygodni 5 dniowych. Pojedyncze braki danych w poszczególnych szeregach czasowych były uzupełniane przy pomocy średnich z obserwacji sąsiednich. Na rysunku 1 przedstawione zostały wykresy indeksów szeregów czasowych tych kursów.

³ Pozostałe waluty to rupia indyjska, real brazylijski, rand południowoafrykański oraz rupia indonezyjska [por. na przykład Goldberg 2014].

⁴ *Turkey Regular Economic Note 2014-2*. The World Bank.
http://www.worldbank.org/content/dam/Worldbank/document/eca/turkey/WB_TREN_2014_2_03042014_ENG.pdf [dostęp 25.08.2014].

⁵ Dostępne na stronie stoq.com.

Rysunek 1. Przebieg indeksów badanych kursów walutowych



Źródło: obliczenia własne

Badania nad kształtowaniem się kursów walutowych wskazują, że są one realizacją procesu błędzenia losowego [por. na przykład Charles i Darne 2009, Kilian i Taylor 2002], co powoduje to, że szeregi czasowe kursów walutowych są niestacjonarne. Analiza przebiegu zmienności badanych kursów walutowych przedstawionych na rysunku 1 hipotezę tę potwierdza, podobnie jak wyniki testów stacjonarności zawarte w Tabeli 1. Na niestacjonarność badanych szeregów czasowych wskazuje zarówno test Grangera, jak i test KPSS⁶. Taka sytuacja „silnie wskazuje na niestacjonarność zmiennych” [Welfe 2009, str. 368].

⁶ Więcej na temat testów ADF i KPSS por. Welfe [2009], str. 360 i dalsze.

Tabela 1. Wyniki testów stacjonarności szeregów czasowych badanych kursów walutowych

Zmienna (kurs)	Test ADF				Test KPSS		Wartość krytyczna ($\alpha = 0,05$)
	Poziomy zmiennych		Pierwsze różnice		Statystyka testu		
	Statystyka testu	p	Statystyka testu	p	Poziomy zmiennych	Pierwsze różnice	
USD/TRY	-0,5631	0,8762	-49,9749	0,0000	18,485	0,120	0,461
EUR/TRY	-0,3340	0,9175	-51,7630	0,0000	21,115	0,122	
GBP/TRY	-0,0013	0,7899	-53,7462	0,0000	8,811	0,182	
CHF/TRY	0,0327	0,9604	-51,8999	0,0000	23,221	0,174	
JPY/TRY	-1,0539	0,7357	-52,8882	0,0000	21,898	0,084	
EUR/USD	-1,2028	0,6755	-53,5662	0,0000	5,309	0,037	
GBP/EUR	-1,4087	0,5795	-49,6499	0,0000	19,319	0,101	
CHF/EUR	-0,6543	0,8559	-51,8172	0,0000	22,288	0,136	
JPY/EUR	-1,4842	0,5417	-55,0394	0,0000	10,973	0,146	
GBP/USD	-1,7067	0,4278	-50,3518	0,0000	14,772	0,075	
CHF/USD	-1,2360	0,6609	-53,8435	0,0000	23,702	0,036	
JPY/USD	-1,3164	0,6241	-54,7301	0,0000	16,683	0,178	

Źródło: obliczenia własne

Niestacjonarność szeregów czasowych badanych kursów walutowych powoduje konieczność zastosowania do analizy ich współzależności metodyki opracowanej przez Engla i Grangera [Engle i Granger 1987], rozwiniętej następnie przez Johansena [Johansen 2000] oraz Johansena i Juselius [Johansen i Juselius 1990]. Idea tej procedury polega na przekształceniu modelu *VAR* (*Vector Auto Regression*) dla poziomów zmiennych o postaci:

$$\mathbf{x}_t = \mathbf{A}_0 \mathbf{d}_t + \sum_{i=1}^r \mathbf{A}_i \mathbf{x}_{t-i} + \mathbf{e}_t \quad (1)$$

gdzie: $\mathbf{x}_t = [x_{t1} \dots x_{tk}]^T$ – wektor obserwacji na bieżących wartościach pierwszych różnic zmiennych objaśnianych, $\mathbf{d}_t = [d_0 \ d_t \ d_{1\ t-1} \ d_{1\ t-r} \ d_{1\ t-r}]^T$ – wektor egzogenicznych składników równań, którego składowymi są odpowiednio: stała równania oraz bieżące i opóźnione wartości zmiennych egzogenicznych, \mathbf{A}_0 – macierz parametrów przy zmiennych wektora \mathbf{d}_t , \mathbf{A}_i – macierz parametrów przy opóźnionych zmiennych wektora \mathbf{x}_t , $\mathbf{e}_t = [\mathbf{e}_{1t} \dots \mathbf{e}_{kt}]^T$ wektory reszt równań modelu, r – rząd opóźnienia, do postaci *VECM* (*Vector Error Correction Model*)⁷:

$$\Delta \mathbf{x}_t = \Psi_0 \mathbf{d}_t + \Pi \mathbf{x}_{t-1} + \sum_{i=1}^r \Pi_i \Delta \mathbf{x}_{t-i} + \varepsilon_t \quad (2)$$

⁷ Sposób otrzymywania modelu *VECM* można znaleźć w pracach Millsa [2002] i Osińskiej [2006].

gdzie: Ψ_0 – macierz parametrów przy zmiennych wektora \mathbf{d}_t , $\Pi = \sum_{j=1}^k \mathbf{A}_j - \mathbf{I}$; $\boldsymbol{\varepsilon}_t$ -

reszty modelu. Macierz Π modelu (2) jest wykorzystywana w teście Johansena, którego ostatecznym efektem testu Johansena jest określenie modelu, w oparciu o który przeprowadzona zostanie analiza współzależności szeregów czasowych⁸. Mągą tu wystąpić trzy przypadki: macierz Π jest pełnego rzędu i wówczas model (2) jest modelem VAR dla poziomów zmiennych, rząd macierzy Π większy od 0 i mniejszy od k wskazuje na liczbę wektorów kointegracyjnych i wymaga zastosowania modelu (2), rząd macierzy Π jest 0 wówczas model (3) jest modelem VAR dla przyrostów zmiennych (pierwszych różnic).

W oparciu o odpowiedni model (VAR dla poziomów zmiennych, VECM lub VAR dla pierwszych różnic) przeprowadzana jest analiza przyczynowości⁹. Potwierdzenie wystąpienia związku przyczynowego pozwala na przeprowadzenie analizy funkcji odpowiedzi na impuls (*Impulse Response Function – IRF*) będącej dopełnieniem analizy przyczynowości dając możliwość określenia kierunku oddziaływania impulsu (przyczyny), siły tego impulsu oraz rozkładu w czasie i szybkości jego wygasania.

Badanie współzależności kursów korony szwedzkiej przeprowadzone zostało w ten sposób, że test Johansena przeprowadzony został dla par szeregów czasowych: kurs liry tureckiej (TRY) do danej waluty i kurs tej waluty do euro, kurs TRY do danej waluty i kurs tej waluty do dolara amerykańskiego. Następnym krokiem była estymacja odpowiednich modeli dla wymienionych par kursów. W dalszej kolejności, w oparciu o wyniki estymacji modeli przeprowadzona została analiza przyczynowości. Do analizy przyczynowości wykorzystany został test F służący badaniu braku restrykcji w modelu ekonometrycznym¹⁰. Podsumowaniem badań będzie analiza funkcji odpowiedzi na impuls (IRF), która pozwala na opis przebiegu w czasie zależności pomiędzy badanymi zmiennymi. Badanie kointegracji i estymacja parametrów modeli opisujących współzależność szeregów czasowych badanych kursów walutowych wykonana została przy pomocy programu GRET, również przy pomocy tego programu wyznaczone zostały wartości funkcji *IRF*.

⁸ Więcej na temat testu Johansena por. Johansen S. [2000].

⁹ Chodzi tu o tak zwaną przyczynowość w sensie Grangera. Więcej na ten temat por. Charemza i Deadman [1997] rozdział 6.3, M. Osińska, op. cit., str.212.

¹⁰ Test ten pozwala na zbadanie czy wprowadzenie do modelu zmiennej lub zbioru zmiennych zmniejszy istotnie wariancję resztową modelu, odpowiada więc na to samo pytanie, na które odpowiada stosowany zazwyczaj test Grangera, a jest znacznie łatwiejszy w stosowaniu, ponieważ w przypadku większości pakietów statystycznych, obsługujących analizę regresji, wartości statystyki tego testu i prawdopodobieństwa błędów I rodzaju podawane są automatycznie.

WYNIKI BADAŃ

Przedział czasowy, w którym badane było kształtowanie się kursu liry tureckiej wobec podstawowych walut światowych musiał być podzielony na okresy ze względu na zmienność parametrów opisujących równania tych zależności. Stabilność tych parametrów badana była przy pomocy testu *QLR*¹¹. W Tabeli 2 przedstawiony został podział badanego przedziału czasowego na okresy, w których parametry równań opisujących wzajemne zależności badanych kursów są stabilne.

Tabela 2. Podział badanego przedziału czasowego na okresy

Numer okresu	Okres (od – do)	Liczba obserwacji
1	2004.01.05 – 2004.04.30	85
2	2004.05.03 – 2004.11.19	145
3	2004.11.22 - 2006.04.07	360
4	2006.04.10 - 2006.12.15	180
Luka	2006.12.18. – 2006.12.30	10
5	2007.01.01 – 2008.02.29	305
Luka	2008.03.03 – 2008.04.04	25
6	2008.04.07 – 2008.09.26	125
7	2008.09.29 – 2009.07.03	200
8	2009.07.06 – 2010.07.02	260
9	2010.07.05 – 2011.07.15	270
Luka	2011.07.18 – 2011.08.12	20
10	2011.08.15 – 2012.06.29	230
11	2012.07.02 – 2013.04.19	210
12	2013.04.22 – 2014.05.16	280

Źródło: obliczenia własne

W Tabeli 2 występują trzy okresy oznaczone jako „luka”. Zostały one wykluczone z badania, gdyż włączenie danych do któregośkolwiek z sąsiednich okresów powodowało, że równania opisujące kształtowanie się kursów liry tureckiej były niestabilne. Niestabilne były również parametry równań opisujące kształtowanie się kursów liry tureckiej w okresach wydzielonych jako „luka”.

Wyniki testu Johansena dla par kursów walutowych (kurs liry tureckiej do danej waluty i kurs tej waluty do euro bądź dolara) dla każdego z wydzielonych okresów przedstawione zostały w Tabeli 3 (w przypadku kursów liry tureckiej do euro i do dolara zmienną objaśniającą ich kształtowanie się był kurs EUR/USD). Wyniki tego testu przedstawione zostały w formie rzędu macierzy. Zgodnie z podanym wcześniej opisem procedury Johansena rząd macierzy Π równy zero oznacza, że do opisu zależności powinien być wykorzystany model VAR dla

¹¹ *Quandt Likelihood Ratio*. Więcej na temat testu QLR por. Stock i Watson [2007], str. 567 i dalsze.

przyrostów zmiennych, gdy rząd tej macierzy jest równy 1 występuje jeden wektor kointegracyjny i należy wykorzystać model *VECM*, a w przypadku gdy jest równy 2, a więc pełny, powinien zostać wykorzystany model *VAR* dla poziomów zmiennych.

Tabela 3. Rząd macierzy Π w teście Johansena dla par kursów walutowych

numer	Okres od – do	USD/TRY		EUR/TRY		GBP/TRY		CHF/TRY		JPY/TRY	
		EUR/USD	EUR/USD	USD	EUR	USD	EUR	USD	EUR	USD	EUR
1	2004.01.05 -	2	2	0	0	0	1	2	2	2	2
2	2004.05.03 -	2	2	2	2	2	2	2	1	1	1
3	2004.11.22 -	2	2	2	2	2	2	2	2	2	2
4	2006.04.10 -	2	2	2	2	2	2	2	2	2	2
5	2007.01.01 -	2	2	2	0	0	1	1	1	1	1
6	2008.04.07 -	2	2	2	2	2	2	2	2	2	2
7	2008.09.29 -	2	2	2	2	2	2	2	2	2	2
8	2009.07.06 -	2	2	0	0	2	0	2	2	2	2
9	2010.07.05 - 2011-	2	2	2	2	0	2	2	2	2	2
10	2011.08.15 -	2	2	2	2	0	0	0	0	0	0
11	2012.07.02 -	2	2	2	2	1	0	2	2	2	2
12	2013.04.22 -	2	2	0	0	2	2	0	0	0	0

Źródło: obliczenia własne

W Tabeli 4 zawarte zostały wyniki testu przyczynowości. Kolorem ciemnym szarym zaznaczone zostały te przypadki, w których kurs danej waluty do euro lub do dolara jest przyczyną (w sensie Grangera) dla kursu liry tureckiej do danej waluty potwierdzoną przez test *F* na poziomie istotności równym 0,05, zaś jasnym szarym, gdy istotność jest równa 0,1. Pola niezaciemnione odnoszą się do przypadków, w których nie można potwierdzić występowania związku przyczynowego między kursem danej waluty do euro lub dolara, a kursem liry tureckiej do tej waluty.

Tabela 4. Wyniki testu przyczynowości

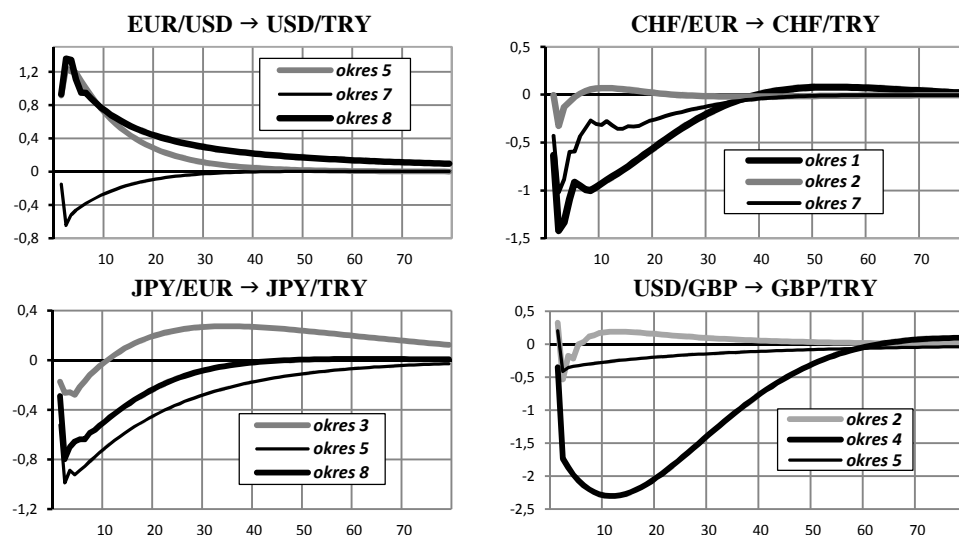
numer	Okres od – do	USD/TRY		EUR/TRY		GBP/TRY		CHF/TRY		JPY/TRY	
		EUR/USD	EUR/USD	USD	EUR	USD	EUR	USD	EUR	USD	EUR
1	2004.01.05 -										
2	2004.05.03 -										
3	2004.11.22 -										
4	2006.04.10 -										
5	2007.01.01 -										
6	2008.04.07 -										
7	2008.09.29 -										
8	2009.07.06 -										
9	2010.07.05 - 2011-										
10	2011.08.15 -										
11	2012.07.02 -										
12	2013.04.22 -										

Źródło: obliczenia własne

Wyniki zawarte w Tabeli 4 dają obraz niezbyt jednoznaczny. Można jednak stwierdzić, że wpływ kursów walut światowych na kursy liry tureckiej potwierdzony przez test przyczynowości, był mocno zróżnicowany w czasie. Nie był widoczny na początku badanego okresu, widać też, iż wpływ ten jest wyraźnie słabszy od połowy roku 2010.

Więcej o kształtowaniu się reakcji badanych zależności pomiędzy kursami walutowymi może powiedzieć funkcja odpowiedzi na impuls (*IRF*). Na Rysunku 2 przedstawione zostały przykładowe wykresy przebiegu funkcji *IRF*, opisujące rozkład w czasie reakcji kursów liry tureckiej na impuls ze strony walut światowych. Wykres funkcji *IRF* skonstruowany został w ten sposób, że na osi rzędnych wykresu odłożona jest wielkość reakcji odpowiedniego kursu korony szwedzkiej na zmianę kursu danej waluty o wielkość równą średniej tego kursu w danym okresie, a na osi odciętych wyrażony w dniach horyzont czasowy oddziaływania tego impulsu.

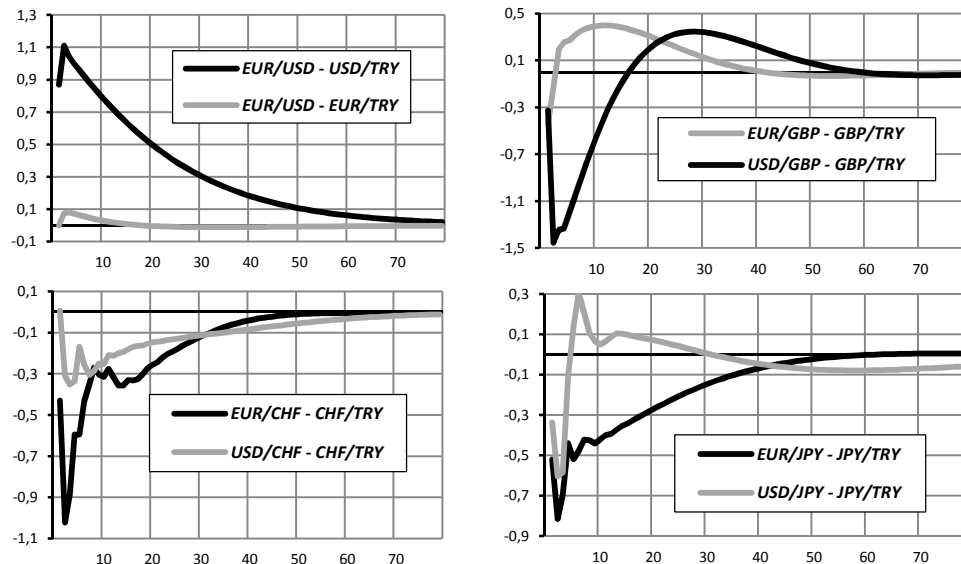
Rysunek 2. Funkcje odpowiedzi na impuls - porównanie reakcji kursów liry tureckiej w różnych okresach



Źródło: obliczenia własne

Przebieg funkcji odpowiedzi na impuls zamieszczonych na rysunku 2 pokazuje, że reakcja kursów liry tureckiej na zmiany kursów walut światowych była bardzo zróżnicowana w czasie w wszystkich trzech wymiarach analizy *IRF*, szybkości wygasania, siły oddziaływania a nawet kierunku oddziaływania. Zróżnicowana, również w tych trzech wymiarach była reakcje w wydzielonych okresach reakcja kursu liry tureckiej do podstawowych walut światowych na zmiany kursu tych walut do euro o do dolara, na co widać wyraźnie wskazuje rysunek 3.

Rysunek 3. Funkcje odpowiedzi na impuls - porównanie reakcji kursów liry tureckiej na zmianę kursów walut światowych do euro i do dolara w okresie 7 (2008.09.29 - 2009.07.03)



Źródło: obliczenia własne

Kursy liry tureckiej do walut światowych, jak wykazała analiza tu przeprowadzona, pozostawały w badanym okresie pod wpływem światowego rynku walutowego. Parametry tego wpływu były jednak bardzo zmienne w czasie, co stawia pod znakiem zapytania możliwość wykorzystania zmian zachodzących na światowym rynku walutowym do prognozowania kursów liry tureckiej.

PODSUMOWANIE

Przeprowadzona analiza zachowania kursów liry tureckiej korony szwedzkiej do podstawowych walut światowych pozwoliła na częściowe potwierdzenie hipotezy postawionej we wstępie. Stwierdzone zostało, iż kursy liry tureckiej pozostawały w części wydzielonych okresów badanego przedziału czasowego pod wpływem kursów podstawowych walut światowych do euro i dolara. Wpływ ten był jednakże zmienny w czasie, na co wskazywał przebieg funkcji odpowiedzi na impuls. Daje to bardzo rozmyty obraz wpływu światowego rynku walutowego na kursy liry tureckiej. Konieczne wydaje się przeprowadzenie badań, w których przedstawione tu modele opisujące zachowanie liry tureckiej zostaną uzupełnione o zmienne charakteryzujące politykę monetarną Turcji.

BIBLIOGRAFIA

- Charemza W.W., Deadman D.F. (1997) Nowa ekonometria. PWE, Warszawa 1997.
- Charles A., Darne O. (2009) Testing for Random Walk Behavior in Euro Exchange Rates, *Economie Internationale*, Tom 119, str. 25 – 45.
- Engle R.F., Granger C.W.J. (1987) Co-integration and Error Correction, *Econometrica*, Tom 55 str. 251 – 276.
- Goldberg P. (2014) Emerging Markets Strategist. 2014 Outlook: Baby Steps on Shaky Grounds, HBSC Research, New York.
- Johansen, S. (1988) Statistical Analysis of Cointegration Vectors' *Journal of Economic Dynamics and Control*, Tom 12, str. 231 – 254.
- Johansen S. (2000) Modelling of cointegration in the vector autoregressive model, *Economic Modelling*, Tom 17 (2000), str. 359 – 373.
- Johansen, S., Juselius, K. (1990) Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics*, Tom 52, str. 169 – 210.
- Karras G., Lee J.M., Stokes H. (2005) Sources of Exchange-Rate Volatility. Impulses or Propagation? , *International Review of Economics and Finance* 14, str. 213–226.
- Kilian L., Taylor M.P. (2002) Why is it so difficult to beat the random walk forecast of exchange rates?, *Journal of International Economics*, Tom 60, str. 85 – 107.
- Lanne M., Vesala T. (2006) The effect of a transaction tax on exchange rate volatility, *Bank of Finland Research Discussion Papers*, 11/2006.
- Mills T.C. (2002) *The Econometric Modeling of Financial Time Series*, Cambridge University Press, Cambridge.
- Osińska M. (2006) *Ekonometria finansowa*, PWE, Warszawa.
- Stock J.H., Watson M.W., *Introduction to Econometrics*, Addison Wesley, Boston 2007.
- Tarashev N.A. (2007) Speculative Attacks and the Information Role of Interest Rate, *Journal of the European Economic Association*, Tom 5, str. 1-36.
- Welfe A. (2009) *Ekonometria. Metody i ich zastosowanie*, PWE, Warszawa.

THE BEHAVIOR OF TURKISH LIRA EXCHANGE RATES

Abstract: The volatility of the exchange rates is the subject of intensive research, however, the research of mutual influence of exchange rates are taken very rarely. On the other hand, is assumed the existence the convergence of smaller currencies. The Turkish economy is developing dynamically. Turkish Lira is currently stable, the its exchange rates are a good laboratory for study the impact of dominant currency (euro). The subject of the paper is the verification of hypothesis that exchange rate of Turkish lira to main currencies are determined by the exchange rate of this currency to the euro.

Keywords: currency exchange rates, Turkish lira, cointegration, impulse response function

ALTERNATYWNA METODA KALKULACJI PROGU EFEKTYWNOŚCI KOSZTOWEJ W ANALIZIE FARMAKOEKONOMICZNEJ NA PRZYKŁADZIE GRUPY CHORÓB NIEDOKRWIENNYCH SERCA

Waldemar Karpa

Katedra Ekonomii, Akademia Leona Koźmińskiego
e-mail: wkarpa@kozminski.edu.pl

Streszczenie: Artykuł przedstawia alternatywną metodę kalkulacji progów efektywności kosztowej w procesie oceny technologii medycznych. Metoda oparta na teoretycznym modelu funkcji zdrowia wykorzystuje rezultat analizy ekonometrycznej w kalkulacji kosztu życia w oparciu o historyczne dane epidemiologiczne i alokację zasobów finansowych w publicznym systemie ochrony zdrowia. Metodę demonstrujemy na przykładzie grupy chorób niedokrwiennych serca. Wyniki analizy wskazują, iż obecnie stosowany w Polsce sztywny próg efektywności nie odzwierciedla rzeczywistych kosztów związanych z poprawą efektu zdrowotnego.

Słowa kluczowe: analiza farmakoekonomiczna, analiza efektywności kosztów chorób, próg efektywności kosztowej

WPROWADZENIE

W farmakoekonomicznej analizie efektywności kosztowej pozwala ocenić, które z porównywanych technologii medycznych jest bardziej opłacalne, a więc umożliwia określenie względnej wartości tych technologii. Analiza efektywności kosztowej przeprowadzana jest w oparciu o analizę porównawczą stosunku nakładów do efektów zdrowotnych związanych z użyciem alternatywnych technologii medycznych. Powyższa analiza jest istotnym etapem procesu oceny technologii medycznych i zarazem nieodzownym składnikiem nowoczesnego i efektywnego systemu opieki zdrowotnej. Ten interdyscyplinarny proces uwzględniający m.in. medyczne, społeczne, ekonomiczne, prawne i etyczne aspekty zastosowania danej technologii dostarcza naukowych podstaw do podejmowania racjonalnych decyzji dotyczących stosowania i finansowania świadczeń

zdrowotnych. Kryteria oceny jak i stosowana metodologia mają więc ogromny wpływ na sytuację pacjentów zmagających się z wyzwaniem jakim jest choroba. System opieki zdrowotnej nie jest jednak doskonały, gdyż ograniczone zasoby wymuszają konieczność niekoniecznie egalitarnego podziału środków. Aspekt skuteczności zdrowotnej oraz wydajności ekonomicznej w dużej mierze warunkują szansę na zdrowie. Nadrzędnym celem oceny technologii medycznych jest zatem prowadzenie bezpiecznej oraz skutecznej polityki zdrowotnej zorientowanej na pacjenta, przy jednoczesnej efektywnej alokacji zasobów.

W Polsce organem, którego zadaniem jest ewaluacja technologii medycznych jest Agencja Oceny Technologii Medycznych (AOTM). W ramach stosowanej analizy ekonomicznej AOTM zainteresowana jest wartością monetarną efektu zdrowotnego związanego z nowym lekiem (procedurą) - na tej podstawie bowiem Agencja dokonuje rekomendacji w sprawie refundacji leków przez publicznego płatnika. W ramach obowiązującej ustawy refundacyjnej¹, w Polsce obowiązuje sztywny próg opłacalności w wysokości trzykrotności PKB per capita (111 381 zł). Jest on zdefiniowany jako monetarna granica uzyskania dodatkowego roku życia skorygowanego o jakość (QALY)². Ustanowienie w 2012 r. jednakowego progu opłacalności dla wszystkich chorób i procedur, przyczyniło się do wykluczenia wielu nowoczesnych i kosztownych terapii, programów zdrowotnych i leków, eliminując tym samym szansę wielu pacjentów na powrót do zdrowia. Szczególnie zagrożeni zostali pacjenci cierpiący na choroby onkologiczne oraz rzadkie, których terapia wymaga terapii bazujących na lekach sierocych.

W zaistniałej sytuacji istnieje potrzeba zaproponowania i oszacowania alternatywnego progu efektywności kosztowej, uwzględniającej specyfikę chorób (choroba powszechne, choroby przewlekłe oraz choroby rzadkie). Celem niniejszej pracy jest oszacowanie progu opłacalności na przykładzie grupy schorzeń (choroby niedokrwienne serca) na podstawie historycznych danych Narodowego Funduszu Zdrowia (NFZ). Nasze opracowanie bazuje na ekonometrycznej analizie funkcji zdrowia, która łączy wydatki NFZ oraz dane epidemiologiczne odnośnie umieralności. Otrzymany w ten sposób próg opłacalności odnosimy do efektu zdrowotnego.

Opracowanie składa się z trzech części. W części pierwszej przedstawiamy założenia teoretyczne modelu zdrowia; kolejna część opracowania poświęcona jest analizie ekonometrycznej modelu teoretycznego wykorzystującej dane empiryczne;

¹ Ustawa z dnia 12 maja 2011 r. o refundacji leków, środków spożywczych specjalnego przeznaczenia żywieniowego oraz wyrobów medycznych (Dz. U. z dnia 13 czerwca 2011 r.).

² QALY (ang. *Quality-Adjusted Life Year*): Metoda polega na pomiarze użyteczności procedur medycznych w systemie opieki zdrowotnej. Wskaźnik QALY opiera się na: (a) liczbie lat zyskanej dzięki danej interwencji medycznej oraz (b) jakości życia w uzyskanych dzięki terapii latach życia. Wartość współczynnika otrzymujemy poprzez pomnożenie obu współczynników.

część trzecią poświęcamy na dyskusję wyników analizy w perspektywie efektywności systemu opieki zdrowotnej w naszym kraju.

METODOLOGIA BADANIA

Model teoretyczny

Modelowanie progu efektywności kosztowej w naszym badaniu oparte jest o relację pomiędzy wydatkami NFZ a epidemiologicznymi danymi o śmiertelności. W badaniu częściowo bazujemy na modelu brytyjskiego National Institute for Health and Clinical Excellence (NICE) [Claxton i in. 2013], instytucji, w której narodziła się idea oparcia wyceny efektywności kosztowej technologii medycznych na koszcie roku życia skorygowanego o jakość.

U podstaw modelu leży funkcja zdrowia $H(\cdot)$, będąca funkcją indywidualnych funkcji użyteczności:

$$H = [u_1(h), \dots, u_n(h)] \quad (1)$$

gdzie h jest kondycją zdrowotną członków populacji. Funkcja zdrowia w swojej formie i założeniach nawiązuje do funkcji dobrobytu społecznego i może być uważana za jej szczególną formę. Zgodnie z użytkowym podejściem, jedyną społecznie sprawiedliwą alokacją zasobów jest ta, która maksymalizuje sumę użyteczności indywidualnych, a więc w przypadku publicznego systemu opieki zdrowotnej, funkcja zdrowia przyjmuje postać:

$$H(u_1, \dots, u_n) = \sum_{i=1}^n u_i \quad (2)$$

gdzie u_i jest użytecznością i -tej osoby. Kondycja zdrowotna populacji (a więc efekty zdrowotne w poszczególnych obszarach systemu zdrowia)³ zależy od wielu czynników, wśród których najważniejszymi są: wielkość wydatków publicznych na opiekę zdrowotną (x) oraz wielkość popytu na świadczenia w publicznym systemie opieki zdrowotnej (n). Ponadto, na kondycję zdrowotną populacji wpływają inne czynniki (z), wśród których wymienić należy przede wszystkim poziom współpłacenia za usługi zdrowotne. Kondycję zdrowotną możemy więc przedstawić jako:

$$h_j = f(x_j, n_j, z_j), j = 1, \dots, m \quad (3)$$

gdzie j identyfikuje poszczególne obszary zdrowotne. Jeżeli przyjmiemy, że w publicznym systemie opieki zdrowotnej funkcję alokacyjną zasobów pełni NFZ,

³ Przyjmujemy, że cała populacja należy do publicznego systemu opieki zdrowotnej, a efekty zdrowotne odzwierciedlone są w danych epidemiologicznych.

możemy założyć, iż zadaniem Funduszu jest wybór optymalnego poziomu wydatków (x^*) tak aby zmaksymalizować użyteczność zdrowotną opisaną w równaniu (2), przy jednoczesnym uwzględnieniu czynników wpływających na efekt zdrowotny (3) oraz ograniczoności środków. Owe ograniczenie budżetowe przyjmuje poniższą formę:

$$x_1 + x_2 + \dots + x_m \leq y \quad (4)$$

gdzie y wyraża całkowitą wielkość zasobów Funduszu.

Specyfikacja ekonometryczna i źródła danych

Przedstawiony powyżej model teoretyczny wskazywałby na konieczność użycia równań współzależnych, odzwierciedlających kompleksowość i złożoność opisywanych procesów. Poziom agregacji danych stosowany w sprawozdawczości Funduszu jak i dane epidemiologiczne stanowią jednak istotną przeszkodę w modelowaniu. Odpowiedni stopień dezagregacji danych nie jest zjawiskiem lokalnym, właściwym wyłącznie dla Polski – jest to problem często spotykany w innym krajach europejskich, na co wskazuje analiza literatury [Cremieux i in. 2009]. Zdajemy sobie także sprawę, iż kierunek zależności pomiędzy wydatkami a wskaźnikiem efektu zdrowotnego w postaci umieralności może budzić kontrowersje. Rozważania nt. trudności metodologicznych związanych z badaniem zależności pomiędzy stanem zdrowotności i wydatkami na ochronę zdrowia spotykamy m.in. u Gravelle'a i Backhouse'a [Gravelle, Backhouse 1987] oraz Nixona i Ulmanna [Nixon, Ulmann 2006]. Autorzy wskazują na możliwy problem endogeniczności relacji (aktualny poziom wydatków jest najprawdopodobniej dostosowany do zastanych efektów zdrowotnych w przeszłości) oraz opóźnień przyczynowo-skutkowych (np. wzrost wydatków nie przekłada się w sposób natychmiastowy na poprawę efektu zdrowotnego). W naszym badaniu wyeliminowaliśmy częściowo problem endogeniczności odrzucając model równań współzależnych. Założyliśmy także, iż istnieje luka czasowa pomiędzy oczekiwanym efektem zdrowotnym a momentem wydatkowania. Dysponując danymi z 16 oddziałów wojewódzkich NFZ oraz regionalnymi wskaźnikami śmiertelności, forma funkcjonalna specyfikacji ekonometrycznej ma charakter danych panelowych i jest wyrażona jako:

$$m_{k,t} = \beta_0 + \beta_1 x_{k,t-1} + \beta_2 n_{k,t} + \beta_3 staff_{k,t} + a_i + u_{k,t} \quad (5)$$

gdzie k jest indeksem porządkowym oddziałów wojewódzkich NFZ, t – indeksem czasu, a_i – opisuje niezaobserwowaną heterogeniczność regionów (a więc wszystkie niezmiennie czasowo pominięte w modelu zmienne charakterystyczne dla poszczególnych regionów, takie jak np. poziom uprzemysłowienia czy jakość środowiska naturalnego, mające wpływ na zdrowie populacji), a $u_{k,t}$ – określa

pozostałe niezaobserwowane czynniki (zmiennie w czasie), wpływające na efekt zdrowotny. Miernikiem owego efektu w naszym badaniu jest wskaźnik umieralności (m), rozumiany jako bezwzględna liczba zgonów wg. jednostek chorobowych (Międzynarodowa Statystyczna Klasyfikacja Chorób i Problemów Zdrowotnych ICD-10). Informację w ujęciu regionalnym zawierają roczniki demograficzne GUS. Wydatki NFZ (x) są wydatkami Funduszu w podziale na jednostki chorobowe (ICD-10). Wielkość popytu na świadczenia (n) odzwierciedla wartość kontraktów Funduszu, które pochodzą z rocznych Sprawozdań z działalności Funduszu. Zbiór danych zawiera również informację o liczbie praktykujących lekarzy w regionie (*staff*).

WYNIKI BADANIA

Empiryczna analiza modelu ekonometrycznego została przeprowadzona w oparciu o grupę chorób niedokrwiennych serca (I20-I25)⁴ i obejmuje lata 2009 - 2011⁵. Wybór jest nieprzypadkowy, gdyż wraz ze wszystkim podtypami, ta grupa schorzeń stanowi najczęstszą przyczynę zgonów w większości państw zachodnich⁶. Wyniki analizy przedstawiamy w dwóch częściach: część pierwsza poświęcona jest rezultatom estymacji ekonometrycznej; w części drugiej wyprowadzamy próg efektywności kosztowej w oparciu o wyniki szacowanego równania.

Rezultat procedury ekonometrycznej

Proces estymacji przeprowadziliśmy w oparciu o model ‘log-log’, wprowadzając wartości zmiennych modelowych w postaci logarytmicznej. Zabieg ten pozwala na interpretację oszacowanych wartości współczynników zmiennych modelowych jako elastyczności. Ponadto, wielkość wydatków oraz wartość kontraktów NFZ przedstawione są per capita, co gwarantuje wewnętrzną spójność danych i jednocześnie pozwala kontrolować efekt skali. Rezultat szacowania modelu (5) na podstawie uogólnionego modelu regresji liniowej (GLS) dla panelu 2009-2011 przedstawia tabela poniżej:

⁴ Grupę stanowią następujące jednostki chorobowe: I20-dusznicza bolesna, I21-ostry zawał serca, I22-ponowny zawał serca, I23- niektóre powikłania występujące w czasie ostrego zawału serca, I24- inne ostre postacie choroby niedokrwiennej serca oraz I25- przewlekła choroba niedokrwienne serca.

⁵ Dane o wielkości wydatków zgodnie z modelem są opóźnione.

⁶ Światowa Organizacja Zdrowia (2013) The top ten causes of death', Fact sheet No. 310, World Health Organization, <http://who.int/mediacentre/factsheets/fs310/en/>, dostęp: 02.06.2014

Tabela 1. Wyniki estymacji modelu (5)

	ln(m)
ln(x)	-.0415** (.029)
ln(n)	-.0009* (.010)
ln(staff)	-.015** (.023)
Const.	9.34***
N=48	
R ² =.35	

Źródło: obliczenia własne przy wykorzystaniu narzędzia STATA 12;

***, **, * - istotność statystyczna na poziomie 1%, 5%, 10%

Współczynnik determinacji R^2 wynosi 0,35, co w przypadku modelu panelowego jest dobrym wynikiem i oznacza, że zaproponowana przez nas specyfikacja wyjaśnia prawie 35 procent zmienności zbioru. Ponadto przeprowadzony test Hausmana potwierdza słuszność wyboru estymacji modelu panelowego z efektami indywidualnymi, a w konsekwencji zastosowania uogólnionej metody najmniejszych kwadratów. Estymowane współczynniki regresji dla zmiennych x oraz $Staff$ posiadają oczekiwane znaki oraz są statystycznie istotne. Ujemna wartość współczynnika β_1 świadczy o pozytywnym wpływie wydatków zdrowotnych na zmniejszenie umieralności, a więc wpływa bezpośrednio na polepszenie efektu zdrowotnego. Widoczna jest także rola personelu medycznego, gdyż model prognozuje, iż wzrost liczby lekarzy w danym regionie powinien także przyczynić się do polepszenia kondycji zdrowotnej populacji. Estymowana wartość parametru popytowego usług zdrowotnych posiada jednak znak przeciwny do oczekiwanego: możemy zatem wysnuć wniosek, iż kontraktowanie NFZ w pewnym stopniu charakteryzuje się nieefektywnością alokacyjną.

Kalkulacja alternatywnego progu efektywności kosztowej opartej na wyniku analizy ekonometrycznej

Jak wspomnieliśmy wcześniej, model (5) został oszacowany w postaci logarymicznej, co pozwala na interpretację parametru wydatków zdrowotnych jako elastyczności. Swoisty "koszt życia", rozumiany jako wartość monetarną związaną z ratowaniem zdrowia i związany ze specyficzną jednostką chorobową (grupą chorobową) możemy zdefiniować jako:

$$\text{koszt życia} = \frac{\Delta \text{wydatki zdrowotne}}{\Delta \text{umieralność}} = \frac{\text{wydatki zdrowotne} * \text{elastyczność wydatków}}{\Delta \text{umieralność}} \quad (6)$$

i zaproponować jako alternatywną (i realistyczną) metodę kalkulacji progu efektywności kosztowej. Jako ilustrację przedstawiamy szacowany "koszt życia" dla grupy chorób niedokrwiennych serca.

Całkowite wydatki na leczenie chorób niedokrwiennych serca poniesione przez oddziały wojewódzkie NFZ w 2010 r. opiewały na sumę 2 199 782 583 zł. Jednocześnie ilość zgonów z powodu chorób niedokrwiennych serca w 2011 r. zmalała o 4 131 osób w stosunku do roku poprzedniego. Podstawiając powyższe wartości do wzoru (6) oraz wykorzystując estymowaną ekonometrycznie elastyczność wydatków, swoista "cena życia" ww. grupy chorób to 218 327,49 zł. Wartość ta niemal dwukrotnie przewyższa ustawową wartość progu efektywności kosztowej. Jeśli więc przyjmiemy, iż wzrost wydatków na ochronę zdrowia jest uzasadniony poprawą efektu zdrowotnego, alternatywna (bądź innowacyjna) technologia medyczna w przypadku chorób niedokrwiennych serca powinna być oceniana względem progu efektywności, którego wartość przytoczyliśmy powyżej.

DYSKUSJA ORAZ WNIOSKI KOŃCOWE

Analizując wyniki badania należy podkreślić, iż zaproponowana metoda kalkulacji progu kosztowego w procesie oceny technologii medycznych może być podstawą decyzji rekomendacyjnych jedynie w sytuacji dostępności i wiarygodności danych epidemiologicznych wykorzystywanych do budowy wskaźnika efektu zdrowotnego. Wybór owego wskaźnika jest także niezwykle istotnym elementem wiarygodnej analizy farmakoekonomicznej. Jego trafny wybór powinien być determinowany partykularną charakterystyką jednostki chorobowej. Dla przykładu, literatura oraz zalecana przez Światową Organizację Zdrowia metodologia obejmuje kilka wskaźników efektu zdrowotnego, wśród których wymienić przede wszystkim wskaźnik YLL (utrata lat życia spowodowana chorobą)⁷, YLD (lata przeżyte w niepełnosprawności)⁸ oraz wspomniany wcześniej wskaźnik QALY. W przypadku chorób obarczonych wysokim ryzykiem zgonu (schorzenia onkologiczne, hematologiczne) obowiązujące rozwiązania definiujące próg efektywności kosztowej pomijają najważniejsze efekty terapii stosowanych w chorobach śmiertelnych, ponieważ w przypadku tej grupy schorzeń najistotniejszym parametrem w procesie badań klinicznych jest całkowite przeżycie. Oczywiście jest więc, iż stosowanie wskaźnika QALY (uwzględniającego jakość życia,) nie jest właściwe. Zdaniem onkologów, ewaluacja terapii w oparciu

⁷ Ang. *Years of Life Lost*, wyliczana jako iloczyn liczby zgonów i przeciętnego dalszego trwania życia dla wieku w którym nastąpił zgon.

⁸ Ang. *Years Lived with Disability, Years Lost due to Disability*, wyliczana jako iloczyn liczby przypadków choroby i średniego czasu jej trwania oraz wagi przypisanej chorobie w zależności od ciężkości schorzenia (0-zupełne zdrowie, 1-zgon).

o wskaźniki jakości życia może stanowić najważniejsze kryterium oceny jedynie w przypadku stosowania metod leczenia chorych „w bardzo zaawansowanych stadiach choroby i z przewidywanym krótkim czasem przeżycia”⁹. Ponadto, zastrzeżenia budzi sposób wypełniania ankiet przez chorych, będących często pod wpływem narkotycznych środków przeciwbólowych, a więc nieświadomych swojego rzeczywistego stanu. Kontrowersje budzi także metodologiczna niespójność wynikająca z braku jednolitych ankiet jakości życia QALY. Reasumując, w przypadku chorób onkologicznych ocena ekonomiczna terapii powinna uwzględniać ocenę przeżycia całkowitego. Użycie wskaźnika QALY jest zatem uzasadnione w przypadku schorzeń nieobarczonych wysokim ryzykiem zgonu, w przypadku których poprawa jakości pacjentów jest zasadniczym zadaniem.

Sztywny próg opłacalności wydaje się także całkowicie nie do zaakceptowania w przypadku chorób rzadkich i ultraradkich. Z uwagi na niewielką liczbę chorych na daną rzadką jednostkę chorobową, która stanowi wskazanie do ich stosowania, leki sieroce¹⁰ używane w terapii tych schorzeń nie stanowią one interesującego rynku inwestycyjnego dla firm farmaceutycznych. Nawet jeśli odpowiedni lek sierocy jest opracowany i dopuszczony do obrotu, zazwyczaj koszt związanej z ich użyciem terapii jest niewspółmiernie wysoki w stosunku do skuteczności zdrowotnej. Zdarza się także, iż stosowanie niektórych leków wiąże się z wystąpieniem poważnych skutków ubocznych, obniżając jakość życia pacjentów cierpiących na rzadkie schorzenie. Niska efektywność kosztowa leków sierocych (bądź jej brak) jest źródłem kontrowersji w przypadku publicznego finansowania ww. terapii. W przypadku chorób rzadkich stosowany obecnie próg opłacalności jest wysoce dyskryminujący i społecznie nie do zaakceptowania. Należy także zauważyć, iż obecne rozwiązania pomijają wysokie koszty społeczne związane z leczeniem rzadkich schorzeń, związane z zaangażowaniem rodzin pacjentów lub osób trzecich w opiekę nad chorym.

BIBLIOGRAFIA

- Claxton K., Martin S., Soares M., Rice N., Spackman E., Hinde S., Devlin N., Smith P.C., Sculpher M. (2013) Methods for the estimation of the NICE Cost Effectiveness Threshold, Centre for Health Economics, University of York, Working paper no.081.
- Cremieux, P.Y., Ouellette P., Pilon C. (1999), Health care spending as determinants of health outcomes. *Health Economics*, 8(7), str. 627-639.
- Gravelle H.S.E., Backhouse M.E. (1987) International cross-section analysis of the determination of mortality, *Social Science & Medicine*, 25(5), str. 427-441.

⁹ Jakubiak L. (2011), Jak wiarygodnie ocenić efektywność kosztową leku, <http://www.rynekapteku.pl/farmakologia/jak-wiarygodnie-ocenic-efektywnosc-kosztowa-leku,4909.html>, dostęp: 10.06.2014

¹⁰ Z ang. *orphan drugs*: leki stosowane w leczeniu chorób rzadko występujących.

- Jakubiak L. (2011), Jak wiarygodnie ocenić efektywność kosztową leku, <http://www.rynekaptek.pl/farmakologia/jak-wiarygodnie-ocenic-efektywnosc-kosztowa-leku,4909.html>, dostęp: 10.06.2014
- Nixon J., Ulmann P. (2006), The relationship between health care expenditure and health outcomes. Evidence and caveats for a causal link, *The European journal of health economics: HEPAC : health economics in prevention and care*, 7(1), str. 7-18.
- Narodowy Fundusz Zdrowia (2010) Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2009 rok, NFZ Warszawa.
- Narodowy Fundusz Zdrowia (2011) Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2010 rok, NFZ Warszawa.
- Narodowy Fundusz Zdrowia (2012) Sprawozdanie z działalności Narodowego Funduszu Zdrowia za 2011 rok, NFZ Warszawa.
- Światowa Organizacja Zdrowia (2013) 'The top ten causes of death', Fact sheet No. 310, World Health Organization, <http://who.int/mediacentre/factsheets/fs310/en/>, dostęp: 02.06.2014
- Ustawa z dnia 12 maja 2011 r. o refundacji leków, środków spożywczych specjalnego przeznaczenia żywieniowego oraz wyrobów medycznych, *Dziennik Ustaw.* z dnia 13 czerwca 2011 r.

**ESTIMATING THE COST-EFFECTIVENESS THRESHOLD
FOR THE ISCHEMIC HEART DISEASES:
AN ALTERNATIVE CALCULATION METHOD DESIGNED
FOR HEALTH TECHNOLOGY ASSESSMENT PURPOSES**

Abstract: This article presents an alternative method for calculating the cost-effectiveness threshold in the process of health technology assessment (HTA). The method is based on a theoretical model of health function and builds upon the econometric analysis addressing the impact of marginal change in overall NHS expenditure on changes in mortality for the ischemic heart diseases. Results of the analysis indicate that currently used in Poland rigid threshold performance does not reflect the actual costs associated with the improvement of health condition.

Keywords: health economics, cost-effectiveness analysis, cost of disease

EFEKTYWNOŚĆ INWESTYCYJNA FUNDUSZY EMERYTALNYCH NA TLE FIO STABILNEGO WZROSTU¹

Andrzej Karpio, Dorota Żebrowska – Suchodolska
Katedra Ekonometrii i Statystyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: andrzej_karpio@sggw.pl, zdorota@o2.pl

Streszczenie: Ze względu na zmiany ustawowe, płacący ubezpieczenie emerytalne mają prawo wyboru wpłat do ZUS lub do OFE, decyzje można podejmować co dwa lata. Ponadto, fundusze emerytalne straciły prawo inwestowania w obligacje skarbowe, w konsekwencji składy portfeli będą uzupełni inne niż dotychczas. W szczególności będą bardziej ryzykowne i w swojej strukturze będą podobne do portfeli funduszy zrównoważonych. Przez cały okres funkcjonowania reformy emerytalnej, inwestycje funduszy emerytalnych przypominały inwestycje otwartych funduszy stabilnego wzrostu. W pracy dokonano analizy porównawczej efektywności inwestycyjnej OFE i otwartych funduszy stabilnego wzrostu na przestrzeni ostatnich lat. Narzędziem porównawczym jest wskaźnik zysków i strat oraz wskaźniki będące modyfikacją wskaźnika Sharpe'a. Jako czynnik rynkowy zaproponowano wskaźnik będący kombinacją liniową indeksu WIG oraz indeksu rynku obligacji TBSP.

Słowa kluczowe: otwarte fundusze inwestycyjne, otwarte fundusze emerytalne, efektywność inwestycyjna, portfele inwestycyjne, wskaźniki efektywności

WSTĘP

Problem zabezpieczenia emerytalnego jest niewątpliwie jednym z podstawowych przed jakimi stają obywatele wszystkich państw. Zmiana ustrojowa, jaka dokonała się w Polsce na początku lat 90-tych ubiegłego wieku, wymusiła reformę tego systemu. Składka pobierana z naszych wynagrodzeń

¹ Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/09/B/HS4/00493

dzielona była pomiędzy dwa filary, pierwszy oparty na obietnicach (ZUS) i drugi oparty na inwestycjach kapitałowych (OFE). Przedmiotem rozważań prowadzonych w niniejszej pracy jest drugi filar ściśle powiązany z rynkiem kapitałowym. Jego efektywność zależy nie tylko od umiejętności zarządzających, ale również od koniunktury gospodarczej, stabilności systemu prawnego regulującego zasady jego działania, czy wreszcie niedotrzymanej obietnicy likwidacji uprzywilejowanych grup pracowniczych, na których emerytury „łożą” wszyscy płacący składki. Już pobieżny rzut oka na wymienione czynniki nasuwa wniosek, że wiele z nich jest dalekie od deklarowanych zasad. Jednak nie ten problem będzie poruszany w tej pracy, ale efektywność funkcjonowania OFE na tle rynku funduszy inwestycyjnych stabilnego wzrostu. Ponieważ ustawowe ograniczenia składu portfeli funduszy emerytalnych przypominają portfele otwartych funduszy inwestycyjnych stabilnego wzrostu, zatem wyniki inwestycyjne tych drugich będą służyły jako swego rodzaju benchmark. Podstawą porównania są: miary będące modyfikacją wskaźnika Sharpe’a oraz wskaźnik zysków i strat, oparte na innej „filozofii” niż standardowo wykorzystywane w rodzaju wskaźników: Treynora, Jensena i Sharpe’a. W badaniach bierze się pod uwagę miesięczne procentowe zmiany jednostek uczestnictwa, nie uwzględnia się pobieranej prowizji i opłat za zarządzanie. Realizowanym celem pracy jest pokazanie, poprzez miary efektywności, że rynki funduszy stabilnego wzrostu i funduszy emerytalnych są do siebie podobne, zarówno z punktu widzenia wartości wskaźników, jak i stabilności rankingów oraz braku długoterminowych liderów.

ZAŁOŻENIA METODOLOGICZNE

Okres badań obejmuje lata 2000 – 2013 i został podzielony na dwa podokresy: pierwszy trwa od 31 stycznia 2000 roku do 31 grudnia 2004 roku, a drugi od 31 stycznia 2005 roku do 31 grudnia 2013 roku. Należy dodać, że fundusze emerytalne rozpoczęły działalność w kwietniu 1999 roku i w początkowym okresie zajmowały się przede wszystkim akwizycją. Dlatego jako datę początkową przyjęto 31 stycznia 2000 roku, gdy ich aktywa znacznie wzrosły umożliwiając inwestycje kapitałowe w pełnym tego słowa znaczeniu. Drugi okres, rozdzielony nieco umowną datą, obejmuje lata dobrej koniunktury giełdowej oraz okres kryzysu. Ma to na celu uwzględnienie w badaniach różnej sytuacji gospodarczej, która jest immanentną cechą gospodarki rynkowej. W badanych okresach na rynku funduszy emerytalnych zachodziło wiele zmian. Pomijając aspekty prawne (zmiany: benchmarku, wielkości przekazywanej składki itp.), z punktu widzenia niniejszej pracy najbardziej istotna była zmieniająca się liczba funduszy. Ponieważ dochodziło do przejęć i połączeń (nie było upadłości, w zasadzie niemożliwej z punktu widzenia obowiązującego prawa). w pracy użyto nazwy funduszu „dominującego”, pod szyldem którego dalej fundusz funkcjonował. W konsekwencji badania dotyczyły 14 funduszy: AEGON OFE (AEGON), Allianz Polska OFE (Allianz), Amplico OFE (Amplico), Aviva OFE

(Aviva), AXA OFE (AXA), Generali OFE (Generali), ING OFE (ING), Nordea OFE (Nordea), PKO BP Bankowy OFE (Bankowy), OFE Pocztylion (Pocztylion), Pekao OFE (Pekao), OFE Polsat (Polsat), PZU „Złota Jesień” (PZU), OFE WARTA (Warta), w nawiasach podano skrócone nazwy wykorzystywane w dalszej części pracy.

Możliwości porównawcze stwarzają jedynie te fundusze stabilnego wzrostu, które funkcjonowały w całym okresie badań. Okazuje się, ich liczba zmieniała się dosyć znacznie. Spośród nich, w pierwszym okresie istniało na rynku jedynie pięć, a w całym drugim okresie (i pierwszym) aż osiemnaście. Dlatego szczegółowe wyniki będą podane jedynie dla okresu drugiego, a okres pierwszy zostanie potraktowany skrótowo. W konsekwencji, fundusze stabilnego wzrostu poddane badaniu to: Allianz FIO Subfundusz Allianz Stabilnego Wzrostu (Allianz), Amplico FIO Parasol Krajowy Amplico Subfundusz Stabilnego Wzrostu (Amplico), Arka BZ WBK Stabilnego Wzrostu FIO (Arka), Aviva Investors FIO subfundusz Aviva Investors Ochrony Kapitału Plus (Aviva Ochrony Kapitału), Aviva Investors FIO subfundusz Aviva Investors Stabilnego Inwestowania (Aviva Stabilnego Inwestowania), BPH FIO Parasolowy Subfundusz BPH Stabilnego Wzrostu (BPH), DWS Polska FIO Zabezpieczenia Emerytalnego (DWS), Fortis FIO subfundusz FORTIS stabilnego wzrostu (Fortis), ING Parasol FIO subfundusz Stabilnego Wzrostu (ING), KBC Parasol FIO Subfundusz Stabilny (KBC), Legg Mason Senior SFIO (Legg Mason), Lukas FIO Subfundusz Lukas Stabilnego Wzrostu (Lukas), Millennium FIO Subfundusz Stabilnego Wzrostu (Millennium), Novo FIO Subfundusz Novo Stabilnego Wzrostu (Novo), Pioneer FIO Subfundusz Pioneer Stabilnego Wzrostu (Pioneer), PKO Stabilnego Wzrostu (PKO), PZU FIO Stabilnego Wzrostu MAZUREK (PZU), Skarbiec FIO Subfundusz Skarbiec III Filar (Skarbiec). Podobnie jak w przypadku funduszy emerytalnych w nawiasach podano nazwy skrócone.

Podstawą oceny efektywności inwestycyjnej funduszy są wskaźniki zysków i strat (wskaźniki Calmara, Omega, potencjału nadwyżkowej stopy zwrotu UPR, Sortino) oraz wskaźnik będący modyfikacją wskaźnika Sharpe’a (Sharpe’a-Israelseña). Ich szczegółowe omówienie wraz z interpretacją można znaleźć w wielu opracowaniach, dlatego w tym miejscu ograniczymy się do tabeli zawierającej podstawowe informacje o wykorzystanych wskaźnikach. Oprócz nazwy i przyjętej definicji Tabela 1 zawiera również odnośnik do literatury, z której korzystano w niniejszej pracy.

Tabela 1. Wskaźniki zysków i strat

Nazwa wskaźnika	Definicja	Źródło
Calmara	$\frac{R}{ MDD_T }$	[Pedersen, Rudholm-Alfvín 2003]
Omega	$\frac{\sum_t \max(R_t - m)}{\sum_t \max(m - R_t)}$	[Shadwick, Keating, 2002]
Wskaźnik potencjału nadwyżkowej stopy zwrotu UPR	$\frac{\frac{1}{T-1} \sum_{t:R_t > m} (R_t - m)}{\sqrt{\frac{1}{T-1} \sum_{t:R_t < m} (R_t - m)^2}}$	[Sortino i inni, 1999]
Sortino	$\frac{\bar{R} - m}{\overline{\theta(m)}}$	[Sortino, Price, 1994]
Sharpe'a - Israelsena	$\frac{\bar{R} - \bar{R}_b}{TR^{\pm 1}}$	[Israelsen 2005]

Źródło: opracowanie własne

W przyjętych wzorach R jest roczną stopą zwrotu z jednostek funduszu, R_t jest zmianą w okresie t , m jest progiem rentowności czyli minimalną akceptowaną stopą zwrotu (przyjmowaną mniej lub bardziej arbitralnie), \bar{R} i \bar{R}_b są odpowiednio średnimi stopami zmian jednostek funduszu i benchmarku, pozostałe wielkości to: obsunięcie kapitału

$$MDD_T = \min_u \left\{ \frac{S_u - \max_t S_t}{\max_t S_t} \right\} \quad (1)$$

gdzie S jest wartością jednostki uczestnictwa (ceną, a nie procentową zmianą); Tracking Error

$$TR = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (R_t - R_{b,t} - (\bar{R} - \bar{R}_b))^2} \quad (2)$$

oraz miara ryzyka zadana wzorem:

$$\overline{\theta(m)} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T [(R_t - m)^-]} \quad (3)$$

gdzie $(R_t - m)^-$ przyjmuje wartość równą $R_t - m$ gdy $R_t < m$ i zero w przeciwnym przypadku.

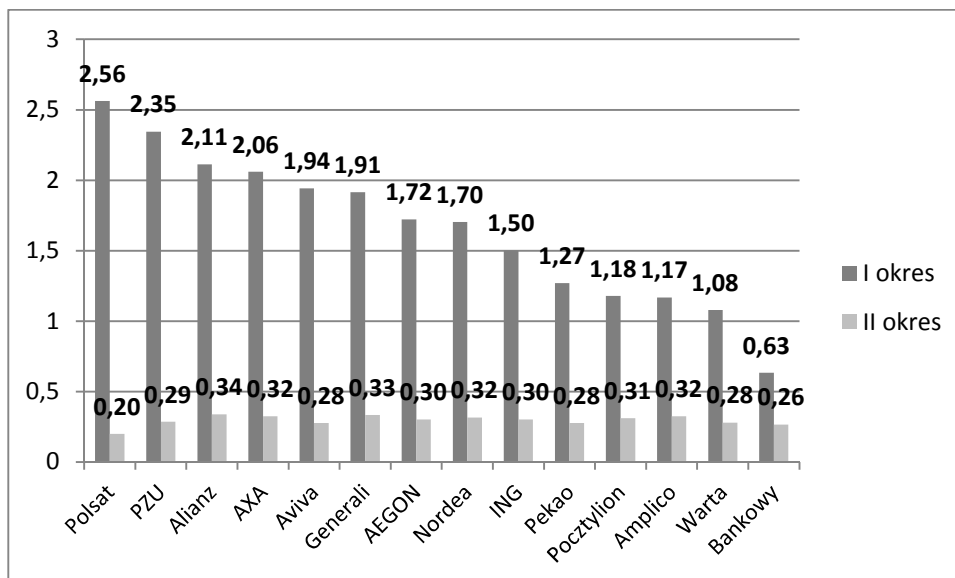
W pracy przyjęto, że próg rentowności jest równy zero ($m = 0$), natomiast benchmarkiem jest średnia ważona indeksu rynku obligacji TBSP oraz WIG z wagami odpowiednio: 0,6 i 0,4. Należy dodać, że datą bazową indeksu TBSP jest 29 grudnia 2006 roku, zatem wskaźnik Sharpe'a - Israelsena liczony był dla okresu zaczynającego się w styczniu 2007 roku, a nie w styczniu 2005 roku. Przyjęte wagi odzwierciedlają dotychczasowe (do końca 2013 r.) ograniczenia nałożone na portfele funduszy emerytalnych, w myśl których mogły one inwestować na rynku

akcji do 40% aktywów. Również w przypadku dwóch funduszy emerytalnych (Polsat i Bankowy) oraz dwóch funduszy stabilnego wzrostu (BPH, ING) szeregi miesięcznych stóp zwrotu kończyły się wcześniej niż 31 grudnia 2013.

RANKINGI FUNDUSZY EMERYTALNYCH, A RYNEK FUNDUSZY STABILNEGO WZROSTU

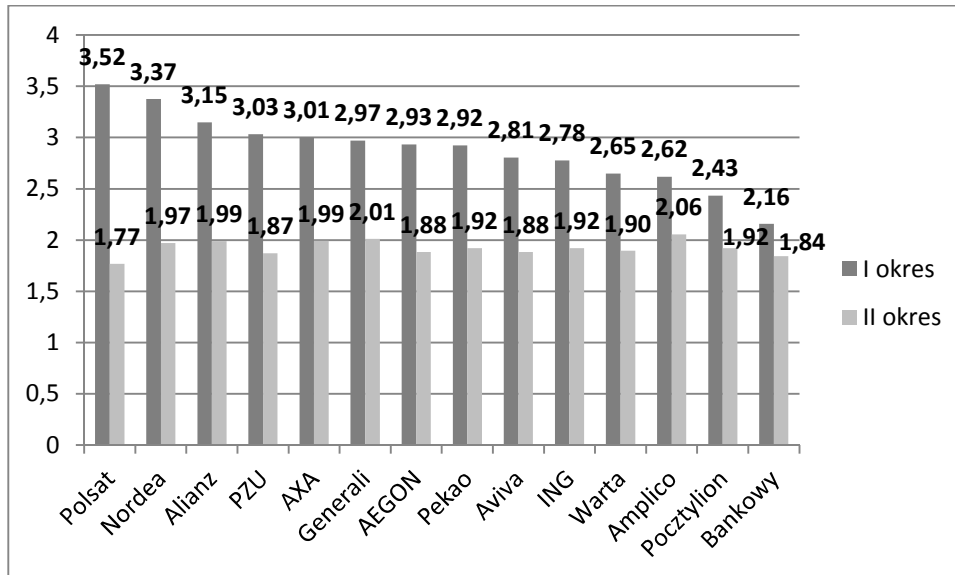
Poniższe wykresy ukazują rankingi funduszy emerytalnych tworzone na podstawie wyżej omówionych wskaźników, pozycje funduszy uporządkowano malejąco w pierwszym okresie (31 stycznia 2000 r. do 31 grudnia 2004 r.). Już teraz warto zwrócić uwagę, że we wszystkich przypadkach zróżnicowanie wartości wskaźników w pierwszym okresie jest znacznie większe niż w drugim. Niektóre fundusze pojawiają się na początkowych pozycjach (np. Polsat, Nordea, PZU), a inne na końcowych (np. Bankowy, Pocztylion, Amplico), co oznacza, że zastosowane miary nie wprowadzają istotnych różnic pomiędzy efektywnością funkcjonowania funduszy. Ponieważ definicje wskaźników korzystają z innych miar ryzyka, zatem sytuację taką można tłumaczyć lepszym i gorszym zarządzaniem portfelami inwestycyjnymi funduszy.

Rysunek 1. Ranking funduszy emerytalnych oparty na wskaźniku Calmara



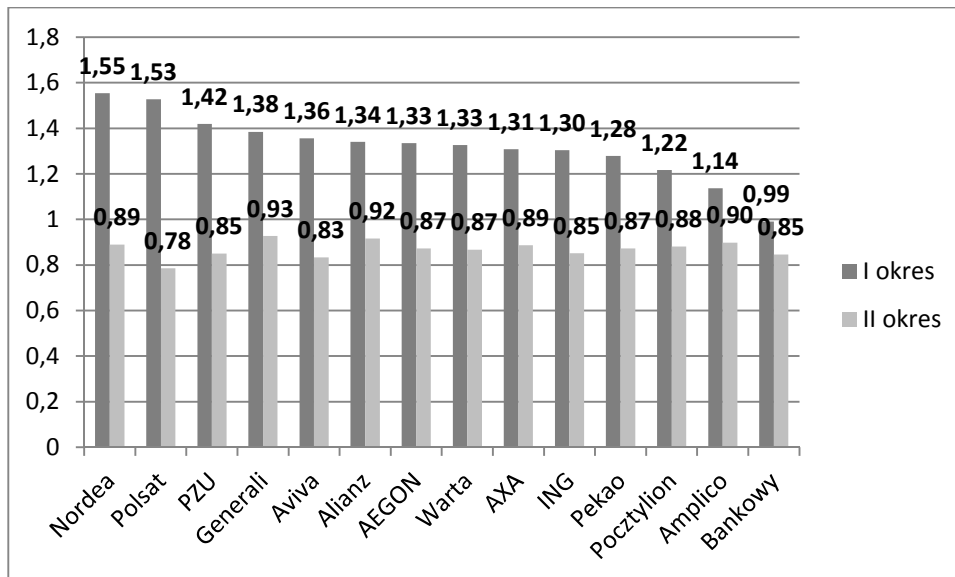
Źródło: opracowanie własne

Rysunek 2. Ranking funduszy emerytalnych oparty na wskaźniku Omega



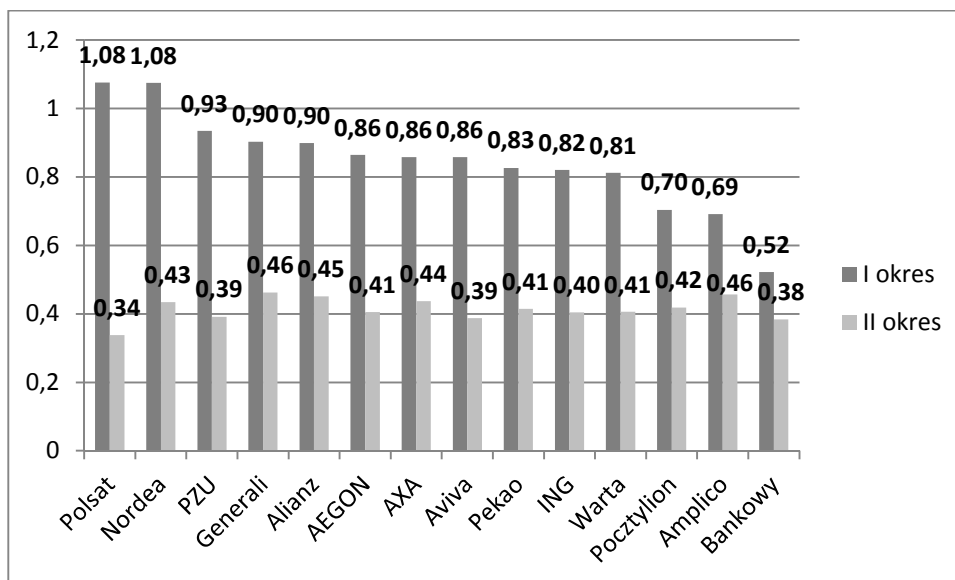
Źródło: opracowanie własne

Rysunek 3. Ranking funduszy emerytalnych oparty na wskaźniku UPR



Źródło: opracowanie własne

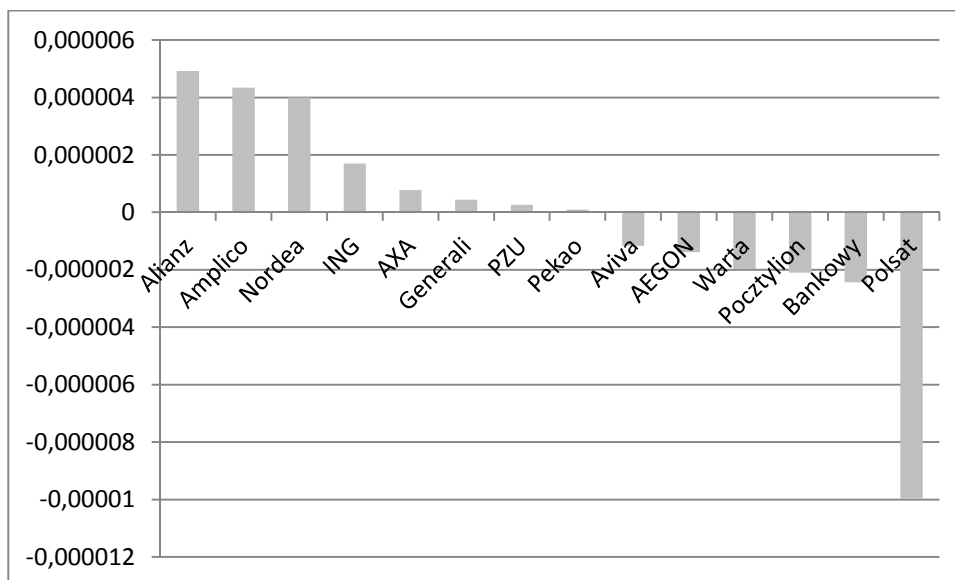
Rysunek 4. Ranking funduszy emerytalnych oparty na wskaźniku Sortino



Źródło: opracowanie własne

Sytuacja zmienia się radykalnie, gdy skorzystamy ze wskaźnika Sharpe'a-Isrealsena, który w istotny sposób zależy od benchmarku. W niniejszej pracy jest nim kombinacja liniowa indeksu WIG i indeksu obligacji TBSP odzwierciedlająca ustawowe ograniczenia nałożone na portfele funduszy emerytalnych, wynik przedstawia Rysunek 5. Lider poprzednich rankingów, fundusz Polsat, tym razem plasuje się na ostatniej pozycji, natomiast Nordea nadal pozostaje w czołówce, a fundusz PZU znalazł się w środku rankingów. Fundusze Bankowy i Pocztylion nadal wypadają bardzo słabo. Wydaje się, że istotne różnice metodologiczne wyżej wykorzystanych wskaźników oraz wskaźnika Sharpe'a-Isrealsena mogą być podstawą wyboru funduszy najlepszych i najgorszych. Ten drugi wskaźnik w istotny sposób wykorzystuje benchmark, który jest niezależny od wyników osiągniętych przez same fundusze, ale opisuje sytuację rynkową, zatem związany jest z ryzykiem rynkowym. Jednak należy zwrócić uwagę na wartości wskaźnika Sharpe'a-Isrealsena, są one bardzo małe i różnice widoczne na wykresie wynikają jedynie z przyjętej skali. Dlatego wyżej podane wnioski mają raczej charakter teoretyczny i są mało przydatne z praktycznego punktu widzenia.

Rysunek 5. Ranking funduszy emerytalnych oparty na wskaźniku Sharpe'a-Isrealsena



Źródło: opracowanie własne

W przypadku funduszy stabilnego wzrostu szczupłość miejsca nie pozwala na pełną prezentację wyników, dlatego ograniczymy się do ich pobieżnego omówienia. Jak wspomniano wcześniej, w pierwszym okresie badań funkcjonowało jedynie pięć podmiotów, które „dotrwały” do końca 2013 roku, dlatego wnioski dotyczą jedynie lat 2005-2013. Wskaźniki Omega i UPR obliczone dla funduszy stabilnego wzrostu i funduszy emerytalnych przyjmują wartości bardzo do siebie zbliżone. Inaczej jest w przypadku wskaźników Calmara i Sortino, wówczas najgorsze fundusze stabilnego wzrostu mają wartości wskaźników o rząd wielkości mniejsze niż fundusze emerytalne. Dla wszystkich funduszy emerytalnych i stabilnego wzrostu najlepsze fundusze mają wartości wszystkich wskaźników bardzo podobne. W konsekwencji, można stwierdzić, że z punktu widzenia współczynników Calmara i Sortino fundusze stabilnego wzrostu charakteryzują się znacznie większą rozpiętością osiągniętych wyników inwestycyjnych niż fundusze emerytalne. Natomiast przy wykorzystaniu wskaźników Omega i UPR rynki obu rodzajów funduszy są do siebie podobne. Jeśli wziąć pod uwagę wskaźnik Sharpe'a-Isrealsena, to jego wartości dla wszystkich funduszy stabilnego wzrostu są ujemne, ale ich moduły przewyższają o rząd wielkości wartości wskaźnika dla funduszy emerytalnych. Są zatem zbliżone do wartości charakteryzującej fundusz emerytalny Polsat.

Kolejne pytanie jakie się nasuwa dotyczy różnic w pozycjach rankingowych wynikających z zastosowanych miar. W pracy wykorzystano współczynnik korelacji rangowej Spearmana [Luszniewicz, Słaby 2003] i obliczono go dla par

rankingów opartych na czterech wskaźnikach, istotność badano na poziomie 0,05. Wyniki dla funduszy emerytalnych dotyczące lat 2005-2013 zebrano w Tabeli 2, a dla funduszy stabilnego wzrostu w Tabeli 3.

Tabela 2. Wartości współczynnika korelacji rangowej Spearmana dla funduszy emerytalnych

	Calmara	Omega	UPR	Sortino
Calmara	1	0,850549	0,947253	0,916484
Omega		1	0,863736	0,925275
UPR			1	0,973626
Sortino				1

Źródło: opracowanie własne

Tabela 3. Wartości współczynnika korelacji rangowej Spearmana dla funduszy stabilnego wzrostu

	Calmara	Omega	UPR	Sortino
Calmara	1	0,958720	0,911249	0,981424
Omega		1	0,907121	0,960784
UPR			1	0,915377
Sortino				1

Źródło: opracowanie własne

W podsumowaniu można stwierdzić, że korelacje pomiędzy rankingami stworzonymi na podstawie wskaźników Calmara, Omega, UPR i Sortino świadczą o tym, że dostarczają one jednakowych informacji o efektywności funkcjonowania funduszy. Wniosek dotyczy zarówno funduszy emerytalnych, jak i stabilnego wzrostu. Zatem jest w zasadzie bez znaczenia, który wskaźnik zastosujemy chcąc wybrać fundusz najlepszy, niewielkie różnice w pozycjach rankingowych mają raczej charakter przypadkowy. Co więcej, poza dużymi podobieństwami między rynkami funduszy emerytalnych i stabilnego wzrostu wynikającymi z omówionych wyżej wskaźników pojawia się ich dodatkowa właściwość. Jest nią niewielka czułość na wybór metody oceny efektywności. W konsekwencji współczynnik korelacji rangowej Spearmana dodatkowo potwierdza wcześniej sformułowany wniosek o podobieństwie pomiędzy rynkami obu rodzajów funduszy.

PODSUMOWANIE

Prezentowana praca stanowi wstęp do badania rynku funduszy emerytalnych po ostatnich zmianach ustawowych. Trudno jest w tym miejscu przewidzieć jak wkrótce będzie wyglądał ten rynek. Zmieniły się zasady funkcjonowania funduszy emerytalnych na tyle radykalnie, że można oczekiwać

bardziej ryzykownych inwestycji (zakaz nabywania obligacji skarbowych), a co za tym idzie bardziej zróżnicowanych wyników inwestycyjnych. Wprawdzie ponad 2,5 miliona osób wybrało fundusze emerytalne zamiast ZUS, co kompletnie wszystkich zaskoczyło, to bez wątpienia rynek ulegnie konsolidacji. Obecna liczba 14 funduszy emerytalnych jest zbyt duża by „obsługiwać” tych wszystkich, którzy zdecydowali się na kapitałowy II filar. Natomiast same fundusze, aby funkcjonować, muszą zachęcić kolejne osoby do wyboru tego wariantu ubezpieczenia emerytalnego, a jest tylko jedna metoda – dobre wyniki inwestycyjne. Zatem dotychczasowe badania rynku funduszy emerytalnych nabierają charakteru historycznego, który prawdopodobnie będzie miał niewielkie przełożenie na strukturę i efektywność przyszłego rynku tych podmiotów. Na zakończenie warto zwrócić uwagę, że z prezentowanej pracy wynika ważny wniosek związany z rynkiem funduszy stabilnego wzrostu. W badanym okresie obejmującym lata 2005-2013 rynek tych funduszy bardzo przypominał rynek funduszy emerytalnych. W szczególności, wartości zastosowanych wskaźników były zbliżone do siebie, jak i pozycje rankingowe na nich oparte były „stabilne”. Jednak dotychczasowa struktura portfeli inwestycyjnych obu rodzajów funduszy ulegnie radykalnej zmianie. Ze względu na zmiany ustawowe, fundusze emerytalne zaczną raczej przypominać fundusze zrównoważone. Zatem badania prowadzone w przyszłości za bazę porównawczą będą musiały brać właśnie te podmioty, bardziej ryzykowne niż fundusze stabilnego wzrostu. Obserwując rynek funduszy zrównoważonych można stwierdzić, że porównywany z nim rynek funduszy emerytalnych stanie się rynkiem znacznie różniącym funkcjonujące na nim podmioty.

BIBLIOGRAFIA

- Israelsen C. L. (2005) A Refinement to the Sharpe Ratio and Information Ratio, *Journal of Asset Management*, vol. 5(6).
- Domański Cz. (red.) (2011) *Nieklasyczne metody oceny efektywności i ryzyka*, PWE, Warszawa.
- Luszniewicz A., Słaby T. (2003) *Statystyka z pakietem komputerowym STATISTICA PL*, Wydawnictwo C.H. BECK, Warszawa.
- Pedersen C. S., Rudholm-Alfvén T. (2003) Selecting a Risk-Adjusted Shareholder Performance Measure, *Journal of Asset Management*, vol. 4(3), str. 152-172.
- Shadwick W., Keating C. (2002) A universal performance measure, *Journal of Performance*, str. 59-84.
- Sortino F., Van der Meer R., Plantinga A. (1999) The Upside Potential Ratio, *Journal of Performance Measurement* 4(1), str. 10–15.
- Sortino F., Price L. (1994) Performance Measurement in a Downside Risk Framework, *Journal of Investing* 3(3), str. 59–64.

**INVESTMENT EFFECTIVENESS
OF THE MUTUAL FUNDS IN COMPARISON WITH
THE STABLE GROWTH OPEN-END MUTUAL FUNDS**

Abstract: Due to the legal changes, people paying pension insurance have the right to choose between ZUS and OFE. The decision may be made every two years. In addition, open pension funds have lost the right to invest in treasury bonds, consequently investment portfolios will be totally different. They will be particularly more risky and they will be similar to stable growth mutual funds portfolios in terms of structure. During the period of the functioning of the pension reform, open pension funds investments resembled open-end stable growth mutual funds investments. In this work, a comparative analysis of the investment effectiveness of OFE and open-end stable growth funds was carried out over the last couple of years. The comparative tool, used in the analysis, are the measures of the gains and losses, as well as the modification of the Sharpe's indicator. As a market factor, an investment indicator which is a linear combination of WIG index and the index of TBSP was proposed.

Keywords: open-end mutual funds, open pension funds, investment effectiveness, investment portfolios, effectiveness indicators

ESTYMACJA MIARY MARTYNGAŁOWEJ NA PODSTAWIE CEN OPCJI Z GIEŁDY PAPIERÓW WARTOŚCIOWYCH W WARSZAWIE

Paweł Kliber

Katedra Ekonomii Matematycznej, Uniwersytet Ekonomiczny w Poznaniu
e-mail: p.kliber@ue.poznan.pl

Streszczenie: W artykule prezentujemy zastosowanie szacowania miary martyngałowej dla indeksu WIG20 z Giełdy Papierów Wartościowych w Warszawie. Miarę martyngałową szacujemy na podstawie cen opcji na ten indeks. Przyjmujemy, że miara martyngałowa jest mieszaniną rozkładów logarytmiczno-normalnych, a parametry rozkładu szacujemy minimalizując sumę kwadratów błędów wyceny. Otrzymane wyniki porównujemy z modelem zakładającym rozkład logarytmiczno-normalny. Jak przykład rozważamy zmiany miary martyngałowej na początku marca 2014 r., po rozpoczęciu kryzysu na Krymie.

Słowa kluczowe: wycena martyngałowa, rozkład prawdopodobieństwa implikowany przez ceny opcji, awersja do ryzyka, realna miara probabilistyczna, wykrywanie wydarzeń

WPROWADZENIE

Jedną z podstawowych reguł nowoczesnej teorii finansów jest zasada wyceny martyngałowej¹. Zgodnie z tą zasadą cena dowolnego instrumentu finansowego jest równa wartości oczekiwanej zdyskontowanych przepływów związanych z tym instrumentem, przy czym wartość oczekiwana jest wyznaczana nie na podstawie rzeczywistych prawdopodobieństw, ale przy przyjęciu pewnej sztucznie wprowadzonej miary probabilistycznej, nazywanej miarą martyngałową. Miara ta to rozkład prawdopodobieństwa, który przyjmowałby hipotetyczny

¹ Patrz np. Musiela, M., Rutkowski, M. (2008) *Martingale Methods in Financial Modelling*, Springer lub Pliska, S.R. (2005) *Wprowadzenie do matematyki finansowej. Modele z czasem dyskretnym*, WNT, Warszawa.

inwestor pozbawiony awersji do ryzyka, przy założeniu, że jego wycena instrumentów finansowych byłaby zgodna z obserwowaną wyceną rynkową – dlatego określa się ją także jako „miarę neutralną względem ryzyka” (ang. *risk-neutral measure*). Na rynku mamy zatem do czynienia z dwoma różnymi miarami probabilistycznymi: miarą rzeczywistą P , opisującą rzeczywiste prawdopodobieństwa różnych poziomów cen oraz miarą martyngałową Q , zawierającą informacje na temat wyceny rynkowej. Ta druga miara zawiera informacje o oczekiwaniach rynków, co do przyszłego rozwoju sytuacji, oraz o awersji inwestorów wobec ryzyka. Gęstość miary martyngałowej względem miary rzeczywistej nazywana jest rynkową ceną ryzyka² i określa niechęć inwestorów do ryzyka w możliwych przyszłych stanach świata.

W pracy [Breedon, Litzenberg 1978] zauważono, że miarę martyngałową dla dowolnego, przyszłego momentu T można wyznaczyć na podstawie opcji europejskich, których terminem wykonania jest T , jeżeli istnieje odpowiednio dużo opcji o różnej cenie wykonania. W teoretycznym przypadku, jeżeli istnieje nieskończenie wiele opcji (kupna lub sprzedaży) z dowolną ceną wykonania, gęstość miary martyngałowej można otrzymać jako drugą pochodną funkcji ceny opcji w zależności od ceny wykonania. W praktyce opcji jest jednak skończenie wiele, skąd pojawiła się konieczność opracowania innych metod wyznaczania miary martyngałowej, których przegląd można znaleźć w opracowaniu [Bahra 1997].

Oszacowania miary martyngałowej mają dwa podstawowe zastosowania. Pierwszym jest wykrywanie możliwości arbitrażu, na co zwrócono uwagę już w [Breedon, Litzenberg 1978]. Jeśli przy aktualnych cenach rynkowych nie można znaleźć gęstości miary martyngałowej – prawdopodobieństwa pewnych stanów świata są ujemne – oznacza to, że opcje o różnych cenach wykonania są niepoprawnie wycenione. Możliwe jest wówczas stworzenie strategii opcyjnej (portfela złożonego z trzech opcji kupna, taka strategia nazywana jest *spreadem motyla*³) pozwalającej na wykorzystanie możliwości arbitrażu.

Oprócz tego oszacowania takie stosowane są w analizie zdarzeń rynkowych i wykrywaniu oczekiwań rynków. W [Bahra 1997] przedstawiono kilka propozycji wykorzystania tego podejścia przez władze monetarne (m.in. do oceny oczekiwań skuteczności polityki pieniężnej i oczekiwań inflacyjnych). W artykule [Jackwerth, Rubinstein 1996] przedstawiono pewną metodę szacowania miary martyngałowej na podstawie cen opcji na indeks S&P 500. Jak stwierdzono, po kryzysie giełdowym z października 1987 roku martyngałowe prawdopodobieństwa dużych spadków wartości indeksu są wielokrotnie większe niż odpowiednie prawdopodobieństwa w mierze rzeczywistej, wyznaczone przy założeniu normalności stóp zwrotu, co dowodzi zmiany oczekiwań. Kryzys ten badano także

² Patrz np. Bjork, T. (2009) *Arbitrage Theory in Continuous Time*, Oxford University Press.

³ Patrz np. Hull, J. (2009) *Options, Futures and Other Derivatives*, Prentice Hall, s. 225.

w [Bates 1991]. W artykule [Melick, Thomas 1997] zastosowano analizowano zmiany oczekiwania na rynku ropy naftowej po wojnie w Zatoce Perskiej w roku 1990. Autorzy wykazali, że po wojnie rozkład martyngałowy zmienił kształt i pojawił się w nim gruby prawy ogon. Problem analizy rynku ropy naftowej był następnie podejmowany w pracach [Sadorsky 2001] i [Gagnon, Power 2013]. W pracy [Mandler 2002] porównano kształt rozkładu w tygodniach, w których odbywały się zebrania Rady Prezesów Europejskiego Banku Centralnego z rozkładem martyngałowym w tygodniach bez zebrań i wykazano istotne różnice między tymi rozkładami. Zmiany miary martyngałowej, wyznaczonej na podstawie opcji na kontrakty futures LIFFE-Euribor charakteryzowały się dużą różnorodnością i nie zdołano uzyskać jednoznacznych rezultatów. Natomiast w pracy [Birru, Figlewski 2012] oszacowania rozkładów prawdopodobieństwa wykorzystano do analizy kryzysu finansowego 2008 roku. W pracy [Wang 2009] zastosowano oszacowania miary martyngałowej dla indeksu FTSE 100 przy założeniu, że proces stochastyczny opisujący wartość indeksu posiada skoki i pokazano, że zdolność predykcyjna modelu takiego modelu jest lepsza. W artykule [Ait-Sahalia, Jacod 2009] zaproponowano metodę wykrywania skoków w procesie cen akcji opartą na szacowaniu miary martyngałowej na podstawie cen opcji.

W pracy [Chabi-Yo, Garcia, Renault 2008] wykorzystali przekształcenia miar martyngałowych do próby wyjaśnienia zagadki nadmiernej premii za ryzyko⁴. W artykule [Ziegler 2007] podjęto próbę wyjaśnienia tzw. „uśmiechu zmienności” (ang. *volatility smile*), przyjmując, że może on być wywołany przez agregację miar martyngałowych inwestorów o różnych oczekiwaniach.

W pracach [Liu et al. 2007], [de Vincent-Humphereys, Noss 2012] oraz [de Vincent-Humphereys, Puigvert 2012] podjęto próbę powiązania miary martyngałowej z miarą rzeczywistą. Zwrócono uwagę, że miara martyngałowa, oszacowana na podstawie cen opcji, może stanowić uzupełnienie prognoz tworzonych na podstawie danych historycznych, jeśli uda się znaleźć odpowiednią transformację od prawdopodobieństw martyngałowych do prawdopodobieństw rzeczywistych. Jako taką transformację proponowano m.in. funkcję rozkładu beta, uogólnione rozkłady beta oraz przekształcenia oparte na funkcji użyteczności ze stałą awersją do ryzyka.

W artykule podejmujemy próbę oszacowania miar martyngałowych dla zdarzeń na Giełdzie Papierów Wartościowych w Warszawie na podstawie opcji na indeks WIG20. Zastosowanie standardowych metod na rynku polskim związane jest z poważnymi problemami, ponieważ opcje znajdujące się w obrocie na GPW są mało płynne. W związku z tym do szacowania rozkładu martyngałowego należy stosować podejście parametryczne z małą liczbą parametrów. Oszacowane

⁴ Patrz Mehra, R., Prescott, E.C. (1985) The Equity Premium: A Puzzle, *Journal of Monetary Economics*, vol. 15, ss. 145-161.

rozkłady można następnie porównywać z rozkładem martyngałowym otrzymanym z modelu Blacka-Scholesa, podobnie jak robiono to w [Melick, Thomas 1997] i [Mandler 2002], i wnioskować stąd o zmianach oczekiwań inwestorów i ich skłonności do akceptacji ryzyka w odpowiedzi na pewne wydarzenia rynkowe.

ESTYMACJA MIARY MARTYNGAŁOWEJ NA PODSTAWIE CEN OPCJI EUROPEJSKICH

Rozważmy europejską opcję kupna z terminem wykonania T i ceną wykonania K . Zgodnie z zasadą wyceny martyngałowej cena tej opcji w chwili 0 wynosi

$$c(K, T) = e^{-rT} E^Q \left[(S_T - K)^+ \right], \quad (1)$$

gdzie S_T to cena instrumentu, na który opcja jest wystawiona, Q to miara martyngałowa, a r jest stopą zwrotu wolną od ryzyka. Należy podkreślić, że Q związana jest z konkretnym momentem wykonania T - określa rozkład cen instrumentu S w tym właśnie momencie. Zakładając, że rozkład ten jest ciągły, oznaczmy jego gęstość przez $q_T(x)$. Wzór (1) można zatem zapisać w postaci

$$c(K, T) = e^{-rT} \int_K^{+\infty} (x - K) q_T(x) dx. \quad (2)$$

Podobnie, cenę opcji sprzedaży z ceną wykonania K możemy wyrazić następującym wzorem:

$$p(K, T) = e^{-rT} \int_0^K (K - x) q_T(x) dx. \quad (3)$$

Różniczkując dwukrotnie obie strony równania (2) względem K otrzymujemy

$$q_T(x) = e^{rT} \frac{\partial^2 c(K, T)}{\partial K^2} \quad (4)$$

Podobnie, z równania (3) można otrzymać następującą zależność

$$q_T(x) = e^{rT} \frac{\partial^2 p(K, T)}{\partial K^2}. \quad (5)$$

Zatem jeśli posiada się obserwacje cen opcji (kupna lub sprzedaży) przy wszystkich możliwych cenach wykonania, można wyznaczyć rozkład martyngałowy. W praktyce jest to oczywiście niemożliwe – notowanych jest jedynie kilka opcji na dany instrument z różnymi cenami wykonania.

Gęstość miary martyngałowej można jednak oszacować na podstawie obserwowanych cen opcji, przyjmując odpowiednie założenia, co do postaci tej gęstości. Przegląd stosowanych metod można znaleźć w [Aparicio, Hodges 1998] oraz [Fusai, Roncoroni 2008]. Poniżej przedstawiamy krótką charakterystykę różnych podejść.

1. Podejście nieparametryczne, oparte na regresji kernelowej, zostało zaproponowane w [Aït-Sahalia, Lo 1998]. Wymaga ono jednak odpowiednio dużej liczby obserwacji. W warunkach polskich, gdzie rynek opcyjny jest mało płynny, tej metody nie można zastosować.
2. Podejście oparte na dopasowywaniu zmienności w modelu. W pracach [Rubinstein 1994], [Jackwerth, Rubinstein 1996], [Jackwerth 1999] zastosowano w tym celu implikowane drzewa dwumianowe. Natomiast w artykułach [Durpie 1994] i [Derma, Kani 1994] wykorzystano „uśmiech zmienności”. W obu tych podejściach przyjmuje się pewne założenia co do przyszłych zmienności instrumentu podstawowego, aby otrzymać w modelu obserwowane na rynku ceny opcji.
3. Podejście oparte na przyjęciu modelu dyfuzji ze skokami do opisu zmian cen instrumentu podstawowego zostało przedstawione w artykułach [Hull, White 1987], [Heston 1993], [Bates 1991], [Wang 2009]. Głównym problemem w tym podejściu jest brak wzorów analitycznych na cenę opcji w modelu dyfuzji ze skokami. Opcje są wyceniane numerycznie, co znacznie wydłuża czas obliczeń.
4. W pracach [Shimko 1993], [Dumas, et al. 1998] przyjęto pewną sparametryzowaną postać powierzchni zmienności (wyrażającej zmienność implikowaną w zależności od ceny wykonania i terminu wykonania). Na podstawie oszacowanej powierzchni zmienności wyznaczono miarę martyngałową, korzystając z wzorów (4) i (5).
5. W artykułach [Melick, Thomas 1997] i [Söderlind, Svensson 1997] zastosowano parametryczną postać funkcji gęstości miary martyngałowej, a następnie szacowano parametry minimalizując różnice między teoretycznymi i obserwowanymi cenami opcji.

W obliczeniach w tym artykule przyjmujemy to ostatnie podejście. Podobnie, jak w [Melick, Thomas 1997] przyjmujemy, że martyngałowy rozkład prawdopodobieństwa cen instrumentu bazowego w chwili wykonania jest mieszaniną rozkładów logarytmiczno-normalnych. Gęstość rozkładu ma zatem postać:

$$q_T(x) = \sum_{i=1}^k w_i f(x, \mu_i, \sigma_i), \quad (6)$$

gdzie wagi $w_i > 0$ sumują się do jedności ($\sum_{i=1}^k w_i = 1$), zaś $f(x, \mu, \sigma)$ to gęstość rozkładu logarytmiczno-normalnego z parametrami μ i σ :

$$f(x, \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi T}} \exp\left(-\frac{1}{2\sigma^2 T} \left(\ln x - \ln S_0 - \mu T + \frac{1}{2}\sigma^2 T\right)^2\right) \quad (7)$$

Ceny opcji w takim modelu są średnimi ważonymi cen wyznaczonych na podstawie wzoru Blacka-Scholesa. Ceny opcji kupna wynoszą

$$c(K, T) = e^{-rT} \sum_{i=1}^k w_i \left[S_0 e^{\left(\mu_i + \frac{1}{2}\sigma_i^2\right)T} N(d_{1i}) - KN(d_{2i}) \right], \quad (8)$$

gdzie S_0 jest obecną ceną instrumentu podstawowego, $N(\cdot)$ to funkcja dystrybuanty standardowego rozkładu normalnego, a współczynniki d_{1i} oraz d_{2i} są równe⁵

$$d_{1i} = \frac{\ln S_0 - \ln K + \left(\mu_i + \frac{\sigma_i^2}{2}\right)T}{\sigma_i \sqrt{T}}, \quad (9)$$

$$d_{2i} = \frac{\ln S_0 - \ln K + \left(\mu_i - \frac{\sigma_i^2}{2}\right)T}{\sigma_i \sqrt{T}}. \quad (10)$$

Natomiast cena opcji sprzedaży jest równa

$$p(K, T) = e^{-rT} \sum_{i=1}^k w_i \left[KN(-d_{2i}) - S_0 e^{\left(\mu_i + \frac{1}{2}\sigma_i^2\right)T} N(-d_{1i}) \right]. \quad (11)$$

Ze wzoru (8), podstawiając $K = 0$, można też otrzymać cenę terminową (*futures*) instrumentu podstawowego która wynosi

$$U_T = e^{-rT} \sum_{i=1}^k w_i S_0 e^{\left(\mu_i + \frac{1}{2}\sigma_i^2\right)T} \quad (12)$$

Oszacowanie miary martyngałowej polega na wyznaczeniu odpowiednich wartości parametrów μ_i , σ_i , w_i , $i = 1, \dots, k$. Dla mieszaniny k rozkładów logarytmicznych należy oszacować $3k - 1$ parametrów.

Współczynniki modelu można kalibrować minimalizując sumy kwadratów odchyłeń między teoretycznymi i obserwowanymi cenami opcji. Załóżmy, że dysponujemy obserwacjami n cen opcji kupna c_1, c_2, \dots, c_n o cenach wykonania K_1, K_2, \dots, K_n i m cen opcji sprzedaży p_1, p_2, \dots, p_m o cenach wykonania H_1, H_2, \dots, H_m , a wszystkie opcje mają ten sam termin wykonania T . Wektor

⁵ Jak można zauważyć, wzory te różnią się od standardowej wersji – w liczniku brak jest parametru r (stopy wolnej od ryzyka). Jest to spowodowane tym, że we wzorze Blacka-Scholesa parametry dryfu (μ) i dyfuzji (σ) są wyznaczone w mierze rzeczywistej, natomiast tutaj są to parametry miary martyngałowej.

parametrów $\theta = (\mu_1, \dots, \mu_k, \sigma_1, \dots, \sigma_k, w_1, \dots, w_k)$ dobieramy rozwiązując następujące zadanie optymalizacji nieliniowej

$$\min_{\theta} \left[\sum_{i=1}^n (c_i - c(K_i, T))^2 + \sum_{i=1}^m (p_i - p(H_i, T))^2 + (U_T - e^{-rT} S_0) \right], \quad (13)$$

pod warunkiem, że

$$\sigma_1, \sigma_2, \dots, \sigma_k > 0 \text{ oraz } w_1, w_2, \dots, w_k \geq 0, \sum_{i=1}^k w_i = 1. \quad (14)$$

DANE I PROCEDURA OBLICZENIOWA

Do estymacji rozkładu prawdopodobieństwa miary martyngałowej na Giełdzie Papierów Wartościowych w Warszawie posłużyliśmy się danymi dotyczącymi cen opcji na indeks WIG20. Notowania cen opcji oraz wartości indeksu zaczerpnęliśmy z bazy danych EIKON Thomson Reuters. Oszacowania przeprowadzaliśmy dla czterech momentów, odpowiadających terminom zapadalności czterech serii opcji: 21 marca 2014, 20 czerwca 2014, 19 września 2014 i 19 grudnia 2014 r. Dla każdej serii notowano opcje o różnych cenach wykonania, przy czym ceny wykonania dla opcji kupna i sprzedaży były zawsze takie same. Dla serii o terminie wykonania 21 marca 2014 oraz 20 czerwca 2014 r. notowano opcje o cenach wykonania: 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000. Dla serii o terminie wykonania 19 września 2014 r. ceny wykonania były następujące: 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000. Dla serii o terminie wykonania 19 grudnia 2014 r. ceny wykonania to: 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900.

Rynek opcji w badanym okresie nie był bardzo płynny. W przypadku każdej serii istniały dni sesyjne, w których nie odbywały się żadne transakcje na opcje z tej serii. Na ogół bardziej płynne były tańsze opcje „nie w cenie” (*out-of-the-money*), przy czym opcje najtańsze charakteryzowały się największą płynnością. W przypadku opcji kupna były to zatem opcje z wyższą ceną wykonania, a w przypadku opcji sprzedaży – z niższą ceną wykonania. Na ogół bardziej płynne były opcje o krótszym terminie wykonania (a więc z wcześniejszych serii). W przypadku braku transakcji GPW jako cenę zamknięcia podaje cenę wyznaczoną na podstawie wzoru Blacka-Scholesa. Uwzględniając te notowania przy szacowaniu miary martyngałowej, otrzymalibyśmy oszacowania obciążone w kierunku rozkładu logarytmiczno-normalnego, co obniżyłoby moc testów stosowanych do sprawdzania, czy rozkłady martyngałowe stóp zwrotu różnią się istotnie od rozkładu normalnego. Dlatego do oszacowania miary martyngałowej w każdym dniu uwzględnialiśmy jedynie notowania opcji, którymi w danym dniu handlowano.

Z uwagi na niską płynność i małą liczbę opcji podlegających obrotowi, musieliśmy przyjąć rozkład o małej liczbie parametrów. Przyjęliśmy podejście opisane w [Melick, Thomas 1997], zgodnie z którym miara martyngałowa jest mieszaniną rozkładów logarytmiczno-normalnych. W [Melick, Thomas 1997] analizowano ceny ropy naftowej w okresie pierwszej wojny w Zatoce Perskiej i przyjęto wobec tego, że rozkład martyngałowy jest mieszaniną trzech rozkładów logarytmiczno-normalnych, które były związane z trzema oczekiwanymi przez rynek rozwiązaniami kryzysu politycznego: 1) powrotem do sytuacji sprzed kryzysu, 2) poważnym uszkodzeniem infrastruktury wydobywczej w krajach Zatoki Perskiej, 3) kontynuacją sytuacji kryzysowej.

W przypadku mieszaniny trzech rozkładów logarytmiczno-normalnych liczba parametrów do oszacowania wynosi 10. Z powodu niskiej płynności opcji notowanych na Giełdzie Papierów Wartościowych w Warszawie zdecydowaliśmy się na przyjęcie, że miara martyngałowa stanowi mieszaninę jedynie dwóch rozkładów logarytmiczno-normalnych. Liczba parametrów do oszacowania wynosi wówczas jedynie 7. Przyjęcie mieszaniny dwóch rozkładów ma oprócz tego naturalną interpretację ekonomiczną: inwestorzy mogą oczekiwać kontynuacji obecnego trendu lub mogą się spodziewać zakłóceń na giełdzie warszawskiej lub w całym regionie Europy Wschodniej.

Dla każdego dnia z okresu od 1 lipca 2013 r. do 19 marca 2014 r. wykonywaliśmy obliczenia dla każdej serii opcji, która była notowana w tym dniu. Miarę martyngałową szacowaliśmy na podstawie cen opcji, dla których wykonywano transakcje (dzienny wolumen obrotu był większy od zera). W przypadku, gdy liczba takich opcji była mniejsza niż 8, dla danego dnia nie wykonywaliśmy obliczeń. Tabela 1 przedstawia podstawowe informacje dotyczące próby. Jak można zauważyć, dla opcji z serii czerwcowej, wrześniowej i grudniowej nie można było przeprowadzić obliczeń w większość dni sesyjnych z powodu zbyt małej liczby opcji, na które odbywały się transakcje. Znacznie bardziej płynne były opcje z serii marcowej. Możliwe było dokonanie oszacowań dla 136 dni sesyjnych (z łącznej próby 179 dni).

Tabela 1. Informacje dotyczące uwzględnionych w badaniu serii opcji

Seria opcji (data wykonania)	Okres obserwacji	Liczba dni w próbie	Liczba dni w badaniu	Liczba aktywnych opcji (średnia)	Aktywnych opcji sprzedaży (średnia)
marcowa (21.3.2014)	1.7.2013- 19.3.2014	179	136	5,3	6,2
czerwcowe (20.12.2014)	1.7.2013- 20.3.2014	180	76	3,2	3,2
wrześniowa (19.9.2014)	23.9.2013- 20.3.2014	121	19	2,0	1,1
grudniowa (19.12.2014)	23.12.2013- 20.3.2014	58	15	1,9	1,8

Źródło: opracowanie własne

Uwagi: Do badania brano tylko „aktywne” opcje, tj. takie, dla których odbywały się transakcje w danym dniu sesyjnym. Ostatnie dwie kolumny podają, ile takich opcji (kupna i sprzedaży) przypadało średnio na dzień sesyjny w okresie obserwacji. W próbie, na podstawie której dokonywano oszacowań, uwzględniono jedynie dni sesyjne, w których liczba „aktywnych” opcji przekraczała 8.

WYNIKI

Prawdopodobieństwa martyngałowe wyznaczaliśmy minimalizując sumę kwadratów odchyłeń cen teoretycznych od cen rynkowych, czyli rozwiązując zadanie (13)-(14). Obliczenia przeprowadzono numerycznie za pomocą funkcji przygotowanych w pakiecie R⁶. Dla porównania, oprócz szacowania miary martyngałowej w modelu z mieszaniną dwóch rozkładów logarytmiczno-normalnych (model ten oznaczamy literą M), wyznaczaliśmy też tę miarę przyjmując, że jej rozkład jest logarytmiczno-normalny (model ten oznaczamy literą L), czyli że spełnione są założenia modelu Blacka-Scholesa. Tabela 2 zawiera informacje na temat dopasowania obu modeli. Kolumny druga i trzecia zawierają dane dotyczące średniej wartości minimalizowanej funkcji celu (13) oraz odchylenia standardowego tej wielkości, obliczone na podstawie wszystkich dni, dla których dokonywano obliczeń. Tabela zawiera też informacje o błędzie średniokwadratowym (MSE) oraz średnim błędzie względnym (MAPE).

Tabela 2. Jakość dopasowania modelu logarytmiczno-normalnego i modelu z mieszaniną rozkładów.

Seria	Suma kwadratów odchyłeń		MSE		MAPE	
	Model L	Model M	Model L	Model M	Model L	Model M
marcowa	1700,2 (1923,9)	826,6 (1295,1)	9,98	7,07	0,31	0,40
czerwcową	8448,2 (23299,6)	4541,4 (11911,2)	24,51	18,48	0,49	0,32
wrześniowa	1360,0 (651,4)	109,9 (191,9)	12,44	3,80	0,29	0,09
grudniowa	1881,3 (2733,6)	1203,1 (2182,8)	15,46	12,66	0,09	0,04

Źródło: obliczenia własne

Tabela 3 zawiera statystyki oszacowanych parametrów dla modelu miary martyngałowej opartym na mieszaninie rozkładów logarytmiczno-normalnych. Przedstawione wartości maksymalne, minimalne, średnie i odchylenia standardowe zostały obliczone na podstawie wyników ze wszystkich dni, w który możliwe było

⁶ R Core Team (2013) R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

wykonanie oszacowań. Jak można zauważyć rozkład pierwszy w mieszaninie rozkładów charakteryzuje się mniejszą wartością oczekiwaną ($\mu_1 < \mu_2$) i większym odchyleniem standardowym ($\sigma_1 > \sigma_2$). Można go zatem zinterpretować jako oczekiwany przez inwestorów „gorszy” stan świata w przyszłości, w którym realizują się niższe stopy zwrotu przy większej zmienności. Parametr w_1 , określający udział pierwszego rozkładu w mieszaninie rozkładów można zatem zinterpretować jako „współczynnik niepokoju”, oznaczający oczekiwane przez inwestorów prawdopodobieństwo realizacji tego gorszego stanu.

Tabela 3. Statystyki oszacowanych parametrów modelu M

Seria		μ_1	μ_2	σ_1	σ_2	w_1
marzec	Max	-0,029	5,467	0,518	0,217	0,805
	Średnia	-1,161	0,268	0,148	0,121	0,338
	Min	-24,083	-0,012	0,000	0,000	0,019
	Odch. std.	3,159	0,484	0,095	0,037	0,170
czerwiec	Max	-0,109	10,211	5,145	2,423	1,000
	Średnia	-2,071	0,878	0,526	0,367	0,504
	Min	-20,879	-2,214	0,000	0,000	0,000
	Odch. std.	3,755	1,306	0,710	0,450	0,209
wrzesień	Max	-0,108	4,627	13,668	7,051	1,000
	Średnia	-11,742	0,309	3,710	0,753	0,282
	Min	-68,533	-1,097	0,354	0,109	0,016
	Odch. std.	18,132	1,220	4,011	1,550	0,299
grudzień	Max	-0,121	2,028	9,733	7,876	1,000
	Średnia	-5,641	0,175	2,632	1,365	0,368
	Min	-35,932	-0,653	0,011	0,186	0,000
	Odch. std.	9,686	0,689	3,019	2,339	0,375

Źródło: obliczenia własne

W Tabeli 4 znajdują się charakterystyki otrzymanych rozkładów. Tabela zawiera informacje na temat wartości oczekiwanej, odchylenia standardowego, współczynnika skośności i kurtozy nadwyżkowej zarówno oszacowanych rozkładów logarytmiczno-normalnych, jak i mieszaniny takich rozkładów (model M). Tabela zawiera średnie wartości charakterystyk, a w nawiasach podano odchylenia standardowe. Jak łatwo zauważyć, stosując mieszaninę rozkładów logarytmiczno-normalnych otrzymuje się na ogół wyższe oszacowania odchylenia standardowego i kurtozy.

Tabela 4. Charakterystyki otrzymanych rozkładów

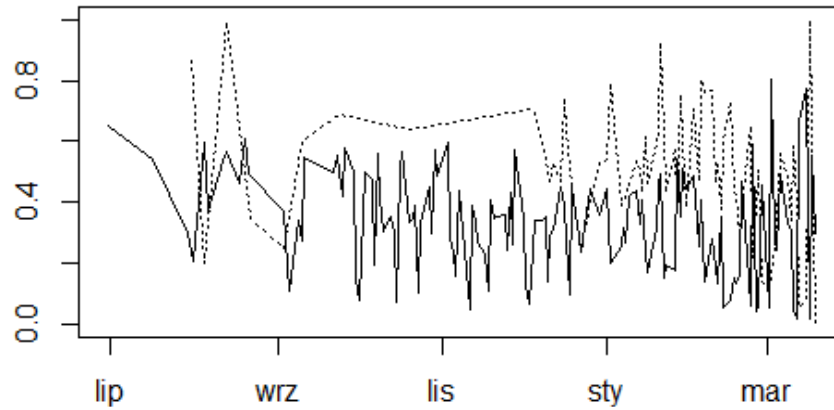
	marcowa	czerwcowe	wrześniowa	grudniowa
	Model L			
Wartość oczekiwana	2430,49 (84,29)	2356,82 (119,22)	2376,70 (70,38)	2343,88 (62,74)
Odchylenie std.	245,43 (81,22)	505,78 (274,24)	403,42 (31,58)	470,84 (18,91)
Wsp. Skośności	83,15 (48,22)	498,58 (1049,85)	208,89 (35,23)	288,37 (25,85)
Kurtoza (nadw.)	0,19 (0,11)	1,70 (6,28)	0,48 (0,09)	0,67 (0,07)
	Model M			
Wartość oczekiwana	2433,94 (85,63)	2381,45 (111,95)	2406,07 (85,64)	2355,33 (49,54)
Odchylenie std.	250,40 (82,07)	525,51 (418,20)	433,74 (43,37)	532,58 (46,62)
Wsp. Skośności	-58,67 (115,40)	940,19 (614,09)	58,33 (293,46)	1189,23 (1224,78)
Kurtoza (nadw.)	0,50 (3,23)	12,82 (99,53)	5,16 (11,79)	62,37 (81,77)

Źródło: obliczenia własne

Rysunek 1 przedstawia wartości parametru w_1 dla opcji serii marcowej i czerwcowej. Największe wartości parametr ten osiąga na początku marca – po rozpoczęciu się niepokojów na Krymie (1 marca Rada Federacji Rosji wyraziła zgodę na użycie wojsk rosyjskich na Krymie, a 11 marca parlament Krymu ogłosił niepodległość). Podobnie, największe różnice między wartościami oczekiwanymi oraz między odchyleniami standardowymi rozkładów stanowiących mieszaninę, również występują na początku marca. Nie dotyczy to opcji z serii wrześniowej i grudniowej; być może z uwagi na ich małą płynność.

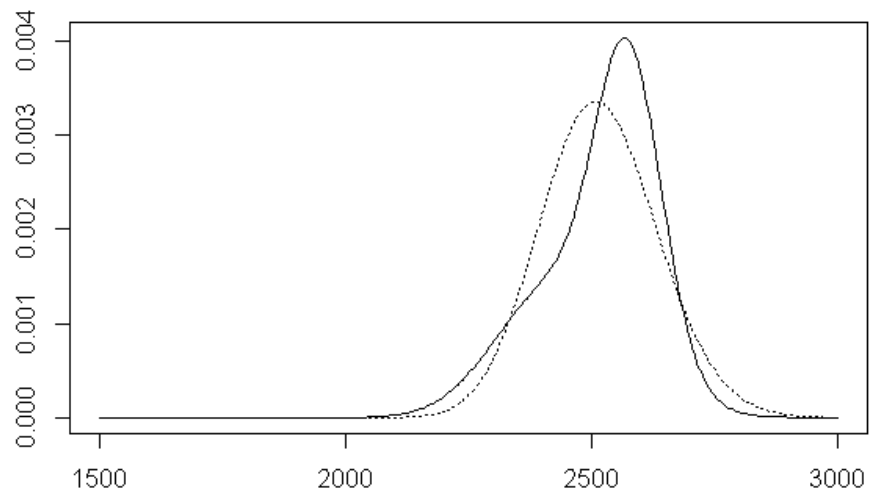
Reakcję na wydarzenia można także wykryć w momentach szacowanych rozkładów. Poniedziałek 3.3.2014 to dzień gwałtownego skoku zarówno skośności, jak i kurtozy dla marcowej serii opcji. Współczynnik skośności spadł do -552,27, a nadwyżkowa kurtoza wzrosła do 8,1, podczas gdy poprzedniego dnia sesyjnego, 28.2.2014 r. wielkości te wynosiły odpowiednio -75, 8 i 0,28, a we wcześniejszych dniach utrzymywały się na zbliżonym poziomie. Rysunki 3 i 4 przedstawiają wykresy funkcji gęstości oszacowanych rozkładów, zarówno dla modelu M, jak i dla modelu L, odpowiednio w dniach 28.3.2014 i 3.3.2014r.

Rysunek 2. Wartość parametru w_1 dla opcji z serii marcowej (linia ciągła) i czerwcowej (linia przerywana)



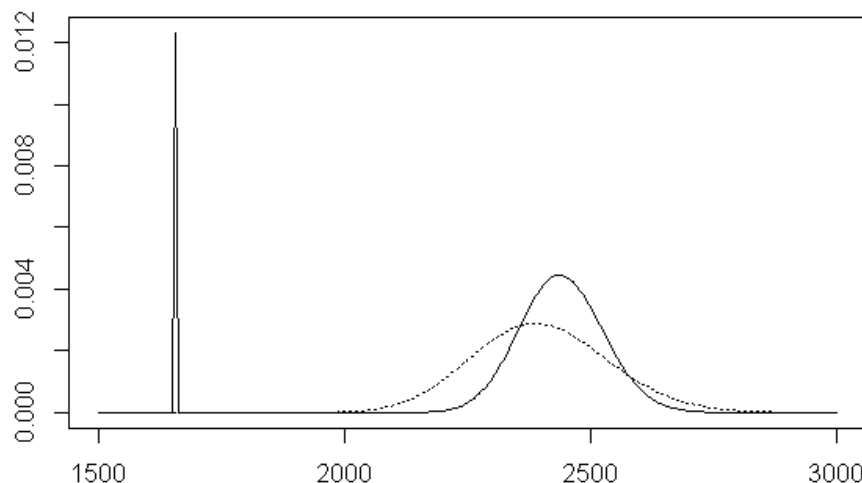
Źródło. opracowanie własne

Rysunek 3. Mieszana rozkładów logarytmiczno-normalnych (linia ciągła) i pojedynczy (linia przerywana) rozkład logarytmiczno-normalny dla cen opcji z dnia 28.2.2014



Źródło: opracowanie własne

Rysunek 4. Mieszanka rozkładów logarytmiczno-normalnych (linia ciągła) i pojedynczy rozkład logarytmiczno-normalny (linia przerywana) dla cen opcji z dnia 3.3.2014



Źródło: opracowanie własne

WNIOSKI

W artykule oszacowaliśmy miary martyngałowe, reprezentujące rynkowe oczekiwania odnośnie przyszłych cen, na podstawie cen opcji na indeks WIG20. Rozważyliśmy dwa modele: w pierwszym z nich przyjęto rozkład logarytmiczno-normalny, a w drugim - mieszaninę dwóch takich rozkładów. Drugi z tych modeli ma oczywistą interpretację: rozkłady w mieszaninie reprezentują dwa możliwe stany świata w przyszłości, których oczekuje rynek. Parametr przedstawiający procentowy udział rozkładów w mieszaninie można wówczas interpretować jako „współczynnik niepokoju” co do przyszłego rozwoju sytuacji rynkowej.

Mimo niskiej płynności opcji notowanych na Giełdzie Papierów Wartościowych w Warszawie, próba oszacowania parametrów miary martyngałowej zakończyła się powodzeniem dla wielu dni sesyjnych. Należy jednak zwrócić uwagę, że dotyczy to opcji o krótkim okresie do wykonania. Opcje o dłuższym okresie są zbyt mało płynne. Próbuąc wykorzystać szacunki miary martyngałowej do prognozowania ekonomicznego, jak to zaproponowano to np. w [de Vincent-Humphereys, Noss 2012], należy ograniczyć się jedynie do prognoz krótkoterminowych.

Wyniki badań pokazały, że zastosowanie mieszaniny rozkładów daje zazwyczaj wyższe charakterystyki rozproszenia rozkładu (wariancję i kurtozę). Na ogół model z mieszaniną rozkładów był lepiej dopasowany do danych niż model oparty na pojedynczym rozkładzie logarytmiczno-normalnym.

BIBLIOGRAFIA

- Aït-Sahalia, Y., Lo, A.W. (2000) Nonparametric risk management and implied risk aversion, *Journal of Econometrics*, vol. 94, ss. 9-51.
- Aït-Sahalia, Y., Jacod, J. (2009) Testing for jumps in a discretely observed process. *The Annals of Statistics*, vol. 37, ss. 184-222.
- Aparicio, S., Hodges, S. (1998) Implied risk-neutral distribution: a comparison of estimation methods, working paper, Warwick University.
- Bahra B. (1997) Implied risk-neutral probability density functions from option prices: theory and application, Bank of England Working Paper No. 66.
- Bates D.S. (1991) The crash of '87: was it expected? The evidence from option markets, *Journal of Finance*, vol. 46, ss. 1009-1044.
- Bjork, T. (2009) *Arbitrage Theory in Continuous Time*, Oxford University Press.
- Birru, J., Figlewski, S. (2012) An anatomy of a meltdown: the risk neutral density from the S&P 500 in the fall of 2008, *Journal of Financial Markets*, vol. 15, ss. 151-180.
- Breeden, D.T., Litzenberg, L.H. (1978) Prices of state-contingent claim implicit in option prices, *Journal of Business*, vol. 51, ss. 621-651.
- Chabi-Yo, F., Garcia, R., Renault, E. (2008) State dependence can explain the risk aversion puzzle, *Review of Financial Studies*, vol. 21, ss. 973-1011.
- Derman E., Kani I. (1994) Riding on a smile, *Risk*, vol. 7, ss. 32-39.
- Dumas, B., Fleming, J., Whaley, R. (1998) Implied volatility functions: empirical test, *Journal of Finance*, vol. 53, ss. 2059-2106.
- Durpie, B. (1994) Pricing with a smile, *Risk*, vol. 7, ss. 18-20.
- Fusai G., Roncoroni, A. (2008) *Implementing Models in Quantitative Finance: Methods and Cases*, Springer.
- Gagnon, M.-H., Power, G.J. (2013) Investor risk aversion and market shocks: event studied using options on crude oil, working paper, SSRN, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2126343 (dostęp: 22.3.2014).
- Heston, S. (1993) A closed-form solution for options with stochastic volatility application to bond and currency options, *Review of Financial Studies*, vol. 6, ss. 327-343.
- Hull, J. (2009) *Options, Futures and Other Derivatives*, Prentice Hall.
- Hull, J., White, A. (1987) The pricing of options on assets with stochastic volatilities, *Journal of Finance*, vol. 42, ss. 281-300.
- Jackwerth, J. (1999) Option implied risk-neutral distributions and implied binomial trees: a literature review, *Journal of Derivatives*, vol. 7, ss. 66-82.
- Jackwerth, J., Rubinstein M. (1996) Recovering probability distributions from option prices, *Journal of Finance*, vol. 51, ss. 1611-1631.
- Liu, X., Shackleton, M.B., Taylor, S.J., Xu, X. (2007) Closed-form transformations from risk-neutral to real-world distributions, *Journal of Banking and Finance*, vol. 31, ss. 1501-1520.
- Mandler, M. (2002) Comparing risk-neutral probability density functions implied by option prices – market uncertainty and ECB-council meetings, referat prezentowany na 29th Annual Meeting of the European Finance Association, Berlin.
- Mehra, R., Prescott, E.C. (1985) The Equity Premium: A Puzzle, *Journal of Monetary Economics*, vol. 15, ss. 145-161.

- Melick, W.R., Thomas C.P. (1997) Recovering an asset's implied PDF from option prices: an application to crude oil during Gulf crisis, *Journal of Financial and Quantitative Analysis*, vol. 32, ss. 91-115.
- Musiela, M., Rutkowski, M. (2008) *Martingale Methods in Financial Modelling*, Springer.
- Pliska, S.R. (2005) *Wprowadzenie do matematyki finansowej. Modele z czasem dyskretnym*, WNT, Warszawa.
- R Core Team (2013) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Rubinstein, M. (1994) Implied binomial trees, *Journal of Finance*, vol. 49, ss. 771-818.
- Sadorsky, P. (2001) Risk factor in stock returns of Canadian oil and gas companies, *Energy Economics*, vol. 23, ss. 17-28.
- Shimko, D. (1993) Bounds of probability, *Risk*, vol. 6, ss. 33-37.
- Söderlind, P., Svensson, L. (1997) New techniques to extract market expectations from financial instruments, *Journal of Monetary Economics*, vol. 40, ss. 383-429.
- de Vincent-Humphereys, R., Puigvert, J. (2012) A quantitative mirror on the Euribor market using implied probability density functions, *Eurasian Economic Review*, vol. 2, ss. 1-31.
- de Vincent-Humphereys, R., Noss, R. (2012) Estimating probability distributions of future asset prices: empirical transformations from option-implies risk-neutral to real-world density function, *Bank of England Working Paper No. 455*.
- Wang, Y.-H. (2009). The impact of jump dynamics on the predictive power of option-implied densities. *Journal of Derivatives*, vol. 16, ss. 9-22.
- Ziegler, A. (2007). Why does implied risk aversion smile? *Review of Financial Studies*, vol. 20, ss. 859-904.

ESTIMATION OF RISK NEUTRAL MEASURE FOR POLISH STOCK MARKET

Abstract: In the paper we present the usage of risk neutral measure estimation to the analysis of the index WIG20 from Polish stock market. The risk neutral measure is calculated from the prices of the options on that index. We assume that risk neutral measure is the mixture of lognormal distributions. The parameters of the distributions are estimated by minimizing the sum of squares of pricing errors. Obtained results are then compared with the model based on a single lognormal distribution. As an example we consider changes in risk neutral distribution at the beginning of March 2014, after the outbreak of political crisis in the Crimea.

Keywords: risk-neutral pricing, option-implied density, risk aversion, real-world measure, event study

**ZASTOSOWANIE MIERNIKÓW
TAKSONOMICZNYCH DO OCENY
EFEKTYWNOŚCI FINANSOWEJ GIEŁD EUROPEJSKICH
W LATACH 2002-2011**

Krzysztof Kompa

Katedra Ekonometrii i Statystyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
krzysztof_kompa@sggw.pl

Streszczenie: Celem artykułu jest ocena efektywności finansowej europejskich giełd papierów wartościowych traktowanych jako przedsiębiorstwa. Efektywność finansowa jest mierzona za pomocą mierników taksonomicznych, do budowy których, jako zmienne diagnostyczne, wykorzystano podstawowe wskaźniki finansowe giełd-przedsiębiorstw: przychody, koszty i EBITDA. Badania zostały przeprowadzone na podstawie danych FESE z lat 2002-2011.

Słowa kluczowe: giełda papierów wartościowych, efektywność finansowa, FESE, mierniki taksonomiczne

WPROWADZENIE

Globalizacja i integracja gospodarcza są determinantami sytuacji we współczesnym świecie. Z jednej strony - usuwają bariery formalne i techniczne dla przepływu ludzi, kapitału i pracy, dynamizując rozwój gospodarczy i wymianę gospodarczą, rozwój naukowy, techniczny i społeczny. Z drugiej jednak, tworząc z państw, gospodarek i społeczeństw system naczyń połączonych, ułatwiają propagację zjawisk społecznie niepożądanych. Sprawiają, iż procesy o ograniczonym dotąd terytorialnie charakterze i znaczeniu, nabierają rangi ponadpaństwowej - regionalnej, kontynentalnej lub wręcz globalnej. Każde to patrzeć na procesy globalizacji i integracji także w kategoriach bezpieczeństwa alokacji kapitału na dowolnie wybranym rynku finansowym.

Bezpieczeństwo i pewność obrotu gospodarczego są zazwyczaj pochodną lokalnego rozwoju społeczno-gospodarczego, stopnia dojrzałości gospodarki rynkowej, a w konsekwencji - rozwoju lokalnych instytucji rynku kapitałowego i ich osadzenia w strukturach państwa i prawa. Zamożność i efektywność ekonomiczna giełdy, jako przedsiębiorstwa, organizatora obrotu na rynku, może mieć istotny wpływ na spełnienie w/w warunków, implikując tym samym atrakcyjność alokacyjną obsługiwanego rynku i atrakcyjność inwestycyjną własnych aktywów. Porządkowanie względem siebie giełd-przedsiębiorstw wydaje się mieć zatem dla praktyki inwestowania znaczenie szczególne, a pozycjonowanie Giełdy Papierów Wartościowych w Warszawie na tle innych giełd europejskich – kluczowe, w kontekście aspiracji GPW do roli regionalnego centrum finansowego (*Warsaw - CEE Financial Hub*) oraz prywatyzacji i upublicznienia obrotu jej akcjami.

W dostępnej literaturze stosunkowo dużo miejsca¹ poświęca się zagadnieniom klasyfikacji giełd, analizując pozycjonowanie rynków pod względem różnych ich własności – kapitalizacji i jej dynamiki (w odniesieniu do rynków, ich segmentów i sektorów), wolumenów i wartości obrotów, liczby i dynamiki IPO, stóp zwrotów z indeksów, portfeli i spółek, zmienności notowanych aktywów, powiązań przyczynowych i in. Wykorzystuje się przy tym szeroki wachlarz metod badawczych, w tym – ekonometryczne, statystyczne i wielowymiarowej analizy porównawczej (WAP). Niewiele natomiast opracowań pozycjonuje względem siebie giełdy jako przedsiębiorstwa.

W konsekwencji, celem prezentowanego badania jest próba budowy rankingu europejskich giełd-przedsiębiorstw, ze szczególnym uwzględnieniem miejsca GPW w tym uszeregowaniu. Analiza zostanie przeprowadzona za pomocą trzech różnych mierników taksonomicznych na podstawie danych za lata 2002-2011.

MIERNIKI TAKSONOMICZNE

Metody wielowymiarowej analizy statystycznej są jednym z podstawowych narzędzi umożliwiających porównywanie obiektów niejednorodnych, których rozwój warunkowany jest czynnikami o różnym charakterze, a literatura przedmiotu jest dość bogata (por. [Łuniewska, Tarczyński 2006], [Młodak 2006], [Nowak 1990], [Pluta 1977], [Pociecha i in. 1988]). Analizy prowadzone są w kilku etapach, wśród których wyróżnia się: (1) wybór zmiennych diagnostycznych, (2) normalizację zmiennych, (3) wybór miernika, (4) wyznaczenie wartości miernika dla wszystkich analizowanych obiektów, (5) pozycjonowanie i grupowanie obiektów na podstawie wartości miernika. Etap pierwszy sprowadza się do oceny merytorycznej czynników wpływających na rozwój badanego obiektu i przeprowadzenia analiz statystycznych, koniecznych zwłaszcza w przypadku znacznej liczby

¹ por. np. [Majewska 2004; Łuniewska 2005; Łuniewska, Tarczyński 2006; Malinowska 2007; Egert, Kočenda 2007; Witkowska, Zdziarski 2008]; szeroki przegląd piśmiennictwa zamieszczono w [Witkowska, Matuszewska, Kompa 2008], [Kompa 2010a].

zmiennych. Zadaniem drugiego etapu jest pozabawienie zmiennych miana i sprowadzenie ich do względnej porównywalności. Przy wyborze miernika głównie określa się czy w jego konstrukcji uwzględniać się będzie tzw. wzorec rozwoju (jeśli tak, to należy również zdefiniować wzorec) oraz czy korzystać się będzie z metod porządkowania liniowego lub metod wykorzystujących funkcje nieliniowe. Pozostałe dwa etapy są konsekwencją wcześniej podjętych decyzji i ewentualnie wymagają dodatkowo ustalenia reguł grupowania obiektów.

W przeprowadzonym badaniu wykorzystano trzy mierniki taksonomiczne:

1. syntetyczny miernik rozwoju (SM), wykorzystujący wzorec z próby,
2. wskaźnik względnego poziomu rozwoju (W) – metodę sum standaryzowanych,
3. wektorową miarę syntetyczną (WMS).

Miara SM wywodzi się z koncepcji Hellwiga (1968) i jest wyznaczana jako:

$$SM_{it} = 1 - \frac{q_{it}}{\bar{q}_t + 2 \cdot S_{qt}} \quad i = 1, 2, \dots, n; t = 1, 2, \dots, T \quad (1)$$

$$q_{it} = \sqrt{\frac{\sum_{j=1}^k (z_{it}^j - \bar{z}_{0t}^j)^2}{k}} \quad j = 1, 2, \dots, k \quad (2)$$

$$z_{it}^j = \frac{x_{it}^j - \bar{x}_t^j}{S_{xt}^j}, \bar{x}_t^j = \frac{\sum_{i=1}^n x_{it}^j}{n}, S_{xt}^j = \sqrt{\frac{\sum_{i=1}^n (x_{it}^j - \bar{x}_t^j)^2}{n}}, z_{0t}^j = \begin{cases} \min_{i=1, 2, \dots, n} \{z_{it}^j\} & \text{if } x_{it}^j \in D \\ \max_{i=1, 2, \dots, n} \{z_{it}^j\} & \text{if } x_{it}^j \in S \end{cases} \quad (3)$$

gdzie dla każdego okresu t : q_{it} – odległość Euklidesowa analizowanego i -tego obiektu od wzorca, \bar{q}_t, S_{qt} – średnia i odchylenie standardowe odległości, x_{it}^j, z_{it}^j – oryginalne i standaryzowane zmienne opisujące j -tą cechę diagnostyczną w i -tym obiekcie, \bar{x}_t^j, S_{xt}^j – średnia i odchylenie standardowe j -tej cechy diagnostycznej, z_{0t}^j – wartości j -tej zmiennej wzorca rozwoju, D, S – ozn. destymulant i stymulant.

Metoda sum standaryzowanych jest metodą bez wzorca i polega na prostym sumowaniu zmiennych po ich normalizacji. Przyjmując poprzednie oznaczenia, wskaźnik względnego rozwoju wyznacza się na podstawie wzorów:

$$W_i = \frac{\sum_{j=1}^m z_{ij}^*}{\sum_{j=1}^m \max_i \{z_{ij}^*\}}; z_{ij}^* = z_{ij} + \left| \max_i \{z_{ij}^*\} \right| \quad (4)$$

Wektorowa miara syntetyczna WSM pochodzi z opracowania Nermenda (2008). Jest wynikiem analizy m składowych wektora cech \vec{x}_t^i klasyfikowanego

(w chwili t) i -tego obiektu, wzdłuż pewnego wektora kierunkowego $\vec{\omega}_t$, rozpiętego pomiędzy końcami wektorów wzorca i antywzorca,

$$WSM_{it} = \frac{\sum_{j=1}^m [(z_{jt}^i - z_{jt_0}^{aw}) \omega_{jt_0}]}{\sum_{j=1}^m (\omega_{jt_0})^2}, i = 1, 2, \dots, n; j = 1, 2, \dots, m; t = 1, 2, \dots, T \quad (5)$$

gdzie: i – numer obiektu (i -ta giełda), n – liczba analizowanych obiektów (giełd), j – numer cechy, m – liczba cech (wymiar przestrzeni), t – numer okresu badania, t_0 – okres odniesienia, T – liczba okresów badania, ω_{jt_0} – różnica j -tych współrzędnych wzorca i antywzorca w przestrzeni m -wymiarowej: $\omega_{jt_0} = z_{jt_0}^w - z_{jt_0}^{aw}$, a ponadto: $z_{jt}^i, z_{jt_0}^i, z_{jt_0}^w, z_{jt_0}^{aw}$ – standaryzowane, dla okresu odniesienia t_0 , wartości j -tych składowych wektorów wzorca, antywzorca, i -tej giełdy oraz – znormalizowane względem okresu odniesienia t_0 – składowe wektorów i -tych giełd dla pozostałych okresów badania t , wyznaczone ze zbiorów wartości stymulant S i destymulant D :

$$z_{jt_0}^i = (x_{jt_0}^i - \bar{x}_{jt_0}) / S_{jt_0} \quad z_{jt}^i = (x_{jt}^i - \bar{x}_{jt_0}) / S_{jt_0}, t \neq t_0$$

$$z_{jt_0}^w = \begin{cases} \min_{i=1,2,\dots,n} \{z_{jt_0}^i\} & \text{dla } x_{jt_0}^i \in D \\ \max_{i=1,2,\dots,n} \{z_{jt_0}^i\} & \text{dla } x_{jt_0}^i \in S \end{cases} \quad z_{jt_0}^{aw} = \begin{cases} \max_{i=1,2,\dots,n} \{z_{jt_0}^i\} & \text{dla } x_{jt_0}^i \in D \\ \min_{i=1,2,\dots,n} \{z_{jt_0}^i\} & \text{dla } x_{jt_0}^i \in S \end{cases} \quad (6)$$

W przeciwieństwie do miar SM i W, których wartości należą do przedziału $\langle 0; 1 \rangle$ i bliższe jedności oznaczają wyższy poziom efektywności przedsiębiorstwa, miara wektorowa przyjmuje wartości z tego przedziału jedynie dla obiektów lepszych (lub takich samych) jak antywzorzec i gorszych (lub takich samych) jak wzorzec. Dopuszcza zarazem wartości spoza przedziału $\langle 0; 1 \rangle$ dla obiektów lepszych od wzorca i gorszych od antywzorca

Przy wykorzystaniu różnych miar syntetycznych pojawia się problem stabilności wyników, a w szczególności odpowiedzi na pytanie czy porządkowania obiektów na podstawie różnych mierników taksonomicznych dają zbieżne wyniki. Klasycznie wykorzystuje się w tym celu współczynniki korelacji tau Kendalla i rang Spearmana (por. [Łuniewska, Tarczyński 2006, s. 90-91]).

Inną kwestią jest zdolność dyskryminacyjna mierników, określająca na ile są one w stanie poprawnie rozpoznać obiekty, z punktu widzenia poziomu ich rozwoju. Przykładem miary służącej ocenie przydatności różnych mierników rozwoju jest współczynnik G_i Sokołowskiego [Sokołowski 1984, Kompa 2010b]:

$$G_i = 1 - \sum_{i=1}^{n-1} \min \left\{ \frac{MR_i - MR_{i+1}}{R}, \frac{1}{n-1} \right\}; R = \max_i \{MR_i\} - \min_i \{MR_i\} \quad (7)$$

gdzie: MR_i , MR_{i+1} - uporządkowane niemalejąco wartości miernika, G_i przyjmuje wartości z przedziału $\langle 0, 1 - [1/(n-1)] \rangle$. Wysokie wartości wskaźnika wskazują na dużą zdolność miernika taksonomicznego do grupowania.

WYNIKI KLASYFIKACJI GIEŁD EUROPEJSKICH

Giełdy papierów wartościowych są zazwyczaj traktowane jako reprezentacja krajowego, bądź regionalnego rynku kapitałowego. Zatem w analizach uwzględnia się m. in. ich kapitalizację, obroty, płynność, liczbę notowanych spółek i instrumentów oraz ich znaczenie w gospodarce (por. [Witkowska, Zdziarski 2008; Krawczyk 2008; Kompa, Witkowska 2014]). Zasadniczo inne podejście stosuje się do wyboru cech charakteryzujących wartość przedsiębiorstwa w ujęciu fundamentalnym². Jednym z lepiej ugruntowanych jest podejście prezentowane przez Tarczyńskiego (1994, 2002) przy wyborze zmiennych do taksonomicznej miary atrakcyjności inwestycji *TMAI*. W ślad za tymi pracami, uwzględniając dostępność danych na podstawie raportów FESE, do analizy pozycjonowania giełd europejskich zostały wybrane³: przychody analizowanych przedsiębiorstw, ich koszty oraz EBITDA (zysk przed opodatkowaniem i amortyzacją).

W badaniach uwzględniono 19 europejskich giełd papierów wartościowych zrzeszonych w FESE, dla których istniała możliwość pozyskania kompletnych obserwacji za okres 2002-2011. Wszystkie zmienne mierzone są w tys. EUR na koniec grudnia każdego roku badania. Wyjątek stanowi londyńska LSEG (*London Stock Exchange Group*), powstała z połączenia London Stock Exchange i Borsa Italiana w 2007r., dla której dane za lata 2002-2006 zostały wygenerowane na podstawie informacji FESE.

Pozycjonowanie giełd europejskich, wyznaczone na podstawie obliczonych mierników syntetycznych, zaprezentowano w tabelach 1-3. Wyraźnie widoczny w tych zestawieniach jest stabilny podział analizowanych przedsiębiorstw na trzy grupy typologiczne: giełdy o najwyższym poziomie efektywności, giełdy średnio rozwinięte i giełdy przeciętnie efektywne. We wszystkich analizowanych latach Deutsche Börse charakteryzuje największą efektywność wg wskazań syntetycznych miar rozwoju SM i WMS. Zajmuje także czołową, I-szą lub II-gą (na zmianę z giełdą EuroNEXT) pozycję w rankingu wg miary bezwzorcowej. EuroNEXT pozycjonuje się na drugim miejscu rankingu: wg SM w latach 2002-2005, dla miernika W: w latach 2002-2007, a wg WMS: w 2002-2005, 2008, 2010, 2011. Giełda londyńska jest trzecią, najbardziej efektywną giełdą europejską: utrzymuje drugą pozycję wg SM w latach 2006-2011, WMS: 2006-2007, natomiast wg miary zagregowanej W giełda ta zajmowała 3. pozycję w całym okresie analizy.

² Zob. np. [Rappaport 1999], [Tarczyński 1994, 2002]

³ W podobny sposób skonstruowano mierniki w pracy [Kompa 2010]. Wszystkie zmienne charakteryzowały się wysoką zmiennością.

Tabela 1. Pozycje analizowanych giełd wyznaczone wg miernika SM

Giełda	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Athens Exchange	7	7	7	7	7	7	8	8	8	10
BME	4	4	5	5	4	3	4	3	4	4
Bratislava SE	17	17	18	19	19	19	19	18	18	17
Bulgarian SE	18	18	17	15	18	16	18	19	19	19
CEESEG - Budapest	14	14	13	14	14	13	14	14	14	14
CEESEG - Ljubljana	15	15	15	17	17	17	17	17	17	18
CEESEG - Prague	13	13	14	13	13	12	13	13	12	12
CEESEG - Vienna	11	11	11	9	10	8	9	9	10	9
Cyprus SE	19	19	19	18	15	15	16	16	16	16
Deutsche Börse	1	1	1	1	1	1	1	1	1	1
Irish SE	12	12	12	12	11	14	12	12	13	13
London SE	3	3	3	3	2	2	2	2	2	2
Luxembourg SE	9	9	9	11	12	11	11	11	11	11
Malta SE	16	16	16	16	16	18	15	15	15	15
OMX Nordic	5	6	4	4	5	6	6	5	5	6
NYSE Euronext	2	2	2	2	3	4	3	4	3	3
Oslo Børs	8	8	8	8	8	10	7	7	7	7
SIX Swiss Exchange	6	5	6	6	6	5	5	6	6	5
Warsaw SE	10	10	10	10	9	9	10	10	9	8

Źródło: obliczenia własne

Tabela 2. Pozycje analizowanych giełd wyznaczone wg miernika W

Giełda	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Athens Exchange	7	7	7	7	7	7	8	8	8	10
BME	6	6	6	5	5	4	4	4	4	4
Bratislava SE	18	18	18	19	19	19	19	18	18	18
Bulgarian SE	19	19	19	16	18	16	18	19	19	19
CEESEG - Budapest	14	14	14	14	14	14	14	14	14	14
CEESEG - Ljubljana	15	15	15	17	16	17	17	17	17	17
CEESEG - Prague	12	13	13	13	13	13	13	13	13	12
CEESEG - Vienna	11	11	10	9	9	8	9	9	10	9
Cyprus SE	16	16	16	15	15	15	15	15	15	16
Deutsche Börse	1	1	1	1	1	1	2	2	2	2
Irish SE	13	12	12	12	12	12	12	12	12	13
London SE	3	3	3	3	3	3	3	3	3	3
Luxembourg SE	9	9	9	11	11	11	11	11	11	11
Malta SE	17	17	17	18	17	18	16	16	16	15
OMX Nordic	4	4	4	4	4	6	6	6	5	6
NYSE Euronext	2	2	2	2	2	2	1	1	1	1
Oslo Børs	8	8	8	8	8	10	7	7	7	7
SIX Swiss Exchange	5	5	5	6	6	5	5	5	6	5
Warsaw SE	10	10	11	10	10	9	10	10	9	8

Źródło: obliczenia własne

Tabela 3. Pozycje analizowanych giełd wyznaczone wg miernika WMS

Giełda	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Athens Exchange	5	5	6	7	7	7	7	7	8	10
BME	4	4	4	4	4	3	4	2	3	4
Bratislava SE	14	16	18	17	18	18	16	16	17	16
Bulgarian SE	17	13	16	15	15	15	17	17	18	17
CEESEG - Budapest	11	10	13	13	13	13	14	13	12	13
CEESEG - Ljubljana	12	15	17	18	17	16	18	18	16	18
CEESEG - Prague	18	12	14	14	14	12	13	11	13	12
CEESEG - Vienna	15	11	12	10	10	9	8	8	10	9
Cyprus SE	19	18	19	19	19	14	19	19	19	19
Deutsche Börse	1	1	1	1	1	1	1	1	1	1
Irish SE	10	9	9	11	11	19	11	14	14	14
London SE	3	3	3	3	2	2	3	4	4	3
Luxembourg SE	9	7	10	12	12	11	12	10	11	11
Malta SE	16	14	15	16	16	17	15	15	15	15
OMX Nordic	7	19	5	5	6	5	6	5	5	5
NYSE Euronext	2	2	2	2	3	4	2	3	2	2
Oslo Børs	8	6	8	8	8	10	9	12	7	7
SIX Swiss Exchange	6	17	7	6	5	6	5	6	6	6
Warsaw SE	13	8	11	9	9	8	10	9	9	8

Źródło: obliczenia własne

Do najmniej efektywnych giełd, z punktu widzenia wyznaczonych mierników SM i W, należą giełdy w Bratysławie, Sofii, na Cyprze, Malcie i w Ljublanie, zajmując trzy ostatnie pozycje w rankingu. Natomiast pozycjonowanie najsłabszych giełd wg wskazań miernika wektorowego WMS nie jest już tak jednoznaczne - najsłabszą jest giełda cypryjska, następnie giełdy w Bratysławie i Ljublanie oraz, w pojedynczych okresach, giełdy w Irlandii, Pradze i Sofii, a nawet OMX Nordic (w 2003 r.).

GPW zajmuje pozycje środkowe we wszystkich rankingach (miejsca od 8. do 10.), chociaż w przypadku WMS w dwóch latach znajduje się na pozycjach 11. i 13. Biorąc pod uwagę stabilność pozycji w kolejnych latach zauważa się znaczną monotonię w przypadku miar SM i W oraz zmienność w przypadku miary wektorowej, co wynika zapewne z przyjęcia wspólnego wzorca dla wszystkich okresów badania. W zasadzie nie obserwuje się istotnych zmian w pozycjach poszczególnych giełd wywołanych kryzysem finansowym w wyjątkiem pogorszenia się pozycji giełdy w Atenach w 2011r. Może to świadczyć o tym, że wszystkie giełdy europejskie zostały dotknięte skutkami kryzysu w podobny sposób.

ANALIZA STABILNOŚCI MIERNIKÓW SYNTETYCZNYCH

Wyniki badania zdolności dyskryminacyjnej wykorzystanych mierników syntetycznych zaprezentowano w tabeli 4., zamieszczając wartości współczynnika

Sokołowskiego G_i wyznaczone dla analizowanych mierników w kolejnych latach badania. Z zestawienia wynika, że relatywnie najmniejszą zdolnością do grupowania obiektów charakteryzuje się popularny miernik rozwoju SM, co jest szczególnie widoczne w 2011r. Pozostałe dwa mierniki mają zbliżone własności, chociaż miara bezwzorcowca częściej daje wyższe wskazania miary G_i niż bardziej skomplikowana obliczeniowo miara WMS.

Tabela 4. Wartości miernika G_i

	Rok									
	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
SM	0,582	0,619	0,563	0,609	0,556	0,583	0,558	0,607	0,582	0,366
W	0,701	0,694	0,693	0,689	0,670	0,752	0,760	0,734	0,728	0,749
WMS	0,717	0,673	0,714	0,706	0,677	0,662	0,732	0,749	0,705	0,708

Źródło: obliczenia własne

Tabela 5. Współczynniki korelacji obliczona dla sąsiadujących lat

Miara	Współczynnik	Lata								
		02/03	03/04	04/05	05/06	06/07	07/08	08/09	09/10	10/11
SM	τ -Kendalla	0,988	0,953	0,906	0,895	0,871	0,883	0,965	0,965	0,953
	ρ -Spearmana	0,998	0,991	0,983	0,977	0,972	0,970	0,995	0,995	0,991
W	τ -Kendalla	0,988	0,988	0,918	0,977	0,930	0,918	0,988	0,977	0,942
	ρ -Spearmana	0,998	0,998	0,979	0,995	0,984	0,981	0,998	0,997	0,990
WMS	τ -Kendalla	0,567	0,661	0,918	0,965	0,790	0,754	0,895	0,883	0,930
	ρ -Spearmana	0,661	0,691	0,981	0,995	0,907	0,893	0,968	0,961	0,986

Źródło: obliczenia własne

Tabela 6. Współczynniki korelacji obliczona dla par mierników syntetycznych

Miara	Współczynnik	Lata									
		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
SM	τ -Kendalla	0,813	0,626	0,918	0,930	0,906	0,918	0,883	0,836	0,918	0,918
WMS	ρ -Spearmana	0,921	0,661	0,981	0,988	0,974	0,970	0,974	0,946	0,979	0,983
WMS	τ -Kendalla	0,743	0,556	0,836	0,883	0,848	0,871	0,871	0,801	0,895	0,906
W	ρ -Spearmana	0,881	0,589	0,954	0,970	0,963	0,942	0,967	0,933	0,972	0,979
W	τ -Kendalla	0,930	0,930	0,918	0,953	0,941	0,953	0,965	0,941	0,953	0,965
SM	ρ -Spearmana	0,982	0,982	0,982	0,988	0,991	0,989	0,993	0,986	0,991	0,993

Źródło: obliczenia własne

Badanie stabilności wskazań wykorzystanych mierników za pomocą współczynników korelacji przedstawiono w tabeli 5., porównując wskazania dla sąsiadujących lat, oraz w tabeli 6. – porównując parami obliczone indykatory. Biorąc pod uwagę rankingi giełd w dwóch kolejnych latach, zauważa się dużą ich zbieżność dla wszystkich okresów i typów mierników taksonomicznych z wyjątkiem miary wektorowej w latach 2002-2004, w przypadku której współczynniki korelacji są relatywnie niskie. Najbardziej stabilne rankingi obserwuje się dla miernika bez

wzorca W. Wartości współczynników korelacji wyznaczone dla każdej z trzech par miar są relatywnie wysokie z wyjątkiem roku 2003 dla obu par, w których występowała miara wektorowa, co potwierdza wcześniejsze spostrzeżenia.

PODSUMOWANIE

Celem niniejszego artykułu była analiza kondycji finansowej 19 giełd europejskich w latach 2002-2011 przy wykorzystaniu mierników taksonomicznych. Z przedstawionych analiz wynika, że w ciągu 10 lat analiz nie zaobserwowano zmian liderów wśród giełd zrzeszonych w FESE. Innymi słowy na rozwiniętych i dużych rynkach kapitałowych (tj. paneuropejska NYSE Euronext, połączone giełdy londyńska i włoska, Deutsche Börse) giełdy jako podmioty gospodarcze wykazują efektywność finansową niezależnie od sytuacji na globalnym rynku.

W badaniach wykorzystano trzy mierniki syntetyczne, których budowa jest znacząco zróżnicowana. Wykazano, że rozpoznawanie najbardziej efektywnych instytucji finansowych nie zależy od zastosowanej metody, chociaż wykorzystanie miernika wektorowego daje mniejszą zbieżność wyników z pozostałymi metodami.

BIBLIOGRAFIA

- Egert B., Kočenda E. (2007) Interdependence between Eastern and Western European Stock Markets: Evidence from Intraday Data. *Economic Systems*, 31(2), 184-203. ISSN 0939-3625.
- Hellwig Z. (1968) Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowania kadr, *Przegląd Statystyczny*, nr 4, 307-326.
- Kompa K. (2010a) Capital markets in transitional countries. Comparison of stock exchanges in Europe, [in:] Witkowska D., Nerment K. (eds.) *Regional Analysis: Globalization, Integration, Transformation*, Szczecin: University of Szczecin.
- Kompa K. (2010b) Klasyfikacja giełd europejskich na potrzeby zarządzania ryzykiem alokacyjnym. Zastosowanie wielowymiarowej analizy porównawczej, II konferencja Naukowa „Skuteczność w Biznesie” Gorzów.
- Kompa K., Witkowska D. (2011) Capital Markets in the Baltic States in Years 2000–2010. Preliminary Investigation, [in:] Lacina L., Rozmahel P., Rusek A. (eds.) *Financial and Economic Crisis: Causes, Consequences and the Future*. Bučovice: Martin Stříž Publishing, 244 – 269.
- Kompa K., Witkowska D. (2014) Comparison of European Stock Exchanges One- and Multi-Dimensional Analysis, *Indian Journal of Fundamental and Applied Life Sciences* 2014, Vol. 4 (S1) April-June, pp. 2111-2126, ISSN: 2231–6345, An Open Access, Online International Journal Available at <http://http://www.cibtech.org/sp.ed/jls/2014/01/jls.htm>
- Krawczyk E. (2008) Miejsce Warszawskiej Giełdy Papierów Wartościowych wśród wybranych giełd Europy Środkowej, *Ekonomika i Organizacja Gospodarki Żywnościowej*, 72, 71-180.

- Łuniewska M. (2005) Evaluation of Selected Methods of Classification for the Warsaw Stock Exchange, *International Advances in Economic Research*, 2005, Vol. 11, No 4, 469-481.
- Łuniewska M., Tarczyński W. (2006) *Metody wielowymiarowej analizy porównawczej na rynku kapitałowym*, Warszawa: Wydawnictwo Naukowe PWN.
- Majewska A. (2004) *Pozycja Giełdy Papierów Wartościowych w Warszawie wobec przystąpienia Polski do Unii Europejskiej, Rynek kapitałowy w przededniu integracji z Unią Europejską*, Szczecin: PTE.
- Malinowska A. (2008) *Gospodarcze znaczenie rynku kapitałowego w Polsce i wybranych krajach Europy w latach 2000–2007*, Warszawa: Oficyna Wydawnicza SGH.
- Młodak A. (2006) *Analiza taksonomiczna w statystyce regionalnej*, Warszawa: Difin.
- Nermend K. (2008) *Rachunek wektorowy w analizie rozwoju regionalnego*, Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- Nowak E. (1990) *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, Warszawa: PWE.
- Pluta W. (1977) *Wielowymiarowa analiza porównawcza w badaniach ekonomicznych*, Warszawa: PWE.
- Pociecha J., Podolec B., Sokołowski A., Zając K. (1988) *Metody taksonomiczne w badaniach społeczno-ekonomicznych*, Państwowe Wydawnictwo Naukowe, Warszawa.
- Rappaport A., *Wartość dla Akcjonariuszy*, WIG Press, Warszawa 1999.
- Sokołowski A. (1984) *Wybrane zagadnienia pomiaru i ważenia cech w taksonomii*, Zeszyty Naukowe nr 203, Kraków: Akademia Ekonomiczna w Krakowie.
- Tarczyński W. (1994) *Taksonomiczna miara atrakcyjności inwestycji w papiery wartościowe*, *Przegląd Statystyczny* Nr 3.
- Tarczyński W. (2002) *Fundamentalny portfel papierów wartościowych*. PWE, Warszawa.
- Witkowska D., Matuszewska A., Kompa K. (2008) *Wprowadzenie do ekonometrii dynamicznej i finansowej*, Warszawa: Wydawnictwo SGGW.
- Witkowska D., Zdziarski T. (2008) *Porównanie GPW w Warszawie z giełdami europejskimi za pomocą syntetycznego miernika taksonomicznego*, [w:] Binderman Z. (red.), *Metody ilościowe w badaniach ekonomicznych: Wielowymiarowa analiza danych*, Warszawa: Wydawnictwo SGGW, 155-163.

APPLICATION OF TAXONOMIC MEASURES TO EVALUATION OF THE EUROPEAN STOCK EXCHANGES FINANCIAL EFFICIENCY

Abstract .The aim of the paper is evaluation of financial efficiency of the European stock exchanges that are treated as enterprises. The financial efficiency is measured applying taxonomic measures using revenue, costs and EBITDA as diagnostic variables. The research is provided on the basis of data provided by FESE for years 2002-2011.

Keywords: stock exchange, financial efficiency, FESE, taxonomic measure

WYKORZYSTANIE ANALIZY WIELOKRYTERIALNEJ DO BADANIA POTENCJAŁU GOSPODARCZEGO WOJEWÓDZTWA PODKARPACKIEGO

Michał Kościółek

Katedra Ekonomii

Politechnika Rzeszowska im. Ignacego Łukasiewicza

e-mail: mkos@prz.edu.pl

Streszczenie: W artykule przedstawiono wykorzystanie metod analizy wielokryterialnej za pomocą których został zbadany potencjał gospodarczy województwa podkarpackiego. Z wielu czynników wpływających na rozwój oraz utrzymanie potencjału gospodarczego wyodrębniono 5 cech. Wyodrębnienia dokonano za pomocą metody AHP. Na podstawie wybranych cech dokonano badania metodą unitaryzacji zerowanej oraz metodą Czekanowskiego. Badanie wykazało, że wybór metody do analizy wielokryterialnej ma niewielki wpływ na grupowanie poszczególnych powiatów w województwie podkarpackim.

Słowa kluczowe: analiza wielokryterialna, metoda AHP, metoda unitaryzacji zerowanej, metoda Czekanowskiego, potencjał gospodarczy, województwo podkarpackie, klasyfikacja

WSTĘP

Potencjał gospodarczy jest określany poprzez umiejętność danego kraju – postrzeganego jako jednostka gospodarcza jak i polityczna- do długookresowego tworzenia wartości dodanej[Stern i in. 2000]¹.

Do badania potencjału gospodarczego danego podmiotu można wykorzystać wiele metod analizy wielokryterialnej. Dobór metod jest kluczowy, ponieważ nie wszystkie metody pozwalają na ich porównanie. Celem opracowania jest próba

¹ Stern S., Porter M.E., Furman J.L., (2000), *The Determinants of National Innovative Capacity*, "Working Paper" No 7876, National Bureau of Economic Research, Cambridge, s. 1.

określenia potencjału województwa podkarpackiego za pomocą metod analizy wielokryterialnej. Dla tak określonego celu pracy zostanie zweryfikowana następująca hipoteza: potencjał gospodarczy województwa podkarpackiego kształtuje się w taki sam sposób niezależnie od zastosowanej metody badawczej.

OPIS PORÓWNYWANYCH METOD

W opracowaniu zostały zastosowane trzy metody analizy wielokryterialnej:

- metoda AHP (Analytic Hierarchy Process),
- metoda unitaryzacji zerowanej,
- Diagram Czekanowskiego.

Metoda AHP jest jedna z najbardziej popularnych i najszybciej rozwijających się metod matematycznych, pozwalającej rozwiązywać problemy wielokryterialne. Wspomaga ona rozwiązanie złożonych problemów, poprzez rozłożenie ich na czynniki. Może być też pomocna w przypadku porównywania dwóch lub większej ilości obiektów, w celu wyboru optymalnego rozwiązania.

Metoda ta obejmuje podejście wielokryterialne oparte na preferencjach. Każdy oceniający ocenia badane warianty według własnej, subiektywnej oceny. Ważnym założeniem jest porównywalność badanych wariantów.

Metoda AHP uwzględnia specyfikę procesów wartościowania, które mają charakter hierarchiczny. Jej zastosowanie jest szczególnie widoczne w podejmowaniu decyzji o charakterze ekonomicznym czy technicznym w przypadku występowania jedynie kryteriów jakościowych. Oceniający badane warianty powinien być osobą mającą doświadczenie w danym obszarze [Downarowicz i in. 2000]².

Etapami w przeprowadzaniu analizy metodą AHP są [Saaty 2004]³:

- tworzenie hierarchicznej struktury procesu decyzyjnego,
- dobór osób oceniających (ekspertów) oraz obliczenie ocen dla elementów hierarchii,
- tworzenie macierzy preferencji,
- tworzenie rankingu finalnego.

Pierwszy etap ma charakter analityczny. Od jakości analizy i doboru wariantów (obiektów) zależy jej wynik końcowy. Niezmiernie ważne w tym etapie jest, aby osoba zajmująca się dekompozycją problemu zajmowała się badanym tematem i miała doświadczenie w tym obszarze.

² Downarowicz O., Krause J., Sikorski M., Stachowski W. (2000), *Zastosowanie metody AHP do oceny i sterowania poziomem bezpieczeństwa złożonego obiektu technicznego*, Politechnika Gdańska, Gdańsk, str. 7-9.

³ Saaty T. L. (2004), *Decision Making – The Analytic Hierarchy and Network Processes (AHP/ANP)*. *Journal of Systems Science and Systems Engineering*, Tsinghua University, Beijing, Vol. 13, No. 1, str. 1-11.

Kolejnym etapem jest dobór grona eksperckiego. Osoby te będą porównywać parami czynniki (cele) według skali Saaty'ego. Na tym etapie obliczane są również oceny dla poszczególnych elementów hierarchii. W tabeli 1 została zaprezentowana skala preferencji pomiędzy porównywanymi elementami, której używa się do badania metodą AHP.

Tabela 1. Skala preferencji między porównywanymi elementami według Saaty'ego

Przewaga ważności elementów decyzyjnych	Opisowe wyrażenie ważności elementów decyzyjnych	Przyporządkowane wartości
Równoważność	X ma takie samo znaczenie jak Y (oba elementy przyczyniają się w równym stopniu do osiągnięcia celu)	1
Słaba lub umiarkowana	X ma trochę większe znaczenie niż Y (nieprzekonujące znaczenie lub słaba preferencja jednego elementu nad drugim)	3
Istotna, zasadnicza, mocna	X ma wyraźnie większe znaczenie niż Y (zasadnicze lub mocne znaczenie lub mocna preferencja jednego elementu nad innym)	5
Zdecydowana lub bardzo mocna	X ma bezwzględnie większe znaczenie niż Y (zdecydowane znaczenie lub bardzo mocna preferencja jednego elementu nad innym)	7
Absolutna	X ma absolutnie większe znaczenie niż Y (absolutne znaczenie lub absolutna preferencja jednego elementu nad innym)	9

Źródło: Winnicki K., Jurek A., Landowski M., *Zastosowanie metody analizy hierarchicznej problemu*, str. 156, <http://p-e.up.krakow.pl/article/viewFile/682/554>

Dwa ostatnie etapy to tworzenie i badanie spójności macierzy preferencji i na jej podstawie tworzenie rankingu końcowego. Ranking końcowy ma na celu pomoc w wyborze najlepszego rozwiązania.

Zaletą przedstawionej metody jest jej prosty sposób obliczania. Do przeprowadzenia obliczeń wystarczy jedynie arkusz kalkulacyjny. Metoda AHP ma na celu ułatwienie wyboru optymalnego wariantu (scenariusza).

Drugą z zastosowanych metod jest metoda unitaryzacji zerowanej. Metoda ta opiera się na porównywaniu wybranych elementów pod względem wielu kryteriów, które niekiedy obrazują odmienne od siebie zmienne. Zmienne często wyrażone są poprzez różne od siebie wielkości. Metoda ta ma na celu uprządkowanie i unormowanie tych cech. Obejmuje ona zarówno elementy ujemne

(destymulanty) jak i dodatnie (stymulanty) [Kukuła 2012]⁴. Normowanie cech odbywa się za pomocą wzorów [Miszczyńska]⁵:

$$z_{ij} = \frac{\max_i x_{ij} - x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \quad (i = 1, 2, \dots, r; j = 1, 2, \dots, s) \quad (1)$$

$$z_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \quad (i = 1, 2, \dots, r; j = 1, 2, \dots, s) \quad (2)$$

Wzór (1) odnosi się do destymulant, natomiast wzór (2) do stymulant. Dzięki tym wzorom tworzona jest macierz zawierająca wybrane wskaźniki. Stworzenie macierzy pozwala w późniejszym etapie na stworzenie rankingu obiektów i określenia grup podobnych do siebie.

Metoda unitaryzacji zerowanej ma na celu przedstawienie wielu kryteriów za pomocą jednej unormowanej zmiennej.

Ostatnią zastosowaną metodą jest metoda Czekanowskiego, zwana również diagramem Czekanowskiego. Metoda ta jest najstarszą metodą taksonomiczną i jest to również uniwersalna metoda klasyfikacji statystycznej.

Diagram Czekanowskiego pozwala na określenie powiązań między obiektami (jednostkami) oraz unaocznic związek i podobieństwa między nimi. Dzięki tej metodzie można stworzyć grupy obiektów charakteryzujące się podobnymi cechami [Wójcik 2012]⁶.

Przy stosowaniu tej metody trzeba trzymać się określonych etapów i wykonywać je sumiennie aby uzyskany wynik był miarodajny. Przy przygotowaniu diagramu Czekanowskiego, należy postępować według następującej procedury [Heffner i in. 2007]⁷:

1. zdefiniowanie macierzy danych,
2. doprowadzenie danych do porównywalności poprzez zastosowanie standaryzacji lub unitaryzacji zmiennych,
3. wybór miary podobieństwa,

⁴ Kukuła K. (2012), Propozycja budowy rankingu obiektów z wykorzystaniem cech ilościowych oraz jakościowych, *Metody ilościowe w badaniach ekonomicznych*, tom XIII/1, str. 6-13.

⁵ Miszczyńska D., Metoda unitaryzacji zerowanej, http://www.kbo.uni.lodz.pl/index.php?option=com_docman&task=doc_download&Itemid=33&gid=1017 [dostęp: 20.08.2014]

⁶ Wójcik A. (2013), Zastosowanie diagramu Czekanowskiego do badania podobieństw krajów Unii Europejskiej pod względem pozyskiwania energii ze źródeł odnawialnych, http://zif.wzr.pl/pim/2013_4_4_25.pdf [dostęp: 20.08.2014]

⁷ Heffner K., Gibas P. (2007), *Analiza ekonomiczno-przestrzenna*, WUE, Katowice, str.55.

4. obliczenie macierzy odległości taksonomicznych,
5. wyszukanie w macierzy odległości taksonomicznych wartości minimalnej i maksymalnej,
6. wyznaczenie przedziałów klasowych dla wartości odległości taksonomicznych na podstawie wyznaczonego empirycznego obszaru zmienności,
7. opracowanie nieuporządkowanego diagramu Czekanowskiego,
8. uzyskanie uporządkowanego diagramu Czekanowskiego,
9. interpretacja uzyskanych wyników.

Ostatnim etapem jest interpretacja wyników, może ona zostać przedstawiona w postaci uporządkowanego diagramu Czekanowskiego wraz z opisem, lub w postaci rankingu końcowego grupującego obiekty (jednostki) podobne. Metoda ta pozwala wyodrębnić najlepsze obszary.

ANALIZA WYBRANYMI METODAMI BADAWCZYMI

Pierwszym etapem analizy było wyodrębnienie z dziesięciu wskaźników pięciu najważniejszych określających potencjał gospodarczy. Do tego wykorzystano metodę AHP. Grupa ekspertów (grupa składająca się z pięciu osób) porównywała pary wskaźników określając przewagę jednego elementu nad drugim wykorzystując skalę preferencji według Saaty'ego. Z dziesięciu wskaźników wyodrębniono pięć najważniejszych. Są to wskaźniki:

- podmioty wpisane do rejestru REGON na 10 tys. ludności,
- przeciętne wynagrodzenie miesięczne brutto,
- stopa bezrobocia rejestrowanego,
- nakłady inwestycyjne przedsiębiorstw w przeliczeniu na jednego mieszkańca,
- migracje – zameldowania na pobyt stały.

Gdy zostały wyodrębnione wskaźniki określające w najlepszy sposób potencjał gospodarczy zaczęto analizę metodą unitaryzacji zerowanej i metodą (diagramem) Czekanowskiego. W początkowej fazie analizy metodą unitaryzacji zerowanej należało określić czy dany wskaźnik jest stymulantą czy destymulantą, i tak:

- podmioty wpisane do rejestru REGON na 10 tys. ludności - stymulanta,
- przeciętne wynagrodzenie miesięczne brutto - stymulanta,
- stopa bezrobocia rejestrowanego - destymulanta,
- nakłady inwestycyjne przedsiębiorstw w przeliczeniu na jednego mieszkańca - stymulanta,
- migracje – zameldowania na pobyt stały - stymulanta.

Po tym etapie tworzona jest macierz zawierająca wszystkie wyżej wymienione wskaźniki unormowana przy pomocą wzoru 1 i 2. Wyniki tej analizy zostały przedstawione w tabeli 2.

Tabela 2. Macierz określająca potencjał gospodarczy województwa podkarpackiego za pomocą metody unitaryzacji zerowanej

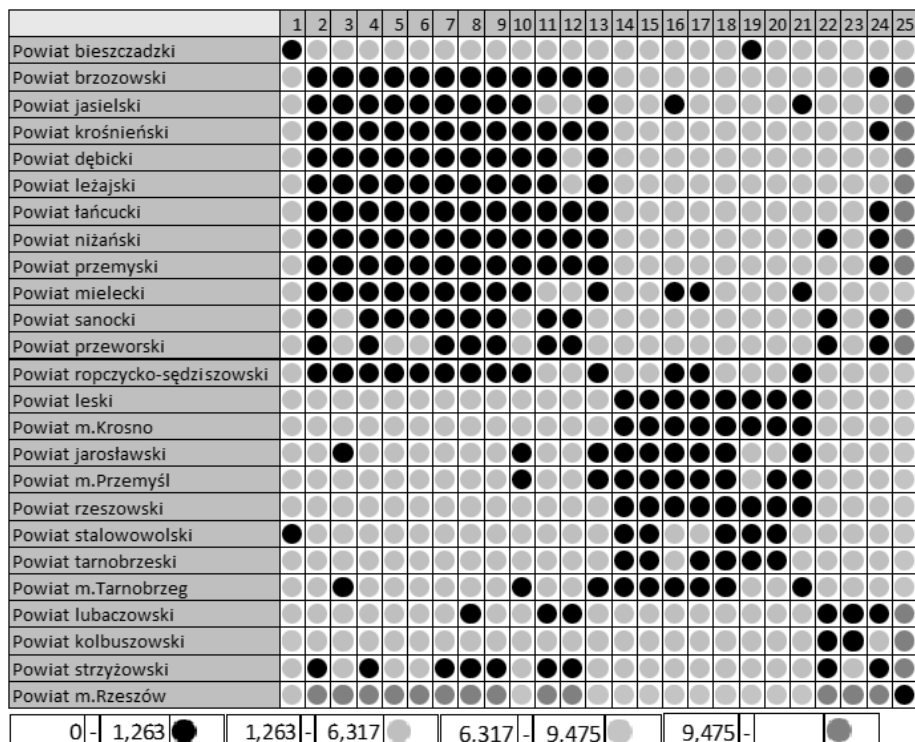
Jednostki terytorialne	x1	x2	x3	x4	x5	Q
Powiat m. Rzeszów	1,00	1,00	1,00	0,75	1,00	0,95
Powiat m. Krosno	0,91	0,08	0,98	1,00	0,13	0,62
Powiat stalowowolski	0,40	0,56	0,60	0,70	0,17	0,48
Powiat tarnobrzeski	0,16	0,47	0,56	0,99	0,10	0,46
Powiat rzeszowski	0,19	0,23	0,63	0,38	0,68	0,42
Powiat m. Tarnobrzeg	0,71	0,35	0,56	0,22	0,09	0,39
Powiat mielecki	0,30	0,48	0,69	0,33	0,11	0,38
Powiat dębicki	0,18	0,31	0,68	0,49	0,11	0,35
Powiat bieszczadzki	0,64	0,40	0,06	0,60	0,00	0,34
Powiat m. Przemyśl	0,59	0,53	0,41	0,01	0,12	0,33
Powiat sanocki	0,28	0,21	0,70	0,28	0,09	0,31
Powiat jarosławski	0,22	0,44	0,34	0,32	0,14	0,29
Powiat leski	0,70	0,41	0,13	0,13	0,00	0,28
Powiat jasielski	0,22	0,29	0,37	0,23	0,12	0,25
Powiat łańcucki	0,25	0,11	0,37	0,26	0,11	0,22
Powiat leżajski	0,09	0,40	0,34	0,20	0,05	0,22
Powiat krośnieński	0,18	0,09	0,34	0,20	0,25	0,21
Powiat ropczycko-sędziszowski	0,16	0,33	0,20	0,24	0,07	0,20
Powiat przemyski	0,00	0,27	0,27	0,11	0,21	0,17
Powiat lubaczowski	0,04	0,18	0,42	0,15	0,02	0,16
Powiat przeworski	0,05	0,15	0,33	0,08	0,11	0,14
Powiat kolbuszowski	0,04	0,04	0,48	0,08	0,02	0,13
Powiat brzozowski	0,06	0,20	0,05	0,00	0,08	0,08
Powiat strzyżowski	0,08	0,00	0,12	0,06	0,08	0,07
Powiat niżański	0,10	0,09	0,00	0,00	0,10	0,06

Źródło: opracowanie własne na podstawie danych z GUS

Wartość Q w tabeli 2 odzwierciedla potencjał każdego z powiatów w województwie podkarpackim w odniesieniu do wszystkich pięciu wskaźników. W analizie wyróżnić można powiat miasto Rzeszów, który swoim wynikiem odbiega od reszty powiatów.

Kolejnym etapem analizy było wykreślenie diagramu Czekanowskiego, który miał za zadanie określenie podobieństw w potencjale gospodarczym w województwie podkarpackim. Diagram został stworzony w programie Microsoft Excel 2007.

Rysunek 1. Diagram Czekanowskiego określający potencjał poszczególnych powiatów w województwie podkarpackim



Źródło: opracowanie własne

Diagram Czekanowskiego zaprezentowany na rysunku 1, przedstawia grupy powiatów o podobnym potencjale gospodarczym na terenie województwa podkarpackiego. Skala zastosowana na rysunku pozwala na wskazanie regionów o podobnym potencjale gospodarczym. Im wartość wyższa tym powiat ma większy potencjał gospodarczy. Wyodrębnić można trzy grupy powiatów które posiadają podobny potencjał gospodarczy. Jedynie powiat miasto Rzeszów odbiega pozytywnie od reszty powiatów, ponieważ posiada wyższy potencjał gospodarczy natomiast powiat bieszczadzki prezentuje się w tym zestawieniu najgorzej.

PODSUMOWANIE

Badanie potencjału gospodarczego w województwie podkarpackim wykazało, że niezależnie od zastosowanej metody powiat miasto Rzeszów wyróżniało się w pozytywny sposób. Rozwój tego powiatu odbiegał od reszty i jest on wzorem do naśladowania przez inne powiaty.

Tabela 3. Porównanie wyników uzyskanych obiema metodami badawczymi

L.p	Powiat	Grupa wg MUZ	Grupa wg DC
1	Powiat m. Rzeszów	I	I
2	Powiat m. Krosno	II	III
3	Powiat stalowowolski	III	III
4	Powiat tarnobrzegi	III	III
5	Powiat rzeszowski	III	III
6	Powiat m. Tarnobrzeg	III	III
7	Powiat mielecki	III	II
8	Powiat dębicki	III	II
9	Powiat bieszczadzki	IV	V
10	Powiat m. Przemyśl	IV	III
11	Powiat sanocki	IV	II
12	Powiat jarosławski	IV	III
13	Powiat leski	IV	III
14	Powiat jasielski	IV	II
15	Powiat łańcucki	IV	II
16	Powiat leżajski	IV	II
17	Powiat krośnieński	IV	II
18	Powiat ropczycko-sędziszowski	IV	II
19	Powiat przemyski	V	II
20	Powiat lubaczowski	V	IV
21	Powiat przeworski	V	II
22	Powiat kolbuszowski	V	IV
23	Powiat brzozowski	V	II
24	Powiat strzyżowski	V	IV
25	Powiat niżański	V	II

Źródło: opracowanie własne

W przypadku większości powiatów zauważyć można, że niezależnie od zastosowanej metody zostały one sklasyfikowane w tej samej grupie lub najbliższej (tabela 3). Brak identycznej grupy może być uwarunkowany tym, iż w przypadku tworzenia grup w obu przypadkach przedziały zostały stworzone niezależnie.

Przyjęta hipoteza została potwierdzona. Niezależnie od wyboru metody badawczej rozkład potencjału gospodarczego w województwie podkarpackim przedstawiał się identycznie.

BIBLIOGRAFIA

- Downarowicz O., Krause J., Sikorski M., Stachowski W. (2000), Zastosowanie metody AHP do oceny i sterowania poziomem bezpieczeństwa złożonego obiektu technicznego, Politechnika Gdańska, Gdańsk, str. 7-9.
- Heffmer K. (2007), Gibas P., Analiza ekonomiczno-przestrzenna, WUE, Katowice, str.55.
- Kukuła K. (2012), Propozycja budowy rankingu obiektów z wykorzystaniem cech ilościowych oraz jakościowych, Metody ilościowe w badaniach ekonomicznych, tom XIII/1, str. 6-13.
- Miszczyńska D., Metoda unitaryzacji zerowanej,
http://www.kbo.uni.lodz.pl/index.php?option=com_docman&task=doc_download&Itemid=33&gid=1017 [dostęp: 20.08.2014].
- Saaty T. L. (2004), Decision Making – The Analytic Hierarchy and Network Processes (AHP/ANP). Journal of Systems Science and Systems Engineering, Tsinghua University, Beijing, Vol. 13, No. 1, str. 1-11.
- Stern S., Porter M.E., Furman J.L., (2000), The Determinants of National Innovative Capacity, “Working Paper” No 7876, National Bureau of Economic Research, Cambridge, s. 1.
- Winnicki K., Jurek A., Landowski M., Zastosowanie metody analizy hierarchicznej problemu, str. 156, <http://p-e.up.krakow.pl/article/viewFile/682/554>.
- Wójcik A. (2013), Zastosowanie diagramu Czekanowskiego do badania podobieństw krajów Unii Europejskiej pod względem pozyskiwania energii ze źródeł odnawialnych, http://zif.wzr.pl/pim/2013_4_4_25.pdf [dostęp: 20.08.2014].

APPLICATION OF MULTI-CRITERIA ANALYSIS IN RESEARCH ON ECONOMIC POTENTIAL OF PODKARPACKIE PROVINCE

Abstract: The article will outline the application of multi-criteria analysis by the use of which the economic potential of Podkarpackie Province will be examined. From numerous factors influencing the development and maintenance of economic potential, 5 features were distinguished. It was done with the aid of selection in the use of AHP method. On the basis of distinguished features the examination was made by the use of zero unitarization method as well as Czekanowski's method. The examination showed that the choice of method for multi-criteria analysis has little effect on grouping of particular communes in Podkarpackie Province.

Keywords: multi-criteria analysis, AHP method, zero unitarization method, Czekanowski's method, economic potential, Podkarpackie Province, classification

WRAŻLIWOŚĆ MIARY SYNTETYCZNEJ NA WIELKOŚCI KRYTYCZNE WSKAŹNIKÓW SŁUŻĄCYCH DO JEJ BUDOWY

Sławomir Lisek

Katedra Statystyki Matematycznej
Uniwersytet Rolniczy w Krakowie
e-mail: sl1@op.pl

Streszczenie: Syntetyczna miara sytuacji finansowej przedsiębiorstwa jest użytecznym narzędziem w analizie sytuacji finansowej podmiotów gospodarczych. Zbudowana została na bazie diagnostycznych wskaźników kondycji ekonomicznej firmy, znormalizowanych, przyjmując za podstawę normalizacji ich wielkość krytyczną. W niniejszym artykule zbadano wrażliwość miary m na prawidłowość ustalenia wartości krytycznych mierników służących do ich budowy. Wykazano iż w przypadku znaczących zmian w otoczeniu przedsiębiorstwa, konieczne jest stałe monitorowanie wartości krytycznych i aktualizowanie ich.

Słowa kluczowe: sytuacja finansowa przedsiębiorstwa, miara syntetyczna, wielowymiarowa analiza porównawcza, wskaźniki diagnostyczne

WSTĘP

Celem artykułu jest zbadanie wrażliwości miary syntetycznej kondycji finansowej przedsiębiorstwa m , na wielkości krytyczne mierników diagnostycznych, przyjętych do jej konstrukcji. Należy zweryfikować tezę, iż właściwe ustalenie wielkości krytycznych wskaźników, na bazie których zbudowano miarę syntetyczną, ma kluczowe znaczenie dla jej przydatności w ocenie ekonomicznej pozycji jednostki gospodarczej.

MIARA SYNTETYCZNA m I WIELKOŚCI UŻYTE DO JEJ UTWORZENIA

Miara syntetyczna m i zasady jej konstrukcji.

Syntetyczną miarę kondycji finansowej m tworzy się na bazie diagnostycznych mierników stanu finansowego firmy. Miernikami diagnostycznymi są:

- zadłużenie kapitału własnego,
- rentowność kapitału własnego.
- rentowność netto majątku,
- płynność szybka¹

Spośród wymienionych powyżej wielkości destymulantą jest zadłużenie kapitału własnego, pozostałe wielkości są stymulantami. Miarę syntetyczną m oblicza się według wzoru (1) lub wzoru (2)²:

$$m = \frac{\frac{Rkw}{0,05} + \frac{Rnm}{0,03} + \frac{Ztr}{1,00} + \frac{1,2}{Zkw}}{4} \quad (1)$$

Gdzie:

Rkw – rentowność kapitału własnego,

Rnm – rentowność netto majątku,

Ztr – płynność szybka,

Zkw – zadłużenie kapitału własnego

Jeśli miarę m obliczamy zgodnie z wzorem (1) to za neutralną przyjmuje się sytuację przedsiębiorstwa gdy ta miara wynosi 1, natomiast jeśli przekracza 1 to oznacza dobrą sytuację finansową przedsiębiorstwa, gdy jest niższa od jedności sytuacja jest zła.

Drugim sposobem obliczania miary m jest³:

$$m = \frac{\frac{Rkw - 0,05}{0,05} + \frac{Rnm - 0,03}{0,03} + \frac{Ztr - 1,00}{1,00} + \frac{1,2 - Zkw}{Zkw}}{4} \quad (2)$$

oznaczenia jak w (1).

¹ S. Lisek „Mierniki Syntetyczne w procesie diagnozowania sytuacji finansowej przedsiębiorstwa” Zeszyty Naukowe Uniwersytetu Szczecińskiego nr 542 z 2009 roku.

² tamże, s. 263-270

³ tamże, s. 263-270

Jeśli miarę m oblicza się zgodnie z wzorem (2) to za neutralną przyjmuje się sytuację przedsiębiorstwa gdy ta miara wynosi 0, natomiast jeśli jest dodatnia to oznacza dobrą sytuację finansową przedsiębiorstwa, gdy jest ujemna sytuacja jest zła. Miary te można uznać za średnią arytmetyczną znormalizowanych cech diagnostycznych. Przy czym za podstawę normalizacji przyjęto wielkość krytyczną danej cechy diagnostycznej, czyli minimalną dopuszczalną w przypadku stymulant, lub też maksymalną tolerowaną w przypadku destymulanty. Odstąpiono konstruując ten wskaźnik od ważenia zmiennych diagnostycznych, ponieważ nie jest możliwe ustalenie obiektywnych wag pomiędzy poszczególnymi zmiennymi. W niniejszym artykule miara m obliczana jest według wzoru (2)

Wielkości krytyczne wielkości diagnostycznych użytych do budowy miary m .

Wielkości krytyczne poszczególnych wskaźników ustalono jako:

- zadłużenie kapitału własnego – maksymalnie 1,2,
- rentowność kapitału własnego – minimum 0,05,
- rentowność netto majątku – minimum 0,03,
- płynność szybka – minimum 1,0

Za podstawę ustalenia wielkości krytycznych przyjęto generalnie wcześniejszy dorobek naukowy dotyczący analizy finansowej. Wskaźnik szybkiej płynności nie powinien być niższy od jedności, by istniała możliwość pełnej spłaty zobowiązań krótkoterminowych z łatwo osiągalnych środków płatniczych.⁴

Zadłużenie kapitału własnego absolutnie nie powinno przekraczać 2⁵, jednak w polskich realiach, przy dużym ryzyku działalności, wielkość krytyczną tego wskaźnika należy obniżyć do 1,2.

Wielkość krytyczną zyskowności kapitału własnego (ROE) przyjęto na poziomie 5% biorąc pod uwagę oferowaną przez banki stopę odsetek od lokat w momencie konstrukcji tej miary (2008 rok), a także stopień ryzyka związany z inwestowaniem kapitału w udziały lub akcje przedsiębiorstwa.⁶

Wartość progową ROA przyjęto biorąc pod uwagę krytyczną wartość ROE i zadłużenia kapitału własnego.

Przesłanki stałości lub zmienności w czasie krytycznych wielkości wskaźników diagnostycznych.

Porównywalność w czasie obliczeń na podstawie miary m przemawia na pierwszy rzut oka za utrzymaniem stałych w czasie wielkości krytycznych, wskaźników diagnostycznych służących budowie miary m . Jednakże bliższe

⁴ Analiza ekonomiczna w przedsiębiorstwie, 2004r., praca zbiorowa pod red. M. Jerzmanowskiej, Polskie Wydawnictwo Ekonomiczne Warszawa s. 138

⁵ tamże, s. 157

⁶ Analiza ekonomiczna przedsiębiorstwa, 1996r., praca zbiorowa po red. B. Wersty, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu.

spojrzenie na procesy zachodzące w gospodarce – szczególnie w gospodarce Polski, nakazują jednak nabrać wątpliwości, czy wobec zmieniającej się sytuacji gospodarczej nie należy wielkości krytycznych aktualizować, szczególnie w dłuższym horyzoncie czasowym.

Oceniając wskaźnik płynności szybkiej należy powiedzieć iż w tym przypadku postulat możliwości spłaty zobowiązań natychmiast wymagalnych z natychmiast dostępnych zasobów, nie zmienia się w czasie. Należy więc przyjąć iż wartość krytyczna tego wskaźnika powinna być stała w czasie.

W odniesieniu do wskaźnika zadłużenia kapitału własnego, można przyjąć iż sytuacja gospodarcza potwierdza ostrożne podejście do tego tematu i nie ma powodów by modyfikować wartość krytyczną tego wskaźnika.

Natomiast wartość krytyczna wskaźnika ROE winna uwzględniać rynkową stopę odsetek i ryzyko inwestycji w przedsiębiorstwo. Stopy procentowe odsetek od lokat wg oprocentowania stałego zgodnie z danymi portalu Money.pl wynoszą:

Tabela 1. Oprocentowanie lokat złotych: oprocentowanie stałe

Nazwa Banku	ROR		1 mies.		3 mies.		6 mies.		12 mies.	
	od	do	od	do	od	od	do	od	do	od
Alior Bank	0,10	-	2,05	2,05	2,10	2,10	-	-	2,20	2,20
Bank BGŻ	-	4,10	0,55	0,95	0,65	1,25	1,15	1,35	1,50	2,10
Bank BPH	0,10	-	-	-	2,00	2,00	2,25	2,25	2,55	2,55
Bank Millennium	-	-	1,20	1,50	1,60	1,90	1,80	2,10	1,80	2,10
Bank Pekao SA	-	-	1,00	1,00	1,00	1,00	1,00	1,00	2,00	2,10
Bank Pocztowy	-	-	-	-	1,75	1,75	-	-	2,25	2,25
BGK	-	-	0,20	0,20	0,20	0,20	0,20	0,20	0,20	0,20
BNP Paribas Bank	-	0,80	2,10	2,20	2,20	2,30	2,40	2,50	2,50	2,60
BOŚ Bank	0,01	7,00	1,95	1,95	2,05	3,10	2,45	3,00	2,30	2,90
BRE Bank Hipoteczny	-	-	1,97	1,97	2,08	2,08	2,21	2,21	-	-
BZ WBK	-	1,25	1,00	1,00	1,25	1,25	1,25	1,25	1,25	1,25
Citi Handlowy	-	-	1,00	1,50	1,75	2,00	1,75	2,00	1,75	2,10
Credit Agricole	-	-	1,95	1,95	2,25	2,25	2,60	2,60	3,00	3,00
Deutsche Bank Polska	-	-	1,00	1,00	1,25	1,25	1,50	1,50	1,75	1,75
DnB NORD	-	0,10	2,30	2,50	2,60	3,20	2,60	2,90	3,00	3,00
Eurobank	-	2,00	2,40	2,65	2,75	3,00	2,80	3,05	3,00	3,25
FM Bank	-	2,00	1,00	3,00	2,00	3,30	2,00	3,50	3,00	3,50
Getin Noble Bank	-	-	1,00	1,00	1,50	1,50	1,75	1,75	2,20	2,20
ING Bank Śląski	-	-	-	-	-	-	1,50	1,50	2,00	2,00
Inteligo	-	-	2,00	2,00	2,10	2,10	-	-	-	-
Kredyt Bank SA	-	5,10	-	-	1,25	1,75	1,25	1,75	1,50	2,00
mBank	-	4,00	-	-	2,00	2,00	2,00	2,00	2,00	2,00
Meritum Bank	-	0,01	3,30	3,30	3,50	3,50	3,50	3,50	3,50	3,50
Noble Bank	-	-	1,00	1,05	1,50	1,55	1,75	1,80	2,20	2,25

Nazwa Banku	ROR		1 mies.		3 mies.		6 mies.		12 mies.	
	od	do	od	do	od	od	do	od	do	od
Nordea Bank	0,01	0,01	2,20	2,30	2,30	2,40	2,40	2,50	-	-
PBP	-	-	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00
PKO BP	0,01	0,01	0,20	0,40	0,40	0,60	0,50	0,70	-	-
PLUS BANK	0,10	4,10	-	-	2,65	2,80	2,60	2,80	2,45	2,65
Raiffeisen Polbank	-	-	1,20	1,90	2,15	2,30	2,00	2,15	2,00	2,15
Santander Consumer Bank	-	-	2,00	2,20	2,45	2,50	3,25	3,25	3,00	3,00
SKOK Kopernik	-	-	-	-	3,30	3,30	4,00	4,00	4,00	4,00
SKOK Wspólnota	-	-	-	-	3,20	3,30	3,35	3,45	3,65	3,75
Toyota Bank	-	-	1,35	1,35	2,75	2,75	2,95	2,95	3,00	3,00
VW Bank direct	0,10	0,10	2,55	2,80	2,65	2,90	-	-	-	-

Źródło: portal internetowy www.money.pl dostęp 25.04.2014r. „-”, oznacza brak oferty

Natomiast przyjmując oprocentowanie zmienne, oferowane odsetki od lokat wynoszą:

Tabela 2. Oprocentowanie lokat złotych: oprocentowanie zmienne

Nazwa Banku	1 mies.		3 mies.		6 mies.		12 mies.		24 mies.	
	od	do	od	do	od	od	Do	od	do	od
Alior Bank	-	-	-	-	-	-	-	-	2,95	2,95
BGK	-	-	0,20	0,20	0,20	0,20	0,20	0,20	0,20	0,20
BNP Paribas Bank	2,10	2,20	2,25	2,35	2,55	2,65	2,95	3,00	-	-
BOŚ Bank	-	-	1,15	1,15	1,75	1,75	2,00	2,00	-	-
Credit Agricole	2,30	2,30	2,90	2,90	3,00	3,00	3,20	3,20	3,20	3,20
Eurobank	1,80	1,85	2,25	2,30	2,45	2,50	2,80	2,85	3,65	3,85
Inteligo	-	-	-	-	2,15	2,15	2,20	2,20	-	-
Nordea Bank	2,10	2,10	2,30	2,30	2,40	2,40	2,50	2,50	2,70	2,70
PKO BP	-	-	-	-	0,50	1,10	1,00	1,60	1,10	1,70
PLUS BANK	-	-	2,25	2,65	2,35	2,75	2,45	2,65	-	-
SK Bank	3,00	3,00	3,40	3,40	4,00	4,00	4,60	4,60	6,00	6,00
SKOK Kopernik	3,00	3,10	4,10	4,20	3,40	3,50	4,50	4,60	3,80	3,90
SKOK Wspólnota	-	-	3,40	3,77	3,57	3,87	3,77	4,07	-	-
VW Bank direct	-	-	-	-	2,85	3,10	3,10	3,10	-	-

Źródło: portal internetowy www.money.pl dostęp 25.04.2014r. „-”, oznacza brak oferty

Średnia arytmetyczna stóp procentowych oferowanych przez banki, dla lokat wynosi 2,20%. W takiej sytuacji przyjęcie wielkości krytycznej wskaźnika ROE na poziomie 5% nie wydaje się właściwe. Jest to wartość zbyt wysoka. W tym stanie zasadnym jest przyjąć wartość krytyczną 3%, co odzwierciedla średnią rynkową stopę procentową w wysokości 2,20% i premię za ryzyko w wysokości 0,8%

Wartość krytyczna ROE determinuje przyjęcie krytycznej wielkości ROA na poziomie 1,8%.

ANALIZA EMPIRYCZNA WRAŻLIWOŚCI MIARY m NA WIELKOŚCI KRYTYCZNE WSKAŹNIKÓW DIAGNOSTYCZNYCH

Miara sytuacji finansowej m obliczona przy zastosowaniu pierwotnych wartości krytycznych wskaźników diagnostycznych i po aktualizacji wartości krytycznych.

Analiza empiryczna zostanie przeprowadzona z wykorzystaniem danych finansowych spółek wymienionych w tab. 3. Są to przedsiębiorstwa przemysłu spożywczego notowane na warszawskiej giełdzie papierów wartościowych. Na podstawie danych ze sprawozdań finansowych tych przedsiębiorstw obliczono miarę m w wariantach wa i wb. Wariant wa oznacza obliczanie m przy zastosowaniu pierwotnych wartości krytycznych wskaźników diagnostycznych, natomiast wb oznacza obliczenie przyjmując wartość krytyczną ROE 3%, natomiast ROA 1,8%.

Tabela 3. Miara syntetyczna sytuacji finansowej m wyszczególnionych przedsiębiorstw w dwóch wariantach

l.p.	przedsiębiorstwo	2013 r.		
		wa	wb	wb-wa
1	Wawel S.A.	2,40	4,00	1,60
2	Graal S.A.	-0,07	0,22	0,29
3	Indykpol S.A.	-0,55	-0,47	0,08
4	Mieszko S.A.	-0,10	0,32	0,42
5	Colian S.A.	0,13	0,41	0,28
6	Duda S.A.	0,10	0,50	0,40
7	Kania S.A.	0,19	0,73	0,54
8	Kofola S.A.	0,13	0,66	0,53
9	Kruszwica S.A.	1,21	2,24	1,03
10	Makarony Polskie S.A.	-0,15	0,14	0,29
11	Otmuchów S.A.	-0,39	-0,31	0,08
12	PBS Finanse S.A.	-0,54	-1,06	-0,52
13	Pepoes S. A.	-0,31	-0,25	0,06
14	SEKO S.A.	0,61	1,23	0,62
15	Wilbo S.A.	-3,29	-4,56	-1,27
	średnia arytmetyczna m	-0,04	0,25	0,30

Źródło: obliczenia własne na podstawie sprawozdań finansowych

Analizując dane zamieszczone w tabeli 3 należy powiedzieć iż zmiana wielkości krytycznej części wskaźników diagnostycznych służących do obliczenia miary syntetycznej m , ma fundamentalny wpływ na obliczoną wielkość tej miary. W przypadku 6 (czyli 40%) przedsiębiorstw, miara syntetyczna obliczona według pierwotnego wariantu wskazywałaby na niekorzystną sytuację finansową, gdy tymczasem jeśli do jej obliczenia przyjmie się właściwe wielkości krytyczne ROE i ROA to okaże się że sytuacja finansowa tych przedsiębiorstw jest dobra.

Jeśli spojrzeć na przedsiębiorstwa, w przypadku których miara m przyjęła dodatnią wartość nawet w wariacie wa, to należy powiedzieć iż wielkość tej miary obliczona w prawidłowym wariacie wb, jest od 67% do 408% większa niż obliczona w wariacie wa. Średnia arytmetyczna miary m dla wszystkich analizowanych przedsiębiorstw odzwierciedla sytuację finansową całej grupy. Należy podkreślić, iż jeśli oblicza się m zgodnie z pierwotnym wariantem, to sytuację finansową całej grupy analizowanych przedsiębiorstw należałoby w 2013 roku uznać za niekorzystną, natomiast jeśli miarę m oblicza się z wykorzystaniem prawidłowo ustalonych wielkości krytycznych wskaźników diagnostycznych, to widać iż sytuacja finansowa analizowanej grupy przedsiębiorstw była na koniec 2013 roku dobra.

Powyższe udowadnia iż miara m jest wrażliwa na właściwy dobór wielkości neutralnych wskaźników diagnostycznych, służących do jej budowy, czyli jej wartość w dużym stopniu zależy od przyjętych wielkości krytycznych. Nie można więc raz na zawsze przyjąć wielkości krytycznych diagnostycznych wskaźników sytuacji finansowej przedsiębiorstwa, a należy je stale monitorować, szczególnie w dłuższym horyzoncie czasowym.

Wnioski wyciągnięte na podstawie 2013 roku, można jeszcze zweryfikować na podstawie 2012 roku. Z tym że według orientacji autora, wielkość krytyczna ROE powinna za 2012 rok wynosić 4%, natomiast ROA 2,25%. Wartość m wyniesie wtedy za 2012 rok w wariantach wa i wb:

Tabela 4. Miara syntetyczna sytuacji finansowej m wyszczególnionych przedsiębiorstw w dwóch wariantach

l.p.	przedsiębiorstwo	2012 r.		
		wa	wb	wb-wa
1	Wawel S.A.	2,26	2,95	0,69
2	Graal S.A.	0,23	0,43	0,20
3	Indykpol S.A.	-0,89	-0,93	-0,04
4	Mieszko S.A.	-0,53	-0,47	0,06
5	Colian S.A.	0,61	0,86	0,25
6	Duda S.A.	-1,08	-1,24	-0,16
7	Kania S.A.	-0,27	-0,16	0,11
8	Kofola S.A.	-0,16	-0,03	0,13
9	Kruszwica S.A.	-0,53	-0,51	0,02

l.p.	przedsiębiorstwo	2012 r.		
		wa	wb	wb-wa
10	Makarony Polskie S.A.	-0,46	-0,41	0,05
11	Otmuchów S.A.	-0,09	0,04	0,13
12	PBS Finanse S.A.	0,53	0,29	-0,24
13	Pepoes S. A.	-0,27	-0,21	0,06
14	SEKO S.A.	-0,04	0,07	0,11
15	Wilbo S.A.	-38,61	-48,56	-9,95
	średnia arytmetyczna	-2,62	-3,19	-0,57

Źródło: obliczenia własne na podstawie sprawozdań finansowych

Tabela 4 informuje iż w przypadku 2012 roku brak aktualizowania neutralnych wielkości mierników diagnostycznych powoduje zniekształcenie syntetycznej miary sytuacji finansowej przedsiębiorstwa m , ale już nie w takim stopniu jak w przypadku 2013 roku. W wypadku dwóch (13,33%) przedsiębiorstw zniekształcona miara pokazuje złą sytuację finansową, zamiast dobrej, podczas gdy w 2013 roku było tak w przypadku 40% przedsiębiorstw. Również zniekształcona jest średnia arytmetyczna syntetycznej miary wszystkich przedsiębiorstw, obrazująca sytuację całej grupy, jednak zniekształcenie nie jest już tak istotne jak w przypadku 2013 roku. Dzieje się tak dlatego iż zmiana wielkości krytycznych ROE i ROA jest znacznie wyraźniejsza w 2013 roku niż w 2012.

Analiza dwóch lat obrotowych pokazuje iż stale należy monitorować prawidłowość ustalenia wartości krytycznych mierników diagnostycznych. Jeśli różnica pomiędzy pierwotnie ustalonymi, a wynikającymi z obecnej sytuacji gospodarczej jest nieistotna, można obliczać miarę m według wzoru pierwotnego. Przy istotnej zmianie sytuacji gospodarczej, należy miarę m obliczać według zaktualizowanego wzoru.

KONKLUZJE

Przeprowadzone badanie wykazało iż syntetyczna miara sytuacji finansowej przedsiębiorstwa obliczana zgodnie z wzorem (1) lub (2) nie jest wrażliwa na krytyczne wielkości wskaźników diagnostycznych służących do jej budowy, gdy nie zmienia się istotnie sytuacja gospodarcza w otoczeniu przedsiębiorstwa.

Natomiast w razie istotnej zmiany w otoczeniu gospodarczym przedsiębiorstwa, miara diagnostyczna m jest bardzo wrażliwa na prawidłowość ustalenia wielkości neutralnych, wskaźników służących do jej konstrukcji.

Dlatego też konieczne jest monitorowanie otoczenia gospodarczego przedsiębiorstwa i w razie istotnych zmian w tym otoczeniu, aktualizacja wzoru służącego do obliczenia miary syntetycznej m .

BIBLIOGRAFIA

- Jerzmanowska M. (red) (2004) Analiza ekonomiczna w przedsiębiorstwie, (2004) Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Lisek S. (2009) Mierniki Syntetyczne w procesie diagnozowania sytuacji finansowej przedsiębiorstwa, Zeszyty Naukowe Uniwersytetu Szczecińskiego nr 542.
- Wersty B. (red) (1996) Analiza ekonomiczna przedsiębiorstwa, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu.

**SENSITIVITY OF SYNTNETHICS MEASURE M ON THE NEUTRAL
VALUES OF THE DIAGNOSTIC INDICATORS
USED IN CREATION MEASURE M**

Abstract: Synthetic measure of the enterprises financial situation is useful tool. In this article is mentioned research results, how neutral values of the diagnostic ratios used I build measure m, effects information, which gives measure m. When changes in the economic situation of the environment aren't significant, one needn't change m measure's formula. When this changes are significant, one have to actualize m measure's formula.

Keywords: financial situation of the enterprises, multivariate comparative analysis, synthetic measure, diagnostic ratios, neutral values.

OCENA POZIOMU ŻYCIA W KRAJACH UNII EUROPEJSKIEJ (UE-28) W ASPEKTCIE ZRÓWNOWAŻONEGO ROZWOJU ZA POMOCA WIELOWYMIAROWEJ ANALIZY PORÓWNAWCZEJ

Anna Murawska

Katedra Ekonomii i Prawa Gospodarczego
Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy
e-mail: anna.murawska@utp.edu.pl

Streszczenie: Poziom i warunki życia w krajach Unii Europejskiej są silnie przestrzennie zróżnicowane. Od początku XXI wieku obserwuje się rozwój działań na rzecz zrównoważonego rozwoju, którego celem jest zapewnienie przyszłym pokoleniom stałego wzrostu poziomu życia i spójności społecznej. Dlatego celem artykułu było porównanie poziomu życia w krajach Unii Europejskiej za pomocą tradycyjnych, nowych oraz syntetycznych wskaźników, a także utworzono grupy krajów o podobnym poziomie życia.

Słowa kluczowe: poziom życia, kraje, Unia Europejska, zrównoważony rozwój, metody wielowymiarowe, wskaźnik

WPROWADZENIE

Poziom życia jest pojęciem złożonym, uwarunkowanym przez szereg różnych cech. Aby zrozumieć zakres, jaki obejmuje jego definicja, należy wrócić do tej sformułowanej przez komisję ekspertów ONZ w 1954 roku, która otrzymała następujące brzmienie: „Pojęcie poziomu życia obejmuje całokształt rzeczywistych warunków życia ludzi oraz stopień ich materialnego i kulturalnego zaspokojenia potrzeb poprzez strumień dóbr i usług odpłatnych, a także pochodzących z funduszy społecznych”¹. Definicja ta stała się podstawą dla wielu innych określeń tego pojęcia [Drewnowski 1974, Żekoński 1974, Luszczewicz 1982, Piasny 1993, Bywalec i Rudnicki 1992, Słaby 2004]. Luszczewicz [1982] wyróżnił siedem podstawowych rodzajów potrzeb: bezpieczeństwo, wyżywienie, ochrona

¹ Rapport sur la definition et evaluation des niveaux de vie du point de vue international, 1954, New York, Nations Unies, s. 5.

zdrowia, warunki mieszkaniowe, komunikacja i transport, edukacja i kultura oraz ochrona środowiska.

Kluczowym elementem we wszystkich definicjach poziomu życia są potrzeby konsumpcyjne oraz ich „zrównoważone” zaspokajanie, co stanowi podstawowy cel działalności ekonomicznej każdego gospodarstwa domowego w skali mikroekonomicznej – lokalnej, oraz podstawowy cel działalności ekonomicznej każdego państwa w skali makroekonomicznej - globalnej.

Zagadnienia dotyczące poziomu życia i spójności społecznej są podstawowymi składnikami wieloaspektowych działań na rzecz zrównoważonego rozwoju. Naczelnym celem strategii zrównoważonego rozwoju Unii Europejskiej jest „określenie i rozwój działań, dzięki którym UE będzie mogła zapewnić obecnym i przyszłym pokoleniom stały wzrost jakości życia przez tworzenie społeczności wydajnie gospodarujących zasobami i z nich korzystających (...), a przez to zapewniających dobrobyt, ochronę środowiska naturalnego i spójność społeczną” [Michaliszyn 2012, Wskaźniki zrównoważonego... 2011].

Po zakończeniu Strategii Lizbońskiej, 17 czerwca 2010 r. Rada Europejska przyjęła Strategię Europa 2020. Europa 2020 jest to program rozwoju społeczno-gospodarczego UE na lata 2010-2020, który ma przyczynić się do wzrostu zatrudnienia, poprawy produktywności oraz spójności społecznej². W strategii określono pięć nadrzędnych celów:

1. Zwiększenie zatrudnienia osób w wieku 20-64 lata do 75%;
2. Zwiększenie nakładów na działalność badawczą i rozwojową do 3% PKB;
3. Redukcja emisji gazów cieplarnianych o 20% względem emisji z 1990 r. (...);
4. Liczba osób niekontynuujących nauki powinna być niższa niż 10% oraz co najmniej 40% ludności powinno mieć wykształcenie wyższe;
5. Redukcja ubóstwa poprzez zmniejszenie liczby ludności zagrożonej ubóstwem lub wykluczeniem społecznym o 20 mln.

Do monitorowania stopnia realizacji przyjętych celów ma służyć obecnie osiem wskaźników³, które zostały uwzględnione w wielowymiarowej analizie poziomu życia w krajach UE-28 w niniejszym artykule.

Dlatego, że poziom życia jest pojęciem wielowymiarowym powinno analizować się go za pomocą szeregu wskaźników jednocześnie i budować syntetyczne mierniki. Jednak z uwagi na dużą czasochłonność takich analiz często stosuje się inne pojedyncze wskaźniki lub mierniki charakteryzujące poziom życia [Murawska 2009]. Przykładowo tradycyjnym i powszechnie stosowanym wskaźnikiem jest produkt krajowy brutto - PKB w przeliczeniu na jednego mieszkańca (ang. *Gross Domestic Product - GDP per capita*) [Gutkowska,

² „Europa 2020 Strategia na rzecz inteligentnego i zrównoważonego rozwoju sprzyjającego włączeniu społecznemu”, komunikat Komisji Europejskiej, Bruksela 3.3.2010, KOM (2010) 2020 wersja ostateczna; http://ec.europa.eu/eu2020/pdf/1_PL_ACT_part1_v1.pdf.

³ Wskaźniki Strategii Europa 2020; http://stat.gov.pl/cps/rde/xbcr/gus/POZ_Wskaźniki_Europa2020.pdf [2.07.2014 r.]

Murawska 2011]. Z kolei najnowszym alternatywnym miernikiem poziomu życia jest faktyczne spożycie indywidualne - FSI przypadające na jednego członka gospodarstwa domowego (ang. *Actual Individual Consumption - AIC per capita*)⁴. Wskaźnik ten został rekomendowany w raporcie opracowanym w 2009 r.⁵, gdzie podważano słuszność stosowania wskaźnika PKB do oceny poziomu życia [Murawska 2013].

Ważność zagadnień dotyczących poziomu życia ludności i jego regionalnego zróżnicowania dostrzeżono i podkreślano już w połowie ubiegłego wieku. Przez ostatnich kilkadziesiąt lat podejmowany był szereg działań w celu zniwelowania regionalnych różnic w poziomie życia zarówno na poziomie lokalnym i krajowym, jednak pomimo tego problem z tym związany nadal istnieje.

CEL, MATERIAŁ I METODY

Celem artykułu była analiza i diagnoza regionalnego zróżnicowania poziomu życia w krajach Unii Europejskiej (UE-28) na podstawie syntetycznej miary, ze szczególnym uwzględnieniem wskaźników charakteryzujących zrównoważony rozwój. Dodatkowo dokonano porządkowania liniowego krajów i powstały ranking krajów na podstawie syntetycznego miernika (SMPŻ) porównano z rankingiem krajów powstałym na podstawie tradycyjnej miary poziomu życia, czyli *GDP per capita (Gross Domestic Product per capita)* oraz nowej miary poziomu życia *AIC per capita (Actual Individual Consumption per capita)*.

Przesłanką do realizacji tego opracowania była akcentowana w ostatnich latach konieczność szukania i uwzględniania nowych wskaźników takich, które w pełni będą odzwierciedlały wszystkie obszary związane z realizacją nowo przyjętej strategii UE Europa 2020. Cel główny został zrealizowany poprzez określenie zmiennych, wyznaczenie cech diagnostycznych i syntetycznego miernika (SMPŻ), sporządzenie rankingu krajów oraz podzielenie ich na grupy, a także dokonanie analizy porównawczej syntetycznego miernika z tradycyjną i nową miarą poziomu życia.

Podstawowym źródłem danych były zbiory gromadzone przez Europejski Urząd Statystyczny Eurostat oraz Główny Urząd Statystyczny w Polsce. Ze względu na trudności z pozyskaniem zakładanego na etapie projektowania badania zbioru informacji statystycznych analizy przestrzennego zróżnicowania poziomu życia w krajach Unii Europejskiej dokonano w kilku etapach.

⁴ GDP per capita varied by one to six across the Member States in 2011. Eurostat newsrelease, 180/2012 -13 December 2012.

⁵ J. E. Stiglitz, A. Sen, J. Fitoussi, 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress, http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.47.

Na pierwszym etapie zebrano dane statystyczne dotyczące siedmiu obszarów badawczych: 1. Bezpieczeństwo, zatrudnienie i spójność społeczna (10 wskaźników), 2. Wyżywienie (2), 3. Warunki mieszkaniowe (4), 4. Ochrona zdrowia (6), 5. Ochrona środowiska (6), 6. Edukacja i kultura (5), 7. Komunikacja i transport, społeczeństwo informacyjne (7). Łącznie utworzona baza danych składała się z 40 zmiennych. Zebrane informacje dotyczyły lat 2000-2012. Po zgromadzeniu danych zaobserwowano szereg brakujących informacji dotyczących wielu krajów członkowskich, dlatego do szczegółowych analiz nie można było uwzględnić wszystkich lat. Ostatecznie w badaniu wykorzystano dane z 2011 roku.

Na drugim etapie badań utworzona baza danych liczbowych została poddana selekcji ze względu na kryteria merytoryczno-formalne własności zmiennych oraz kryteria wartości informacyjnej zmiennych [Zeliaś, 2000, s. 36-37] i dzięki temu dokonano analizy i doboru zmiennych diagnostycznych. Jako, że wśród zgromadzonych 40 wskaźników dla 2011 r. występowały brakujące dane dla niektórych krajów pominięto je podczas dalszych analiz. Następnie ze zbioru potencjalnych cech diagnostycznych usunięto zmienne odznaczające się zbyt niskim stopniem zmienności ($V_s < 10\%$) [Nowak 1997, s. 12] oraz nadmiernie skorelowane ze sobą [Hellwig 1981, s. 57, Nowak 1984, s. 127].

Redukcji zbioru zmiennych diagnostycznych dokonano za pomocą metody Hellwiga. Na podstawie macierzy korelacji obliczono wartość progową współczynnika r^* zgodnie z regułą zaproponowaną przez E. Nowaka:

$$r^* = r_{02} - \lambda(r_{02} - r_{01}), \quad (1)$$

gdzie:

$$r_{01} = \min_i \min_j |r_{ij}|, \quad r_{02} = \max_i \max_j |r_{ij}|, \quad (2)$$

natomiast λ mieści się w przedziale $0 < \lambda < 1$ i jest liczbą obraną przez badacza (przyjęto, że $\lambda=0,5$). Zmienne, dla których współczynnik korelacji co do wartości bezwzględnej był wyższy niż wartość krytyczna, wyeliminowano ze zbioru zmiennych (były to tzw. zmienne satelitarne). Z kolei zmienne centralne i izolowane, czyli te, pomiędzy którymi współczynnik korelacji nie przekraczał przyjętej wartości progowej r^* utworzyły ostateczny zbiór cech diagnostycznych, który składał się z następujących 18-stu zmiennych⁶:

I. Bezpieczeństwo, zatrudnienie i spójność społeczna

X₁₃ Wskaźnik pogłębionej deprivacji materialnej w %

X₁₆ Stopa bezrobocia długookresowego w %

X₁₈ Wskaźnik zatrudnienia kobiet w wieku 20-64 lata w %

II. Wyżywienie

X₂₁ Udział żywności i napojów bezalkoholowych w wydatkach ogółem w %

III. Warunki mieszkaniowe

X₃₁ Średnia liczba pokoi na 1 osobę

⁶ Zmienne X₁₃, X₅₂, X₆₃, X₆₅, X₇₃ należą do grupy ośmiu wskaźników służących do monitorowania stopnia realizacji przyjętych celów Strategii Europa 2020.

- X₃₂ Udział wydatków na użytkowanie i wyposażenie mieszkania w wydatkach ogółem w %
 X₃₄ Odsetek ludności żyjącej w gospodarstwach domowych, w których całkowite koszty mieszkaniowe stanowią ponad 40% całkowitego rozporzadzalnego dochodu gospodarstwa domowego

IV. Ochrona zdrowia

- X₄₄ Liczba lekarzy na 100 tys. mieszkańców
 X₄₅ Liczba łóżek szpitalnych na 100 tys. ludności

V. Ochrona środowiska

- X₅₂ Udział energii ze źródeł odnawialnych w końcowym zużyciu energii brutto w %

VI. Edukacja i kultura

- X₆₁ Studenci szkół wyższych na 10 tys. ludności
 X₆₂ Udział wydatków na zaspokojenie potrzeb z zakresu rekreacji i kultury w wydatkach ogółem w %
 X₆₃ Osoby w wieku 30-34 lata posiadające wykształcenie wyższe w % osób w wieku 30-34 lata
 X₆₅ Młodzież niekontynuująca nauki w % osób w wieku 18-24 lata

VII. Komunikacja i transport, społeczeństwo informacyjne

- X₇₃ Gospodarstwa domowe z dostępem do Internetu w % ogółu gospodarstw domowych
 X₇₄ Liczba abonentów telefonii ruchomej (komórkowej) na 1000 ludności
 X₇₅ Liczba abonentów telefonii stacjonarnej na 1000 ludności
 X₇₆ Samochody osobowe w użytkowaniu na 1000 osób w sztukach.

Wszystkie wskaźniki po umieszczeniu w bazie danych zostały poddane analizie statystycznej. Dla każdej zmiennej wyznaczono miary położenia, miary zmienności, obliczono współczynniki asymetrii oraz przeprowadzono weryfikację hipotezy o normalności rozkładu zmiennych za pomocą testu Shapiro-Wilka. W tabeli 1 przedstawiono wybrane obliczenia statystyczne dla wyselekcjonowanych cech diagnostycznych.

Budowa syntetycznej miary poziomu życia wymaga podziału zbioru zmiennych diagnostycznych na stymulanty i destymulanty. Do zbioru stymulant zostały zaliczone następujące zmienne: $S: \{X_{18}, X_{31}, X_{32}, X_{44}, X_{45}, X_{52}, X_{61}, X_{62}, X_{63}, X_{73}, X_{74}, X_{75}, X_{76}\}$, natomiast zbiór destymulant był następujący: $D: \{X_{13}, X_{16}, X_{21}, X_{34}, X_{65}\}$. Przekształcenia destymulant w stymulanty dokonano za pomocą wzoru:

$$x_{ij} := c_j - x'_{ij}, (i = 1, \dots, n; j = 1, \dots, k), \quad (3)$$

przy czym x_{ij} ($i=1, \dots, n; j=1, \dots, k$) to wartość j -tej zmiennej w i -tym obiekcie wielowymiarowym Q ; x'_{ij} są realizacjami zmiennej destymulandy, gdzie c_j oznacza pewną stałą, przy czym „:=” oznacza podstawienie [Zeliaś 2000].

W dalszym etapie analiz zmienne potraktowano jako równoprawne, przyjmując system wag jednostkowych. Zredukowany, ale zawierający najważniejsze elementy, zestaw danych poddano standaryzacji, zgodnie ze wzorem:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, \quad \text{gdzie: } \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}, \quad s_j = \left[\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \right]^{0.5}, \quad (4)$$

gdzie z_{ij} to zestandaryzowana wartość zmiennej diagnostycznej x_i .

Tabela 1. Statystyczna charakterystyka zmiennych opisujących poziom życia w krajach UE (28) w 2011 r.

Zmienna	$x_{\bar{s}r}$	V_s	x_{min}	x_{max}	A_s
X ₁₃	11,0	92,6	1,2 (Luksemburg, Szwecja)	43,6 (Bułgaria)	1,79
X ₁₆	4,6	60,9	1,1 (Austria)	9,2 (Słowacja)	0,45
X ₁₈	62,1	13,1	43,4 (Malta)	76,5 (Szwecja)	-0,44
X ₂₁	15,5	32,4	9,2 (Luksemburg)	29,1 (Rumunia)	1,08
X ₃₁	1,6	24,6	1,0 (Polska, Rumunia)	2,3 (Belgia)	-0,20
X ₃₂	22,4	17,8	12,5 (Malta)	29,8 (Dania)	-0,74
X ₃₄	9,7	52,7	3,0 (Malta)	24,2 (Grecja)	1,05
X ₄₄	337,9	24,2	218,6 (Polska)	611,4 (Grecja)	1,46
X ₄₅	525,5	30,6	270,6 (Szwecja)	822,2 (Niemcy)	-0,07
X ₅₂	15,5	70,9	0,4 (Malta)	46,8 (Szwecja)	1,01
X ₆₁	445,8	24,5	111,0 (Luksemburg)	646,0 (Litwa)	-0,71
X ₆₂	8,5	22,9	4,9 (Rumunia)	11,5 (Finlandia)	-0,08
X ₆₃	35,3	28,0	20,3 (Włochy)	49,7 (Irlandia)	-0,16
X ₆₅	11,5	51,1	4,1 (Chorwacja)	26,5 (Hiszpania)	1,07
X ₇₃	66,8	21,6	33,0 (Bułgaria)	91,0 (Holandia)	-0,23
X ₇₄	1241,5	15,9	941,0 (Francja)	1659,0 (Finlandia)	0,77
X ₇₅	388,4	33,6	179,0 (Polska)	630,0 (Francja)	-0,03
X ₇₆	454,4	23,1	201,0 (Rumunia)	666,0 (Luksemburg)	-0,45

Źródło: obliczenia własne

V_s – współczynnik zmienności w %, A_s – współczynnik asymetrii.

Następnie ustalono wzorec rozwoju, którego współrzędne określone były przez największe, maksymalne wartości:

$$z_{oj} := \max_i z_{ij}. \quad (5)$$

W stosunku do otrzymanego wzorca dla każdego badanego kraju obliczono wielowymiarowe dystanse, stosując metrykę Euklidesa, którą zapisuje się wzorem:

$$d_{io} = \left[\sum_{j=1}^k (z_{ij} - z_{oj})^2 \right]^{0,5} \quad (6)$$

gdzie d_{io} – odległość (dystans) obiektu Q_i ($i=1, \dots, n$) od hipotetycznego (abstrakcyjnego) obiektu wzorcowego Q_o .

W celu unormowania zmiennej syntetycznej d_{io} , a także w celu otrzymania takiej miary, której rosnące wartości świadczyłyby o rozwoju badanego zjawiska, skonstruowano tzw. względne zmienne syntetyczne (syntetyczne mierniki) o postaci:

$$z_i := 1 - \frac{d_{io}}{d_o}, \quad \text{gdzie: } d_o = \bar{d}_o + 3S_o, \quad (7)$$

przy czym:

$$\bar{d}_o = \frac{1}{n} \sum_{i=1}^n d_{io}, \quad S_o = \left[\frac{1}{n} \sum_{i=1}^n (d_{io} - \bar{d}_o)^2 \right]^{0.5}. \quad (8)$$

Utworzone zmienne syntetyczne z prawdopodobieństwem bliskim jedności przyjęły wartości z przedziału [0,1]. Skonstruowana taksonomiczna miara rozwoju przyjmuje wartości tym wyższe, im bardziej rozwinięty jest dany kraj [Zeliaś 2000].

Na podstawie syntetycznych mierników obliczonych dla wszystkich krajów sporządzono ranking. Z kolei w celu wyodrębnienia grup typologicznych składających się z krajów o podobnym poziomie życia oparto się na analizie różnic w poziomie wartości zmiennej syntetycznej, zgodnie z regułą zaproponowaną przez E. Nowaka [1990]. Podstawą uzyskania klas jednostek przestrzennych są przedziały wartości zmiennej syntetycznej zbudowane opierając się na średniej arytmetycznej \bar{z} i odchyleniu standardowym S_z . Zbiór obiektów jest dzielony na cztery grupy, obejmujące obiekty o wartościach zmiennej syntetycznej z następujących rozłącznych przedziałów (grupy te spełniają warunek rozłączności i zupełności):

$$\text{grupa I: } z_i \geq \bar{z} + s_z, \quad (9)$$

$$\text{grupa II: } \bar{z} + s_z > z_i \geq \bar{z}, \quad (10)$$

$$\text{grupa III: } \bar{z} > z_i \geq \bar{z} - s_z, \quad (11)$$

$$\text{grupa IV: } z_i < \bar{z} - s_z, \quad (12)$$

$$\text{gdzie: } \bar{z} = \frac{1}{n} \sum_{i=1}^n z_i, \quad s_z = \left[\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 \right]^{0.5}.$$

Uzyskanie jednorodnych i rozłącznych grup najbardziej do siebie podobnych obiektów ułatwia wyodrębnienie i analizę merytoryczną zasadniczych ich cech. Obiekty należące do danej grupy typologicznej są uporządkowane według wartości miernika syntetycznego. Możliwe jest także natychmiastowe porównanie różnych grup typologicznych pod względem poziomu życia.

WYNIKI KLASYFIKACJI

Obliczone wartości zmiennych syntetycznych opisujących poziom życia pozwoliły na sporządzenie rankingu krajów od najlepszego do najłabszego pod względem analizowanego zjawiska, a także utworzenie grup krajów o wysokim, średnim, niskim i bardzo niskim poziomie życia (tab. 2).

W świetle przedstawionych wyników badań stwierdzono, że zdecydowanie najwyższy poziom życia występuje w Szwecji. Wartość syntetycznego miernika obliczona dla tego kraju wyniosła $z_i=0,394$. Pomimo tego, że Szwecja znalazła się na pierwszym miejscu w rankingu i tak odbiega od wzorca, czyli jednostki abstrakcyjnej, której współrzędne określone są przez maksymalne wartości.

Wynika to z uwzględnienia do budowy syntetycznego miernika aż 18 cech, które znacznie różnią się w poszczególnych krajach.

Tabela 2. Ranking krajów pod względem poziomu życia w 2011 r. na podstawie trzech wskaźników – syntetycznego miernika oraz GDP i AIC *per capita*

Kraj	d_{io}	z_i	Grupy: poziom życia	SMPŻ	GDP <i>per capita</i>	AIC <i>per capita</i>
				Miejsce w rankingu		
Szwecja	11,22	0,394	wysoki	1	5	5
Finlandia	11,36	0,387		2	9	9
Austria	11,53	0,378		3	4	3
Holandia	12,33	0,334		4	2	6
Dania	13,08	0,294	średni	5	6	7
Francja	13,19	0,288		6	11	8
Niemcy	13,27	0,284		7	7	2
Belgia	13,60	0,265		8	8	10
Luksemburg	13,62	0,265		9	1	1
Słowenia	13,85	0,252		10	16	17
Czechy	13,89	0,250		11	17	19
Estonia	14,16	0,235		12	21	25
Wielka Brytania	14,21	0,233		13	10	4
Irlandia	14,49	0,218		14	3	11
Cypr	14,68	0,208		15	14	13
Litwa	14,69	0,207	niski	16	22	21
Łotwa	15,50	0,163		17	26	26
Polska	15,58	0,159		18	24	22
Portugalia	15,68	0,153		19	19	18
Słowacja	15,74	0,150		20	20	20
Włochy	15,77	0,148		21	12	12
Hiszpania	16,32	0,119		22	13	14
Węgry	16,54	0,107		23	23	23
Malta	16,66	0,101	bardzo niski	24	15	16
Chorwacja	16,83	0,091		25	25	24
Grecja	17,10	0,077		26	18	15
Bułgaria	17,53	0,054		27	28	28
Rumunia	18,74	-0,012		28	27	27

Źródło: obliczenia własne

Kolejne w rankingu znalazły się Finlandia, Austria oraz Holandia. Sytuacja rynku pracy, bezpieczeństwo oraz spójność społeczna w tych krajach jest na najwyższym poziomie w Europie. Są to kraje o najniższym wskaźniku zagrożenia ubóstwem lub wykluczeniem społecznym oraz pogłębionej deprawacji materialnej, najniższa jest stopa bezrobocia, zwłaszcza bezrobocia długookresowego oraz najwyższy wskaźnik zatrudnienia kobiet, a wskaźniki dotyczące warunków mieszkaniowych, ochrony zdrowia, edukacji, kultury czy zaspokojenia potrzeb

w zakresie komunikacji i łączności są najwyższe w Europie. Szwecja jest dodatkowo krajem, który szczególnie dba o ochronę środowiska, dynamicznie redukuje emisję gazów cieplarnianych a także może się pochwalić najwyższym wskaźnikiem w Europie udziału energii ze źródeł odnawialnych w końcowym zużyciu energii brutto.

Na drugim biegunie, na ostatnim 28 miejscu w rankingu znalazła się Rumunia. W kraju tym udział żywności w wydatkach ogółem stanowi 30% (podczas gdy w Luksemburgu niespełna 9,2%). Warunki mieszkaniowe są na najniższym poziomie, świadczy o tym średnia liczba pokoi przypadająca na jedną osobę. Podczas gdy w Belgii jedna osoba w mieszkaniu może dysponować 2,3 pomieszczeniami w mieszkaniu, to w Rumunii zaledwie jednym. Oznacza to, że gospodarstwa domowe w Rumunii są najbardziej przeludnione w całej Europie. Mieszkańcy Rumunii najmniej w całej Unii Europejskiej wydają na zaspokojenie potrzeb z zakresu rekreacji i kultury, co zapewne jest spowodowane niskimi wynagrodzeniami oraz wysokimi wydatkami, które muszą przeznaczać na zaspokojenie potrzeb żywnościowych.

Na szczególną uwagę zasługuje Polska, która w rankingu syntetycznej miary poziomu życia znalazła się na wysokim, 18-stym miejscu. Podczas badania poziomu życia za pomocą wskaźników: *GDP per capita* lub *AIC per capita* Polska znalazła się na niższych pozycjach, odpowiednio na 24 miejscu przed Chorwacją, Łotwą, Rumunią oraz Bułgarią oraz na 22 miejscu dodatkowo jeszcze przed Estonią i Węgrami.

Zbiór wszystkich analizowanych krajów pod względem syntetycznej miary poziomu życia podzielono na 4 grupy. W grupie I, o wysokim poziomie życia znalazły się takie kraje jak: Szwecja, Finlandia, Austria i Holandia; w grupie II, o średnim poziomie życia znalazło się 11 krajów, a mianowicie: Dania, Francja, Niemcy, Belgia, Luksemburg, Słowenia, Czechy, Estonia, Wielka Brytania, Irlandia i Cypr; w grupie III, o niskim poziomie życia zidentyfikowano kraje: Litwa, Łotwa, Polska, Portugalia, Słowacja, Włochy, Hiszpania i Węgry; oraz w IV grupie, o najniższym poziomie życia uwzględniono pięć krajów: Malta, Chorwacja, Grecja, Bułgaria oraz Rumunia.

PODSUMOWANIE

Możliwość rozwoju gospodarczego i społecznego wywarła istotny wpływ na zmiany ekonomiczne i społeczne w krajach europejskich oraz na wzrost poziomu życia i spójności społecznej. Aspekty te mierzone są za pomocą wskaźników takich jak produkt krajowy brutto, faktyczne spożycie indywidualne, a także miar syntetycznych. W krajach Unii Europejskiej istnieją duże regionalne dysproporcje w poziomie życia ludności. Wszystkie kraje znacznie różnią się w zakresie takich aspektów jak: bezpieczeństwo, zatrudnienie i spójność społeczna, warunki mieszkaniowe, ochrona zdrowia, edukacja i kultura, komunikacja i transport, a także ochrona środowiska. Dlatego rankingi krajów pod względem poziomu

życia po uwzględnieniu pojedynczych wskaźników różnią się od rankingów krajów powstałych na podstawie analizy wielowymiarowej. Jednakże niezmiennie na pierwszych miejscach znajdują się kraje takie jak Szwecja, Austria Holandia czy Luksemburg, a na drugim biegunie Rumunia, Bułgaria czy Grecja.

W aspekcie zrównoważonego rozwoju, wyrównanego poziomu życia oraz właściwej realizacji programu strategii Europa 2020, ważne jest, aby ocenić na jakim poziomie zróżnicowań znajdują się kraje europejskie, w jakim obszarze powinno się wyrównywać poziom życia, a w jakim zakresie to jest trudne w realizacji. Przygotowując nowe rozwiązania w celu poprawy życia w Unii Europejskiej i Europie nie można uwzględniać i analizować pojedynczych wskaźników, a należy dokonywać syntez wieloaspektowych uwzględniających wszystkie czynniki sprawcze, bodźce i skutki zmian, jakie zachodzą w krajach europejskich.

BIBLIOGRAFIA

- Bywalec Cz., Rudnicki L (1992) Podstawy teorii i metodyki badań konsumpcji” AE, Kraków.
- Drewnowski J. (1974) On Measuring and Planning the Quality of Life, Institute of Social Studies, Mouton, str. 35.
- Europa 2020 Strategia na rzecz inteligentnego i zrównoważonego rozwoju sprzyjającego włączeniu społecznemu, komunikat Komisji Europejskiej, Bruksela 3.3.2010, KOM (2010) 2020 wersja ostateczna http://ec.europa.eu/eu2020/pdf/1_PL_ACT_part1_v1.pdf.
- GDP per capita varied by one to six across the Member States in 2011. Eurostat newsrelease, 180/2012 -13 December 2012.
- Gutkowska K., Murawska A. (2011) Regionalne zróżnicowanie poziomu i jakości życia Polaków, Handel Wewnętrzny. Konsumpcja a rozwój regionów, IBRKiK, Warszawa, styczeń-luty 2011, str. 91.
- Hellwig Z. (1981) Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych, w: W. Welfe (red.): Metody i modele ekonomiczno-matematyczne w doskonaleniu zarządzania gospodarką socjalistyczną, PWE, Warszawa, str. 57.
- Luszniewicz A. (1982) Statystyka społeczna. Podstawowe problemy i metody, PWE, Warszawa, str. 11.
- Michaliszyn B. (2012) Strategie zrównoważonej konsumpcji w Unii Europejskiej, Wydawnictwo Ekonomia i Środowisko, Białystok, str. 39.
- Murawska A. (2009) Rozwój społeczno-ekonomiczny a poziom życia w gospodarstwach domowych w Polsce, praca doktorska, SGGW, Warszawa (maszynopis niepublikowany), str. 47.
- Murawska A. (2013) Poziom życia w krajach Europy i Unii Europejskiej na podstawie Actual Individual Consumption, Ekonomia i prawo, Tom XII, nr 3/2013, Wydawnictwo Naukowe UMK, Toruń, str. 405.

- Nowak E. (1984) Problemy doboru zmiennych do modelu ekonometrycznego, PWN, Warszawa, 1984, s. 127.
- Nowak E. (1990) Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych, PWE, Warszawa, str. 93.
- Nowak E. (1997) Zarys metod ekonometrii, PWN, Warszawa, str. 12.
- Piasny J. (1993) Poziom i jakość życia ludności oraz źródła i mierniki ich określania, Ruch Prawniczy, Ekonomiczny i Socjologiczny, z. 2, str. 73.
- Rapport sur la definition et evaluation des niveaux de vie du point de vue international, 1954, New York, Nations Unies, str. 5.
- Słaby T. (2004) Poziom i jakość życia (w:) T. Panek i A. Szulc (red.) (2004): Statystyka Społeczna. Wybrane zagadnienia, SGH w Warszawie, Oficyna Wydawnicza, Warszawa, str. 66.
- Stiglitz J. E., Sen A., Fitoussi J., (2009) Report by the Commission on the Measurement of Economic Performance and Social Progress, http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.
- Wskaźniki Strategii Europa 2020; http://stat.gov.pl/cps/rde/xbcr/gus/POZ_Wskazniki_Europa2020.pdf.
- Wskaźniki zrównoważonego rozwoju Polski (2011) GUS, US w Katowicach, Katowice, str. 7.
- Zeliaś A. (red.) (2000) Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym, AE w Krakowie, Kraków, str. 35-101.
- Żekoński Z. (1974) Z problemów metodologicznych sformułowania społeczno-bytowych celów rozwoju, Gospodarka Planowa, nr 6.

**ASSESSMENT OF THE STANDARD OF LIVING
IN THE EUROPEAN UNION (EU-28) IN THE CONTEXT
OF SUSTAINABLE DEVELOPMENT USING MULTIDIMENSIONAL
ANALYSIS OF COMPARATIVE**

Abstract: Level and living conditions in the countries of the European Union are highly spatially variable. Since the beginning of the XXI century has seen the development of sustainable development, which aims to provide future generations a steady increase in living standards and social cohesion. Therefore the aim of the article was to compare the standard of living in the European Union by means of traditional, new and synthetic indicators, and also created a group of countries with a similar standard of living.

Keywords: standard of living, countries, European Union, sustainable development, multivariate methods, indicator

WYKORZYSTANIE WYBRANYCH METOD PORZĄDKOWANIA OBIEKTÓW DO KLASYFIKACJI WOJEWÓDZTW POD KĄTEM ICH POTENCJAŁU INNOWACYJNEGO

Anna M. Olszewska, Alicja E. Gudanowska

Katedra Informatyki Gospodarczej i Logistyki

Politechnika Białostocka

e-mail: a.olszewska@pb.edu.pl, a.gudanowska@pb.edu.pl

Streszczenie: Jednoznaczne określenie definicji oraz determinant innowacyjnego potencjału regionu jest zadaniem trudnym. Mimo to stanowi on jeden z czynników decydujących o rozwoju regionu oraz jego pozycji konkurencyjnej. Istotny w tym kontekście staje się – nawet bardziej niż dostępność danych – problem doboru procedury klasyfikacyjnej. Głównym celem artykułu było dokonanie grupowania województw pod kątem ich potencjału innowacyjnego poprzez zastosowanie różnych metod porządkowania. Dokonano klasyfikacji województw uwypuklając zmiany w otrzymanych wynikach przy uwzględnieniu różnych form porządkowania.

Słowa kluczowe: innowacyjność, metody porządkowania obiektów

WPROWADZENIE

Innowacja jako pierwsza została dostrzeżona i opisana przez J. Schumpetera jako jeden z trzech etapów procesu przemiany: inwencji, innowacji i dyfuzji (*invention-innovation-diffusion*) [Mahdjoubi 1997]. Schumpeter wprowadził tym samym podstawowe ujęcie tematyki innowacji, które chętnie i szeroko opisywane jest przez liczną grupę badaczy [np. Kalinowski 2010; Christensen 2010; Nazarko 2013; Wiśniewska, Janasz 2013]. Powszechnie traktuje się termin ten jako równoznaczny z całym procesem przemiany, w najprostszym ujęciu oznaczającym wykonanie nowego produktu, procesu, usługi bądź też wykonania rzeczy już realizowanych w nowy sposób. Jednak to pierwotne ujęcie innowacji ukształtowało się w gospodarce kapitalistycznej. Obecnie czynnikami o dużym znaczeniu, niezbędnymi w ujęciu istoty innowacji są wiedza i informacja. Stąd też innowacje można utożsamiać również z pomysłem czy nową ideą [Nazarko 2013]. Innowacje

postrzega się już nie tylko jako rezultat pewnych działań, ale także jako cały proces ich realizacji [Hejduk, Grudzewski 2008].

Ogólność podstawowej definicji innowacji pozwala na jej postrzeganie z różnych perspektyw, a co za tym idzie liczne ich podziały. W literaturze wymienia się między innymi innowacje wyróżniane z uwagi na kryterium skali, w jakiej zachodzą. Można wówczas wymienić innowacje przełomowe, znaczące oraz drobne usprawnienia [Kalinowski 2010]. Z kolei w tak zwanej klasyfikacji przedmiotowej wyszczególnia się innowacje produktowe, procesowe i strategiczne (inaczej określane jako organizacyjne czy marketingowe) [Kalinowski 2010]. Za podstawowe źródło innowacji przyjąć można środowisko, które stymuluje, dostrzega oraz reaguje na zmiany. Innowacje powstawać mogą na bazie prac naukowo-badawczych i naukowo-technicznych prowadzonych przez ośrodki naukowe, informacji zamieszczanych w publikacjach i opracowaniach, informacji przekazywanych na targach, wystawach czy zjazdach naukowych, ale też na bazie nowych konstrukcji wyrobów, technologii i organizacji czy wynalazków i udoskonaleń w przedsiębiorstwach [Moczała 2005]. Sygnały dotyczące nadchodzącej innowacji mogą płynąć zarówno bezpośrednio z przedsiębiorstw czy organizacji, ale też ze źródeł zewnętrznych, choć nie będą one wywoływały tak szybkiej reakcji. Należy jednak zauważyć, że większość innowacji prowadzących do zmiany oblicza danej gałęzi przemysłu pochodziło właśnie spoza danej branży [Utterback 1996].

W świetle przytoczonych informacji wybór zmiennych, które mają stanowić podstawę decyzji o potencjale innowacyjnym danego obiektu powinien być zdeterminowany wiedzą o istocie samej innowacji czy źródłach ich powstawania. Ponadto niezbędne do uwzględnienia są również uwarunkowania danego obiektu, które w wypadku województw, wybranych przez autorki jako przedmiot zainteresowania, powinny obejmować między innymi kapitał ludzki, transfer technologii, źródła finansowania, otoczenie instytucjonalne czy politykę innowacyjną [Prystrom 2012]. Wybór województw, jako przedmiotu analizy podyktowany został zarówno dostępnością obszernego zbioru danych o województwach, jak i popularnością zagadnienia tworzenia rankingów w tym zakresie. Zasadniczym celem artykułu była nie tyle budowa nowych rankingów, co porównanie powstałych porządkowań przy zastosowaniu różnych metod ich tworzenia. W artykule dokonano klasyfikacji województw uwypuklając zmiany w otrzymanych wynikach przy uwzględnieniu różnych form porządkowania, bazując każdorazowo na identycznym zestawie danych wejściowych. Wyznaczono dla otrzymanych rankingów wskaźniki rang Spearmana oraz w odniesieniu do utworzonych grup indeks Randa, w celu określenia poziomu ich zgodności. Przeprowadzono ponadto analizę homogeniczności i heterogeniczności utworzonych zestawień oraz wyznaczono miernik poprawności grupowania.

OPIS ZASTOSOWANEJ PROCEDURY BADAWCZEJ

Realizując postawiony cel przeprowadzono eksperyment, którego wynikiem miało być porównanie rankingów województw przy wykorzystaniu różnych metod klasyfikacji. Pierwotnie uzyskane zmienne wykorzystane podczas badania pochodziły z roczników statystycznych i raportów dotyczących innowacyjności regionów publikowanych przez GUS. Dane dotyczyły przede wszystkim roku 2011 oraz w szczególnych przypadkach odnoszone były do roku 2010 lub 2009, co spowodowane było problemem z dostępnością danych z jednego okresu.

Wstępnie zastosowano procedurę doboru zmiennych, która pozwoliła zredukować wyjściowy zestaw do kilku zmiennych. Przebiegała ona w kilku etapach. Podczas pierwszego, dokonano merytoryczno-formalnej redukcji zgromadzonych zmiennych, ograniczając wyjściowy materiał badawczy do 34 wskaźników pozwalających zdefiniować potencjał innowacyjny województw.

Następnie ograniczono otrzymaną uprzednio listę wskaźników za pomocą kryterium ich zdolności dyskryminacyjnej przy zastosowaniu kilku wskaźników zmienności (pozycyjnych oraz klasycznego)¹. Przeprowadzona redukcja doprowadziła do sporządzenia różnych wynikowo zestawów, stanowiących materiał wyjściowy dla kolejnego etapu obejmującego eliminację zmiennych skorelowanych. Na tym etapie zastosowano metodę Hellwiga wraz z jego parametryczną modyfikacją przy uwzględnieniu różnych wartości progowych. Szczegółowy opis przebiegu wymienionych etapów doboru został przedstawiony w pracy [Olszewska 2014].

Zastosowane różnorodne procedury doprowadziły finalnie do wyboru pozycyjnego współczynnika zmienności jako kryterium dyskryminacyjnego. Kolejny etap obejmował zmodyfikowaną parametryczną metodę Hellwiga zaproponowaną przez A. Młodaka, w której zamiast sum bezwzględnych współczynników korelacji liniowej Pearsona stosuje się ich medianę [Młodak 2006]. Przeprowadzone procedury umożliwiły wskazanie zmiennych, dla których po zastosowaniu jednolitej procedury grupowania otrzymane wyniki cechowała najwyższa poprawność. W ten sposób uzyskany zestaw został zredukowany do pięciu następujących:

1. eksport wyrobów wysokiej i średniowysokiej techniki (w % eksportu ogółem);
2. nakłady na działalność badawczo-rozwojową ponoszone przez przedsiębiorstwa (w % nakładów);
3. nakłady na działalność innowacyjną na 10 tys. przedsiębiorstw;
4. przychody netto z eksportu produktów nowych lub istotnie ulepszonych w % przychodów netto ze sprzedaży ogółem;
5. zatrudnienie w działalności badawczo-rozwojowej na 1 tys. aktywnych zawodowo.

¹ Wartości krytyczne wskaźników przyjęto na poziomie 10%.

Otrzymany w wyniku przeprowadzonego badania zestaw zmiennych poddano standaryzacji². Przy wykorzystaniu dwóch metod wzorcowych i dwóch bezwzorcowych dokonano uszeregowania województw pod kątem ich potencjału innowacyjnego. Wybranymi metodami wzorcowymi były: wzorcowa metoda Hellwiga³ i metoda dystansowa⁴. Pozostałymi metodami wybranymi do grupowania obiektów były: metoda rang⁵ i metoda sum⁶.

Z uwagi na fakt, że dane wejściowe zredukowane były przy zastosowaniu miar pozycyjnych (ze względu na znaczną asymetrię i wysoką koncentrację części danych wejściowych), to również taką miarę wybrano do przeprowadzenia standaryzacji. Zastosowano standaryzację pozycyjną bazującą na medianowym odchyleniu bezwzględny⁷. Po przekształceniu zmiennych zbudowano cztery dodatkowe rankingi województw, oparte na wymienionych wyżej metodach. Zestawienie metod wraz z ich oznaczeniami zamieszczono w Tabeli 1.

² Zastosowano klasyczną standaryzację zgodną z następującym wzorem: $z_i = \frac{x_i - \bar{x}}{s}$, gdzie

x_i – zmienna, \bar{x} – średnia zmiennej X, s – odchylenie standardowe zmiennej X.

³ Miernik syntetyczny wzorcowej metody Hellwiga wyznaczany jest jako [Nowak, 1990]:

$$\mu_i = 1 - d_{i0} / d_0 \quad (i=1,2,\dots,n), \text{ gdzie } d_{i0} = \left[\sum_{j=1}^m (z_{ij} - z_{0j})^2 \right]^{0,5}$$

jest odległością każdego obiektu z_{ij} od obiektu wzorca z_{0j} wyznaczonego jako wartość maksymalna dla zmiennych będących stymulantami; $d_0 = \bar{d}_0 + 2s(d_0)$, przy czym $\bar{d}_0 = \frac{1}{n} \sum_{i=1}^n d_{i0}$, zaś $s(d_0) = \left[\frac{1}{n} \sum_{i=1}^n (d_{i0} - \bar{d}_0)^2 \right]^{0,5}$.

⁴ Miernik syntetyczny metody dystansowej jest zgodnie ze wzorem [Panek 2009]:

$$\mu_i = \left(\frac{d_{i0} - \min_i \{d_{i0}\}}{\max_i \{d_{i0}\} - \min_i \{d_{i0}\}} \right)^p$$

gdzie p to dodatni wykładnik potęgowy, zaś d_{i0} – odległość

obektu od wzorca wyznaczona jak w metodzie Hellwiga.

⁵ Miernik syntetyczny wyznaczany jest jako średniej wartość z rang kolejnych zmiennych [Malina, Wanat 1995].

⁶ Miernik syntetyczny wyznaczany jest jako średnia arytmetyczna wszystkich zmiennych, a następnie normuje zgodnie ze wzorem [Malina, Zeliaś 1998]:

$$\mu_i'' = \left(\mu_i - \min_i \{ \mu_i \} \right) / \left(\max_i \{ \mu_i - \min_i \{ \mu_i \} \} \right)$$

⁷ Standaryzacja pozycyjna, zgodną z wzorem [Młodak 2006]: $z_{ij} = \frac{x_{ij} - M(x_j)}{MOB(x_j)}$, gdzie $MOB(x_j)$

jest medianowym odchyleniem bezwzględny j -tej zmiennej wyznaczonym następująco: $MOB(x_j) = M_i |x_{ij} - M(x_j)|$, zaś $M(x_j)$ jej medianą.

Tabela 1. Wybrane metody porządkowania obiektów

Oznaczenie metody	Metoda standaryzacji	Metoda klasyfikacji obiektów
M1	Standaryzacja klasyczna	Wzorcowa metoda Hellwiga
M2		Metoda dystansowa
M3		Metoda rang
M4		Metoda sum
M5	Standaryzacja pozycyjna	Wzorcowa metoda Hellwiga
M6		Metoda dystansowa
M7		Metoda rang
M8		Metoda sum

Źródło: opracowanie własne

Aby porównać otrzymane wyniki rankingów zbadano ich poziom skorelowania wykorzystując w tym celu współczynnik korelacji rang Spearmana. Po zbudowaniu rankingów możliwe było przeprowadzenie kolejnego etapu analizy województw – podziału badanych obiektów na cztery grupy⁸. Chcąc określić na ile wyniki otrzymanych procedur różnią się od siebie, porównano utworzone grupy wykorzystując indeks Randa [Gatnar i Walesiak 2004; Rand 1971]. Zbadano także na ile poprawne są użyte metody grupowania poprzez wyznaczenie miar homogeniczności⁹ i heterogeniczności¹⁰ grup obiektów oraz miernika poprawności grupowania jako ich ilorazu [Panek 2009].

WYNIKI PRZEPROWADZONEGO GRUPOWANIA WOJEWÓDZTW

W wyniku przeprowadzonej procedury uzyskano osiem rankingów przedstawionych syntetycznie w Tabeli 2. W każdym z wierszy odniesionym do konkretnego województwa przytoczono uzyskane przez dane województw miejsce

⁸ Wartości miernika syntetycznego (μ_i dla $i=1,2,\dots,16$) zostały podzielone na cztery grupy [Nowak 1990]: I grupa: $\mu_i \geq \bar{\mu} + s_\mu$; II grupa: $\bar{\mu} \leq \mu_i < \bar{\mu} + s_\mu$; III grupa: $\bar{\mu} - s_\mu \leq \mu_i < \bar{\mu}$; IV grupa: $\mu_i < \bar{\mu} - s_\mu$, gdzie $\bar{\mu}$ jest średnią z mierników syntetycznych, zaś s_μ ich odchyleniem standardowym.

⁹ Miary homogeniczności wyznaczają zgodnie z wzorem [za Panek 2009]: $\bar{d} = \frac{1}{z} \sum_{r=1}^z d_r$, gdzie

$$d_r = \frac{2}{n_r(n_r+1)} \sum_{i=1}^{n_r} \sum_{i'>i}^{n_r} d_{ii'}$$

dla $r=1,2,\dots,z$ i $d_{ii'}$ będącą odległością między obiektami należącymi do badanej grupy.

¹⁰ Miary heterogeniczności wyznaczają zgodnie z wzorem [za Panek 2009]: $\bar{d} = \frac{1}{z} \sum_{r=1}^z d_r$,

gdzie $d_r = \frac{1}{n_r n_r'} \sum_{i=1}^{n_r} \sum_{i'=1}^{n_r'} d_{ii'}$ dla $r=1,2,\dots,z$ i $d_{ii'}$ będącą odległością między obiektami nie należącymi do jednej grupy.

w rankingu oraz przynależność do danej grupy (podaną w nawiasach). W tabeli wskazano również na maksymalną różnicę zachowując ten sam schemat zapisu.

Tabela 2. Porównanie rankingów po standaryzacji klasycznej

Województwo	Metoda kwalifikacji obiektów								Maks. Różnica
	M1	M2	M3	M4	M5	M6	M7	M8	
Dolnośląskie	3(2)	3(2)	3(1)	2(2)	6(2)	6(2)	3(1)	2(2)	4(1)
Kujawsko-pomorskie	11(3)	11(3)	9(3)	11(3)	9(3)	9(3)	9(3)	11(3)	2(0)
Lubelskie	12(3)	12(3)	14(4)	13(3)	14(3)	14(3)	14(4)	13(3)	2(1)
Lubuskie	13(3)	13(3)	13(4)	12(3)	16(4)	16(4)	13(4)	12(3)	4(1)
Łódzkie	8(2)	8(2)	8(2)	7(2)	8(3)	8(3)	8(2)	7(2)	1(1)
Małopolskie	7(2)	7(2)	5(2)	8(2)	4(2)	4(2)	5(2)	8(2)	4(0)
Mazowieckie	4(2)	4(2)	6,5(2)	4(2)	5(2)	5(2)	6,5(2)	4(2)	2,5(0)
Opolskie	9(3)	9(3)	10(3)	9(3)	10(3)	10(3)	10(3)	9(3)	1(0)
Podkarpackie	5(2)	5(2)	6,5(2)	5(2)	3(2)	3(2)	6,5(2)	5(2)	3,5(0)
Podlaskie	10(3)	10(3)	11(3)	10(3)	11(3)	11(3)	11(3)	10(3)	1(0)
Pomorskie	1(1)	1(1)	3(1)	1(1)	1(1)	1(1)	3(1)	1(1)	2(0)
Śląskie	6(2)	6(2)	1(1)	6(2)	7(2)	7(2)	1(1)	6(2)	6(1)
Świętokrzyskie	15(4)	15(4)	15(4)	15(4)	15(3)	15(3)	15(4)	15(4)	0(1)
Warmińsko-mazurskie	16(4)	16(4)	16(4)	16(4)	12(3)	12(3)	16(4)	16(4)	4(1)
Wielkopolskie	2(1)	2(1)	3(1)	3(2)	2(1)	2(1)	3(1)	3(2)	1(1)
Zachodniopomorskie	14(3)	14(3)	12(4)	14(4)	13(3)	13(3)	12(4)	14(4)	2(1)

Źródło: obliczenia własne

Analizując wyniki przedstawione w Tabeli 2 można zauważyć, że maksymalną różnicą pomiędzy uzyskanymi przez dane województwo miejscami w poszczególnych rankingach była wartość 6 odnosząca się do województwa śląskiego. Oznacza to, że w zależności od przyjętej procedury województwo klasyfikowało się na pierwszym, szóstym bądź siódmym miejscu. Przyczyny takiej sytuacji można upatrywać w skrajnie niskich bądź wysokich wartościach niektórych spośród analizowanych zmiennych charakteryzujących województwo. Pojawienie się takich wartości mogło w zależności od wybranej metody wpływać znacząco na otrzymane wyniki. Pojawiło się również województwo, które niezależnie od wybranego rankingu pozostawało na tym samym miejscu (województwo świętokrzyskie). Jednak znaczna część badanych obiektów zmieniała swoją pozycję maksymalnie o jedno bądź dwa miejsca. W aspekcie przyporządkowania do grup różnice nie były już tak znaczące, a zachodzące zmiany obejmowały przesunięcie co najwyżej o jedną grupę.

W celu zbadania stopnia zgodności uzyskanych wyników rankingów wyznaczono współczynnik rang Spearmana, zaś w odniesieniu do grup – indeks Randa. Wyniki zaprezentowano zbiorczo w Tabeli 3. Kilka metod rangowania dało identyczne wyniki – były to metody: wzorcowa Hellwiga (M1 i M5) oraz dystansowa (M2 i M6) – zarówno przy wykorzystaniu standaryzacji klasycznej jak i pozycyjnej. Pozostałe uzyskane wartości współczynników również wskazują na silną korelację pomiędzy wynikami rankingów. Indeks Randa (choć uzyskane wartości były tu zasadniczo niższe) sugeruje znaczne podobieństwo w przypisaniu województw do stworzonych grup.

Tabela 3. Wartości współczynników rang Spearmana i indeksu Randa

	M1	M2	M3	M4	M5	M6	M7	M8	Współczynnik rang Spearmana
M1		1,000	0,917	0,991	0,912	0,912	0,865	0,882	
M2	1,000		0,917	0,991	0,912	0,912	0,865	0,882	
M3	0,775	0,775		0,911	0,917	0,917	1,000	0,911	
M4	0,883	0,883	0,775		0,991	0,991	0,911	1,000	
M5	0,775	0,775	0,700	0,708		1,000	0,917	0,991	
M6	0,775	0,775	0,700	0,708	1,000		0,917	0,991	
M7	0,775	0,775	1,000	0,775	0,700	0,700		0,911	
M8	0,883	0,883	0,775	1,000	0,708	0,708	0,775		
Indeks Randa									

Źródło: opracowanie własne

Wyniki dotyczące poprawności grupowania obiektów według metod poprzedzonych zarówno klasyczną, jak i pozycyjną standaryzacją przedstawiono w Tabeli 4. Pomimo wspólnego zestawienia wszystkich metod należy pamiętać, że analizy dotyczące miar homogeniczności oraz heterogeniczności grup województw powinny być prowadzone odrębnie dla różnych form standaryzacji. Miarą służącą do porównania wszystkich metod może być jedynie miernik poprawności grupowania województw. Na nim skupiono się w dalszej analizie.

Tabela 4. Wyniki miar poprawności grupowania

Miary	Metoda klasyfikacji zmiennych							
	M1	M2	M3	M4	M5	M6	M7	M8
Miara homogeniczności grup województw	1,172	1,172	1,434	0,949	3,173	3,173	4,241	3,269
Miara heterogeniczności grup województw	3,439	3,439	3,112	3,703	11,158	11,158	9,042	11,088
Miernik poprawności grupowania województw	0,341	0,341	0,460	0,256	0,284	0,284	0,469	0,295

Źródło: obliczenia własne

Analizując wyliczenia zgromadzone w Tabeli 4, zauważyć można, że metoda rang (oznaczona jako M3 oraz M7 w zależności od wybranego sposobu standaryzacji) była metodą o najwyższych wartościach miernika, co oznacza, że podczas wykorzystania tej metody grupowanie obiektów cechowało się najmniejszą poprawnością. Biorąc pod uwagę standaryzację klasyczną najlepszą metodą okazała się według wyznaczonych mierników metoda sum. Z kolei przy uwzględnieniu standaryzacji pozycyjnej była to dowolna metoda wzorcowa, choć tutaj również metoda sum uzyskała niewiele gorszy wynik.

PODSUMOWANIE

Podsumowując przeprowadzony eksperyment statystyczny należy pamiętać, że na uzyskane wyniki wpływ miał charakter wybranych zmiennych. Jednak już na opisanym przykładzie można wyciągnąć wnioski przydatne podczas konstrukcji bardziej rozbudowanych rankingów. Istotne jest aby badać, jak poszczególne etapy konstrukcji wpływają na wyniki rankingu, gdyż każdy z etapów może znacząco oddziaływać na finalnie uzyskiwany ranking.

Warto zaznaczyć również, że wybór skomplikowanych metod nie zawsze daje bardziej poprawne wyniki. W wypadku opisanego badania najlepszą okazała się jedna z najprostszych metod – metoda sum.

Chcąc otrzymać wiarygodne wyniki warto jest pokusić się o przygotowanie kilku rankingów, przeprowadzonych różnymi metodami, przy różnych założeniach, a następnie zbadać je pod kątem zgodności otrzymanych wyników. W wypadku analizy województw uzyskanie danego miejsca w rankingu wpływa silnie na postrzeganie regionu w oczach społeczeństwa, a często skutkuje poważnymi konsekwencjami, jak chociażby wysokość przyznawanych dla regionu funduszy. Jak wykazał eksperyment wystąpienie różnicy nawet sześciu miejsc danego województwa w rankingach może pozostać niezauważone czy to z uwagi na stworzenie pojedynczego rankingu, czy nawet w zestawieniu zbiorczym, gdyż porównanie wyników rankingów wskaże na ich silną korelację. Logicznie wydaje się to zadziwiające, szczególnie w aspekcie porównania województw pod kątem oceny ich innowacyjności. Być może bardziej zasadnym w tym kontekście rozwiązaniem byłoby wyciąganie wniosków na bazie przyporządkowania do grup, niż zaś na szczegółowym rankingu.

BIBLIOGRAFIA

- Christensen C. N. (2010), *Przełomowe innowacje*, Wydawnictwo Naukowe PWN, Warszawa.
- Gatnar E., Walesiak M. (red.) (2004), *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wydawnictwo AE we Wrocławiu, Wrocław.
- Hejduk I., Grudzewski W. (2008), *Zarządzanie technologiami: zaawansowane technologie i wyzwanie ich komercjalizacji*, Difin, Warszawa.
- Kalinowski T. B. (2010), *Innowacyjność przedsiębiorstw a systemy zarządzania jakością*, Oficyna a Wolters Kluwer business, Warszawa.
- Mahdjoubi D. (1997), *Schumpeterian Economics and the Trilogy of 'Invention-Innovation-Diffusion'*, [online], dostęp zdalny: www.ischool.utexas.edu/~darius/17-Schumpeter-innovation.pdf, [data wejścia: 01.07.2013].
- Malina A., Wanat S. (1995), *Przestrzenna analiza rozwoju Polski*, *Wiadomości Statystyczne*, nr 5, s. 20-25.
- Malina A., Zeliaś A. (1998), *On Building Taxonomic Measures on Living Conditions*, *Statistic in Transaction*, t. 3, nr 3, s. 523-544.
- Młodak A. (2006), *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa.
- Moczała A. (2005), *Zarządzanie innowacjami. Narzędzia Innowatyki. Transfer Technologii. Zarządzanie Technologią*, Wydawnictwo Akademii Techniczno-Humanistycznej w Bielsku-Białej, Bielsko-Biała.
- Nazarko J. (2013), *Regionalny foresight gospodarczy. Scenariusze rozwoju innowacyjności mazowieckich przedsiębiorstw*, Związek Pracodawców Warszawy i Mazowsza, Warszawa.
- Nowak E. (1990), *Metody taksonomiczne w klasyfikacji obiektów społeczno-ekonomicznych*, Państwowe Wydawnictwo Ekonomiczne, Warszawa.
- Nowakowska A. (red.) (2009), *Zdolności innowacyjne polskich regionów*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Olszewska A. (2014), *Wykorzystanie wybranych metod taksonomicznych do oceny potencjału innowacyjnego województw*, *Taksonomia. Klasyfikacja i analiza danych – teoria i zastosowania*, nr 328, s. 167-176.
- Panek T. (2009), *Statystyczne metody wielowymiarowej analizy porównawczej*, Szkoła Główna Handlowa w Warszawie, Warszawa.
- Pawlik A. (2012), *Potencjał innowacyjny w rozwoju regionalnym*, Wydawnictwo Uniwersytetu Jana Kochanowskiego, Kielce.
- Prystrom J. (2012), *Innowacje w procesie rozwoju gospodarczego. Istota i uwarunkowania*, Difin, Warszawa.
- Rand W. M. (1971), *Objective criteria for the evaluation of clustering methods*, „*Journal of the American Statistical Association*”, vol. 66 nr .336, s. 846–850.
- Utterback J. (1996), *Mastering The Dynamics of Innovation*, Harvard Business Press.
- Wiśniewska J., Janasz K. (2013), *Innowacje i jakość w zarządzaniu organizacjami*, CeDeWu, Warszawa.

**THE USE OF SELECTED METHODS OF ORDERING OBJECTS
FOR CLASSIFICATION OF VOIVODSHIPS
IN TERMS OF THEIR INNOVATIVE POTENTIAL**

Abstract: Defining and identifying determinants of the innovative potential of the region is difficult. Despite this, it is one of the determining factors in terms of the development of the region and its competitive position. Even more than the availability of data, relevant in this context becomes the problem of the selection of the classification procedure. The main objective of this paper is to group voivodships in terms of their innovation potential through the use of different methods of ordering. In the paper the classification of regions highlighting the changes between the results taking into account the different forms of ordering.

Keywords: innovation, methods of organizing objects

ZASTOSOWANIE TECHNIK EKSPLOKACJI TEKSTU DO ANALIZY OPINII KONSUMENCKICH

Marcin Olszewski, Tomasz Ząbkowski

Katedra Informatyki

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mail: marcin_olszewski@sggw.pl, tomasz_zabkowski@sggw.pl

Streszczenie: W niniejszej publikacji zaproponowano jedną z metod eksploracji danych – reguły asocjacyjne do wykrycia zależności w opiniach konsumenckich, na przykładzie opinii jednego z hoteli amerykańskich. Wykorzystanie tej techniki wynikało m.in. z dużej ilości dostępnych danych oraz faktu, że otrzymane reguły w sposób niezwykle czytelny prezentują zależności znalezione w danych. W badaniu odkryto szereg reguł, które mogą stanowić cenne źródło informacji o jakości usług oraz postrzeganiu obiektu przez klientów korzystających z usług hotelowych.

Słowa kluczowe: eksploracja tekstu, reguły asocjacyjne, opinie konsumenckie

WPROWADZENIE I OKREŚLENIE CELU BADANIA

Celem prezentowanych badań było zastosowanie reguł asocjacyjnych do analizy opinii konsumenckich. Tego typu analiza jest narzędziem, które może dostarczyć istotnych informacji mających wpływ na wizerunek firmy, co jest nie bez znaczenia dla firm działających na bardzo konkurencyjnym rynku. Jednakże, pozyskiwanie takich informacji nie jest procesem łatwym i konieczne w tym celu jest przetworzenie ogromnej liczby dokumentów takich jak ankiety, kwestionariusze, opinie. Problem ten wymaga nieco innego spojrzenia na dane i doboru określonego sposobu ich analizowania za pomocą technik eksploracji danych [Larose 2006, Han i Kamber 2001]. Metody te mają kluczowe znaczenie w przypadku badania opinii konsumenckich, umożliwiają one bowiem tworzenie reprezentacji tekstu, która będzie posiadała cechy tekstu wejściowego, a przy tym formę adekwatną do wykorzystania przez algorytmy. Dokumenty tekstowe można traktować jako zbiory z atrybutami o charakterze nominalnym, dlatego też szczególnie użytecznym sposobem ich analizowania może być analiza koszykowa

[Pasztyła 2005], która za pomocą reguł wskazuje na współwystępujące słowa [Amir i in. 2005]. Znalezienie takich zestawów słów wewnątrz opinii konsumenckich, przy uwzględnieniu ich nacechowania semantycznego pozwala na stworzenie bazy wiedzy, w oparciu o którą możliwe jest także dokonywanie klasyfikacji nowych opinii. W wymienionej literaturze przedmiotu opisuje się szeroko zastosowanie analizy koszykowej do wyszukiwania reguł w różnych bazach, jednak w przypadku baz nieustrukturyzowanych, takich jak bazy tekstowe, nie przeprowadzono jak dotąd eksperymentów, które dodatkowo uwzględniałyby rozpatrywane w pracy aspekty dotyczące przetwarzania tekstów pod kątem usprawnienia działania algorytmów asocjacyjnych. To właśnie jest głównym zamierzeniem autorów niniejszego artykułu i stanowi o jego wartości.

METODY ZMNIEJSZANIA REPREZENTACJI TEKSTU OPINII KLIENCKICH

Każda opinia konsumencka jest dokumentem tekstowym, który nie posiada konkretnej struktury. Brak sformalizowanej postaci takiego dokumentu uniemożliwia jego natychmiastową klasyfikację i tym bardziej pozyskanie z niego określonych informacji. Na tym etapie niezbędne jest wstępne przetworzenie opinii, czego efektem będzie odpowiednia dla danego algorytmu postać tekstu, zawierająca ponadto zmniejszoną reprezentację. Możliwości algorytmów eksploracji tekstu są mocno ograniczone jeżeli chodzi o pracę na dużej ilości danych (duża złożoność obliczeniowa i długi czas pracy), dlatego etap ten obejmuje przekształcenie tekstu do zmniejszonej i uproszczonej postaci. Postać taka umożliwia o wiele szybszą i bardziej efektywną analizę danych. W celu zmniejszenia wielkości wejściowych danych tekstowych najczęściej stosuje się podejście tzw. „worka słów” (ang. bag of words), w którym ważniejsza jest częstość występowania słów w poszczególnych dokumentach niż miejsce, czy kolejność ich wystąpienia. Zasadniczo, większość z metod pozyskiwania zmniejszonych reprezentacji tekstu opiera się na prostych statystykach lub zależnościach. Niżej wymieniono metody zmniejszania reprezentacji dokumentów tekstowych, które są istotne dla efektywnego przetwarzania tekstu.

Metoda oparta na prawie Zipf'a

Zgodnie z teorią informacji w każdym języku naturalnym istnieje zależność, która mówi, że rozkład częstości słów występujących w danym języku nie jest losowy, czyli taki, w którym wybrane słowo X występuje z tą samą częstością co słowo Y. Słów najczęściej używanych jest proporcjonalnie wiele razy więcej niż słów najrzadziej używanych w dowolnych rzeczywistym tekście. Ten nierównomierny rozkład słów w językach naturalnych został potwierdzony przez badanie amerykańskiego lingwisty George'a Zipf'a. Prawo to umożliwia odnalezienie zależności w ogromnych ilościach danych tekstowych, które na pierwszy rzut oka mogą wydawać się jednolite. Ponadto prawo to można

wykorzystać do określenia rangi słów. Gdy każdemu słowu z rozkładu Zipf'a przypisze się wartość oznaczającą pozycję w rankingu ważności takiego słowa na podstawie częstości jego wystąpienia, to częstość występowania słów będzie odwrotnie proporcjonalna do pozycji tego słowa w rankingu ważności słów [Ward 1997].

Metoda oparta na stop liście słów

Oprócz własności, którą zauważył Zipf, każdy język charakteryzuje specyficzna konstrukcja posiadająca odpowiednie kryteria składniowe i fleksyjne. Do budowy zdań używa się różnych części mowy i są to (w zależności od języka): zaimki, przyimki, rodzajniki, spójniki, wykrzykniki. Słowa należące do wymienionych kategorii mają bardzo wysoką częstość wystąpień, ale nie niosą żadnej użytecznej wiedzy. Metoda stop listy [Rajaraman i Ullman 2012] (ang. stop list, stop word) polega na pominięciu tych słów na początkowym etapie przygotowania danych w celu usprawnienia pracy algorytmu.

Metoda przycinania

Poza ograniczaniem liczności zbioru słów poprzez tworzenie stop listy, można także zmniejszać reprezentację tekstu za pomocą miar statystycznych – jest to tzw. przycinanie (ang. pruning). Rozwiązanie takie polega na usuwaniu słów najczęściej lub zbyt często występujących w danym dokumencie tekstowym oraz słów, których częstość występowania jest bardzo mała. Określenie progów oddzielających słowa nieistotne z powodu zbyt dużej lub zbyt małej częstości użycia, znacznie zmniejsza rozmiar reprezentacji, dzięki czemu poprawia się efektywność przetwarzania danych oraz redukuje szum informacyjny, nie zmieniając przy tym znacząco wyników działania algorytmu eksploracji tekstu.

ZASTOSOWANE TECHNIKI – REGUŁY ASOCJACYJNE

Wyodrębnienie występujących zestawów słów w treściach opinii może zostać przeprowadzone w oparciu o analizę koszykową. Jej celem jest wyszukiwanie zależności między badanymi obiektami – w tym przypadku między poszczególnymi słowami – i przedstawienia ich w postaci reguł. Reguły te nazywa się regułami asocjacyjnymi i mają one postać:

$[poprzednik: warunek] \Rightarrow [następnik: warunek]$.

Wyszukiwanie reguł asocjacyjnych w opiniach sprowadza się do ilościowego określenia relacji pomiędzy poszczególnymi słowami. Ocenę przydatności reguł asocjacyjnych wyznacza się miarami jakości [Paszyła 2005], [Tan i in. 2005]: wsparciem (ang. support), które mówi jak często otrzymana reguła występuje całej bazie opinii, pewnością (ang. confidence), która mówi, jak silna jest reguła, co odpowiada prawdopodobieństwu warunkowemu, że jeżeli w losowej opinii, pojawiło się słowo X to pojawi się również słowo Y, oraz przyrostem (ang. lift), który mówi jaki wpływ ma pojawienie się w opinii słowa X na wystąpienie w opinii

również słowa Y . Wymienione miary są obiektywnym sposobem na ocenę istotności i siły niesionych przez reguły informacji, a ich formułowanie jest istotne, gdyż często dla określonych danych wejściowych z wieloma atrybutami powstaje bardzo dużo potencjalnych reguł, z których nie wszystkie są użyteczne.

ANALIZA OPINII KONSUMENCKICH

Przygotowanie danych

Przeprowadzony eksperyment polegał na stworzeniu zbioru reguł asocjacyjnych dla przykładowego dokumentu z opiniami na temat jednego z amerykańskich hoteli. Analizę koszykową przeprowadzono algorytmem Apriori [Agrawal i Srikant 1994]. Algorytm ten w każdym kroku iteracji generuje zbiory kandydujące o liczności większej niż w poprzednim kroku, które po spełnieniu określonego progu minimalnego wsparcia stają się zbiorami częstymi. Ze zbiorów częstych wybiera się te, które osiągają wymagany poziom ufności i z nich tworzy się reguły asocjacyjne. Dane ujęte w badaniu pochodzą z repozytorium internetowego, które opracowali Ganesan i Zhai [Ganesan i Zhai 2011]. Cały zbiór danych zawiera listę rzeczywistych opinii hoteli z 10 różnych miast świata. Analiza została ograniczona do przebadania opinii jednego z hoteli, mianowicie hotelu Riviera z Las Vegas. Hotel ten posiada 576 opinii, w których łączna liczba użytych słów wynosi 123727. Treść przykładowych opinii prezentuje Tabela 1. Do analizy wykorzystano treść opinii, gdzie każda opinia to jeden unikalny koszyk, czyli zbiór słów składający się na tę opinię (transakcję). Zanim możliwe było wyszukanie powiązań, dane należało sprowadzić do formatu transakcyjnego.

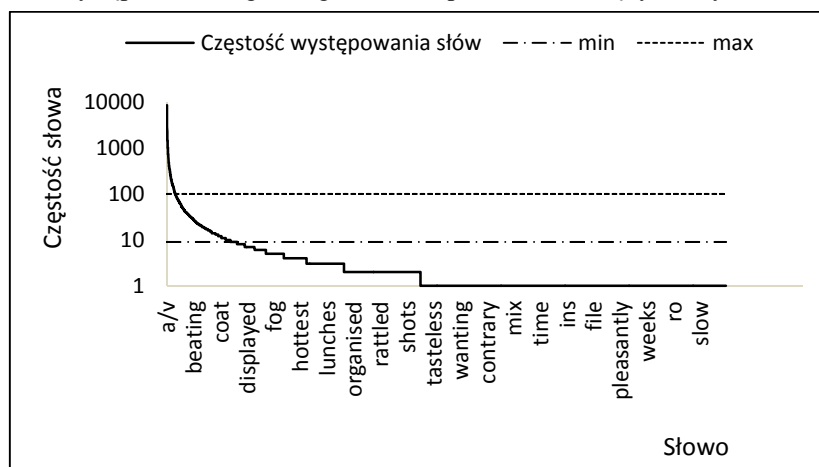
Tabela 1. Przykładowe opinie na temat hotelu

Nr opinii	Treść
1	Hotel Room: Monaco Tower, 4927, king bed. Bed was comfortable and pillows had some support. It won't remind you of home but a solid nights sleep is easily achievable. Room I'd guess is 300 sq ft. Dark brown decor and white sheets. Flat screen tv, 26-32 inches, somewhere around that size. Three dresser drawer, two night stands, desk and an arm chair with foot rest. Safe (I think you have to pay for it) in night stand. View of the Hilton and parking deck behind the hotel. Con: walls seemed to be thin or one of our neighbors was deaf. Wed night we could hear one of our neighbors Tv's like it was in our room.
2	The location is a bit far north of the main action on the strip, and still a long bus ride to Fremont Street. The room was small and below average. The staff are rude, and the casino staff even worse. Find a better place to stay.

Źródło: opracowanie własne

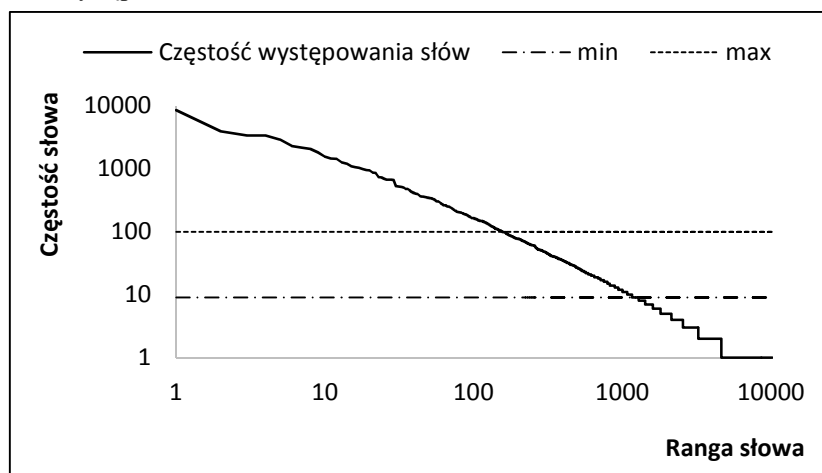
Z każdej opinii usunięto wszelkie znaki interpunkcyjne (kropki, przecinki, średniki, myślniki, znaki zapytania, wykrzykniki, dwukropki, nawiasy). Powstała tabela zawiera tylko słowa języka angielskiego (oraz ewentualne nieznanne wyrażenia wynikające z błędów osób piszących opinię). Szereg słów nie niesie żadnej użytecznej informacji, dlatego zawartość tabeli została poddana zabiegom zmniejszającym liczbę słów we wszystkich opiniach. W tym celu została stworzona stop lista, która zawiera słowa będące zaimkami (31 sztuk), przyimkami (72 sztuk), rodzajnikami (5 sztuk), spójnikami (34 sztuk) i wykrzyknikami (114 sztuk) [Słownik spójników 2013], [Słownik przyimków 2013], [Słownik wykrzykników 2013]. W badaniu uwzględniono także wielkość liter dla słów rozpoczynających zdania. Ten zabieg pozwolił zmniejszyć reprezentację tekstu usuwając łącznie 48002 słów (powtarzających się wystąpień słów ze stop listy), co spowodowało, że w bazie pozostało 75725 słów. W następnym kroku usunięte zostały słowa zbędne w oparciu o metodę przycinania, polegającą na usuwaniu słów o zbyt dużej i zbyt małej częstości wystąpień [Borycki i Sołdacki 2002]. Progi te zostały dobrane eksperymentalnie, w taki sposób, aby w końcowej ocenie miary statystyczne dla reguł asocjacyjnych pozostawały możliwe najwyższe tj. najlepsze w rozpatrywanym przypadku. Zostały one pokazane na wykresach nr 1 i nr 2. Słowa o częstości większej niż 100 to głównie słowa ze stop listy, ale także słowa, które nie niosą żadnej użytecznej informacji. Słowa o częstości niższej niż 9, to najczęściej błędy użytkowników lub słowa tak rzadkie, że usunięcie ich nie wpłynie na wyniki, ale poprawi efektywność przetwarzania, gdyż tych słów jest aż 8918 (17324 słów uwzględniając ich powtórzenia). Schodkowość wykresów wynika z powtarzających się licznosci poszczególnych słów.

Rysunek 1. Wykres zależności poszczególnych słów z opinii o hotelu do częstości ich wystąpień (z uwagi na ograniczenia prezentowane są tylko wybrane słowa)



Źródło: opracowanie własne

Rysunek 2. Wykres zależności rang poszczególnych słów z opinii o hotelu do częstości ich wystąpień



Źródło: opracowanie własne

Dla tak dobranych progów zostało pominiętych 9074 unikatowych słów (uwzględniając stop listę liczba ta zmniejszy się z uwagi na powtarzające się słowa). Metoda przycinania pozwoliła usunąć łącznie 48940 słów, co spowoduje, że w bazie pozostanie 26785 słów. Wewnątrz wszystkich opinii są jednak słowa powtarzające się, gdyż w jednej opinii możliwe było użycie kilku takich samych słów. Do analizy wystarczyła reprezentacja składająca się z unikatowych słów wewnątrz opinii [Morzy 2013]. Po wykluczeniu powtarzających się słów wewnątrz opinii pozostało 23609, co stanowiło zbiór wejściowy do analizy. Tabela 1 prezentuje transakcyjną postać opinii klientów hotelu po przygotowaniu.

Tabela 2. Transakcyjna postać opinii wg przygotowania

Nr opinii	Słowo
1	always
1	anyway
1	ask
⋮	⋮
2	action
2	average
2	far
⋮	⋮
3	bathroom
3	best
⋮	⋮

Źródło: opracowanie własne

Wykrywanie reguł asocjacyjnych w tekstach opinii

W badaniu rozważano reguły dwu i trzejelementowe. Ustalony został także minimalny poziom wsparcia na poziomie 3% oraz pewności reguły na poziomie 30%. Dla tak skonfigurowanego procesu znaleziono 68 reguł. Wyniki dla dziesięciu pierwszych reguł przedstawiono w tabelach 3, 4, 5 uwzględniając szeregowanie według miar wsparcia, pewności i przyrostu.

Na podstawie przeprowadzonej analizy najczęstszymi zwrotami (popartymi pewnością) użytymi w opiniach użytkowników są m.in. pary: reading–reviews, slots–fun, shop–coffee, screen–flat, flat–tv, monte–carlo, minutes–took, line–long, checkin–took. U wielu oceniających padła formuła zawierająca słowa reviews i reading. Może to oznaczać, że użytkownicy chętnie czytają i dzielą się opiniami na temat hotelu, co wskazuje, że ten kanał komunikacji jest niezwykle opiniotwórczy dla wielu użytkowników, w tym także dla potencjalnych klientów.

Tabela 3. Wybrane reguły asocjacyjne dla tabeli słów (posortowane wg wsparcia)

Przyrost	Wsparcie (%)	Pewność(%)	Reguła
5.67	5.73	84.62	reading => reviews
16.94	5.73	100.00	monte => carlo
3.92	5.38	47.69	slots => fun
5.57	4.86	60.87	shop => coffee
13.37	4.51	81.25	screen => flat
5.02	4.34	71.43	flat => tv
3.36	4.34	45.45	minutes => took
2.49	4.34	33.78	checkin => took
3.00	4.17	37.50	line => long
2.24	3.99	30.67	use => bit

Źródło: opracowanie własne

Tabela 4. Wybrane reguły asocjacyjne dla tabeli słów (posortowane wg pewności)

Przyrost	Wsparcie (%)	Pewność(%)	Reguła
16.94	5.73	100.00	monte => carlo
14.31	3.47	86.96	tv & screen => flat
5.67	5.73	84.62	reading => reviews
13.37	4.51	81.25	screen => flat
14.40	3.47	80.00	tv & flat => screen
5.40	3.47	76.92	screen & flat => tv
13.37	4.51	74.29	flat => screen
5.05	3.99	71.88	screen => tv
5.02	4.34	71.43	flat => tv
4.25	4.51	63.41	read => reviews

Źródło: opracowanie własne

Tabela 5. Wybrane reguły dla tabeli słów (posortowane wg przyrostu)

Przyrost	Wsparcie (%)	Pewność(%)	Reguła
16.94	5.73	100.00	monte => carlo
14.40	3.47	80.00	tv & flat => screen
14.31	3.47	86.96	tv & screen => flat
13.37	4.51	81.25	screen => flat
5.67	5.73	84.62	reading => reviews
5.57	4.86	60.87	shop => coffee
5.40	3.47	76.92	screen & flat => tv
5.02	4.34	71.43	flat => tv
3.36	4.34	45.45	minutes => took
3.00	4.17	37.50	line => long

Źródło: opracowanie własne

Współwystępujące pary słów: line–long, minutes–took, checkin–took mogą wskazywać na długi czas oczekiwania w kolejce na zameldowanie lub inne czynności hotelowe, co spowodowało, że wielu użytkowników nie wyraziło się pozytywnie o hotelu. Z kolei zdania zawierające pary słów: shop–coffee, flat–tv, slots–fun mogą świadczyć o zadowoleniu użytkowników z wyposażenia pokoi hotelowych oraz dostępności sklepu, kawiarni czy rozrywki, co zaowocowało pozytywnymi opiniami. Natomiast reguła monte–carlo, nie jest użyteczna, gdyż słowa te stanowią nazwę własną oznaczającą nazwę części budynku hotelowego.

WNIOSKI

Na podstawie przeprowadzonych badań można sformułować następujące wnioski.

Reguły asocjacyjne znajdują zastosowanie do odnajdywania zależności w danych tekstowych. Dane te jednak zazwyczaj mają obszerną reprezentację i nieustrukturyzowaną formę, co powoduje, że wymagają one odpowiedniego przygotowania. Opierając się na prawie Zipf'a, metodzie stop listy oraz przycinaniu, można zmniejszyć reprezentację tekstu, co jest konieczne, aby zastosować wybrany algorytm reguł asocjacyjnych. Analiza taka może być pomocna do odpowiedzi na pytanie jak postrzegany jest przez klientów dany obiekt hotelowy. Otrzymane wyniki analizy tekstu sugerują, że pary słów m.in.: long–line, took–minutes, checkin–took reprezentują negatywną klasę opinii. Pokazuje to, że w tym obszarze działalności hotelu konieczne są usprawnienia, które przyczynią się do podniesienia jakości obsługi. Z kolei teksty zawierające m.in. frazy: shop–coffee, tv–flat, slots–fun mogą posłużyć jako klasyfikator korzystnych dla hotelu opinii.

Przytoczone powyżej reguły stanowią cenne źródło informacji o opiniach i nastrojach klientów korzystających z usług hotelowych. Nie bez znaczenia pozostaje fakt, że otrzymane reguły w sposób niezwykle przystępny ujmują najważniejsze prawidłowości w bardzo dużym zbiorze opinii.

BIBLIOGRAFIA

- Agrawal R., Srikant R. (1994) Fast Algorithms for Mining Association Rules, IBM Research Report RJ9839, IBM Almaden Research Center San Jose, California.
- Amir A., Auman Y., Feldman R., Fresko M. (2005) Maximal Association Rules: A Tool for Mining Associations in Text, *Journal of Int. Information Systems*, str. 333–345.
- Borycki Ł., Sołdacki P. (2002) Automatyczna klasyfikacja tekstów, in: *Mat. III Krajowej Konferencji: Multimedialne i Sieciowe Systemy Informacyjne*, str. 473-481.
- Ganesan K., Zhai C. (2011) Opinion-Based Entity Ranking, *Information Retrieval*. [Online <http://archive.ics.uci.edu/ml/datasets/OpinRank+Review+Dataset>]
- Gunther R., Levitin L., Shapiro B., Wagner P. (1996) Zipf's law and the effect of ranking on probability distributions, *Int. Journal of Theoretical Physics*, 35(2), str. 395-417.
- Han J., Kamber M. (2001) *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publisher.
- Larose D. (2006) *Odkrywanie wiedzy z danych*, Wydawnictwo Naukowe PWN, Warszawa.
- Morzy, T. (2013). *Eksploracja danych*. Wydawnictwo Naukowe PWN, Warszawa.
- Pasztyła A. (2005) Przykład badania wzorców zachowań klientów za pomocą analizy koczowej, *Data mining: poznaj siebie i swoich klientów (pub. elektr.)*. Statsoft, Kraków.
- Rajaraman A., Ullman J.D. (2012) *Data Mining. Mining of Massive Datasets*. Cambridge University Press, New York.
- Słownik przyimków w języku angielskim (2013)
[Online <http://www.englishclub.com/grammar/prepositions-list.htm> (2013-12-30)]
- Słownik spójników w języku angielskim (2013) [Online <http://www.english-grammar-revolution.com/list-of-conjunctions.html> (2013-12-30)]
- Słownik wykrzykników w języku angielskim (2013)
[Online <http://www.vidarholen.net/contents/interjections> (2013-12-30)]
- Tan P., Steinbach M., Kumar V. (2005) *Introduction to Data Mining*, Addison-Wesley, Boston.
- Ward M. (1997) *50 najważniejszych problemów zarządzania*, Wydawnictwo Profesjonalnej Szkoły Biznesu, Kraków.

APPLICATION OF TEXT MINING TECHNIQUES FOR THE CUSTOMER REVIEWS ANALYSIS

Abstract: This paper presents application of one of data mining techniques – association rules to analyze customer reviews, based on the data gathered at one of the American hotels. The application of association rules is due to the large volume of available review data and the fact that the rules can be presented in a very clear and meaningful way. The study resulted in a number of interesting rules that can be a valuable source of information about the quality of services and the perception of the hotel by the clients.

Keywords: text mining, association rules, customer reviews

WYBRANE TESTY STATYSTYCZNE DLA WARTOŚCI NIETYPOWYCH I ICH ZASTOSOWANIE W ANALIZACH EKONOMETRYCZNYCH

Dorota Pekasiewicz

Katedra Metod Statystycznych Uniwersytet Łódzki

e-mail: pekasiewicz@uni.lodz.pl

Streszczenie: Wykrywanie obserwacji nietypowych w próbie losowej stanowi ważne zagadnienie w analizach statystycznych. Jednym ze sposobów badania próby od kątem istnienia wartości odstających jest stosowanie testów statystycznych opartych na statystykach ekstremalnych, do których należą: test Grubbsa i jego uogólnienie, test Dixona oraz testy oparte na asymptotycznych rozkładach minimum i maksimum z próby. Granicznymi rozkładami statystyk ekstremalnych są, w zależności od klasy rozkładu analizowanej zmiennej, rozkład Gumbela, Frecheta lub Weibulla. W artykule, oprócz rozważań teoretycznych, przedstawiono zastosowania wybranych testów do weryfikacji hipotez o wartościach nietypowych przy konstrukcji modeli ekonometrycznych.

Słowa kluczowe: statystyki ekstremalne, minimum, maksimum, wartość nietypowa, rozkład Gumbela

WSTĘP

W konstrukcji modeli ekonometrycznych istotne znaczenie ma wykrywanie i ewentualnie eliminacja wartości nietypowych. W modelach regresji z jedną zmienną objaśniającą najprostszym sposobem ich wyznaczenia jest analiza rozrzutu obserwacji (diagramu korelacyjnego), natomiast w przypadku regresji wielu zmiennych wykresów korelacyjnych par wszystkich zmiennych. Miernikami nietypowości są również reszty regresji oraz standaryzowane reszty regresji.

Ważną grupę metod służących wykrywaniu obserwacji odstających stanowią testy statystyczne. Wykorzystują one pojęcia statystyk ekstremalnych i ich własności, w szczególności twierdzenia o granicznych rozkładach minimum i maksimum z próby. Do testów weryfikujących hipotezy o istnieniu wartości

odstających wykorzystujących wartości ekstremalne należą m.in. test Grubbsa i jego uogólnienie oraz test Dixona. Informacja o klasie rozkładu statystyki minimum, czy maksimum rozważanej zmiennej losowej może być również wykorzystana przy konstrukcji statystyki testowej. Dla rozkładów o tzw. cienkich ogonach, takich jak rozkład normalny, logarytmiczno-normalny, granicznymi rozkładami statystyk ekstremalnych są rozkład Gumbela lub odwrócony Gumbela, natomiast dla rozkładów o grubych ogonach, np. t -Studenta, granicznymi rozkładami maksimum i minimum są, odpowiednio, rozkład Frecheta i odwrócony Frecheta. Dla zmiennych losowych przyjmujących wartości ze skończonego przedziału liczb rzeczywistych rozkłady statystyk ekstremalnych są typu Weibulla.

TEST GRUBBSA I JEGO UOGÓLNIENIE

Niech X_1, X_2, \dots, X_n będzie próbą losową pochodzącą z populacji o rozkładzie normalnym, zaś x_1, x_2, \dots, x_n wartościami wylosowanej próby.

Rozważmy następującą hipotezę:

H_0 : w ciągu obserwacji x_1, x_2, \dots, x_n nie istnieją obserwacje nietypowe (o zbyt dużych lub zbyt małych wartościach).

Hipotezą alternatywną może być jedna z następujących hipotez:

H_1' : minimalna wartość realizacji x_1, x_2, \dots, x_n próby losowej jest wartością nietypową,

H_1'' : maksymalna wartość realizacji x_1, x_2, \dots, x_n próby losowej jest wartością nietypową.

Dla hipotez H_0 i H_1' statystyka testu Grubbsa [Grubbs 1969] wyraża się wzorem:

$$G_m = \frac{\bar{X} - X_{(1)}^{(n)}}{S}, \quad (1)$$

natomiast dla hipotez H_0 i H_1'' ma następującą postać:

$$G_M = \frac{X_{(n)}^{(n)} - \bar{X}}{S}, \quad (2)$$

gdzie $X_{(1)}^{(n)} = \min\{X_1, \dots, X_n\}$ oraz $X_{(n)}^{(n)} = \max\{X_1, \dots, X_n\}$.

Obszar krytyczny dla testu Grubbsa, przy przyjętym poziomie istotności α , wyznaczony jest przez wartość krytyczną:

$$G_\alpha = \frac{n-1}{\sqrt{n}} \frac{t_{(\alpha/n, n-2)}}{\sqrt{n-2 + t_{(\alpha/n, n-2)}^2}}, \quad (3)$$

gdzie $t_{(\alpha/n, n-2)}$ jest wartością krytyczną rozkładu t -Studenta o $n - 2$ stopniach swobody dla $\frac{\alpha}{n}$.

Hipotezę zerową odrzucamy na poziomie istotności α , gdy $G_M > G_\alpha$ lub $G_m > G_\alpha$.

Test Grubbsa rozstrzyga, czy wartość najmniejsza w próbie losowej lub wartość największa może być uznana za nietypową, odstającą od pozostałych elementów próby. Istnieje uogólnienie tego testu pozwalające wyznaczyć $r-1$ wartości nietypowych w r krokach testowania.

Procedurę weryfikacji hipotezy H_0 wobec $\sim H_0$ za pomocą uogólnionego testu Grubbsa przeprowadza się następująco.

W i -tym kroku, $i = 1, 2, \dots$, oblicza się:

- $\bar{X}_i = \frac{1}{n-i+1} \sum_{j \in L_i} X_j$, gdzie L_i jest zbiorem indeksów elementów próby zredukowanej, czyli powstałej poprzez pominięcie $i-1$ wartości,
- S_i - odchylenie standardowe wyznaczone w oparciu o zredukowaną próbę,
- statystykę testu:

$$G_i = \max_j \left(\frac{|X_j - \bar{X}_i|}{S_i} \right). \quad (4)$$

Statystykę tę porównuje się z wartością krytyczną:

$$G_{\alpha i} = \frac{n-i}{\sqrt{n-i+1}} \frac{t_{(p, n-i-1)}}{\sqrt{n-i-1 + t_{(p, n-i-1)}^2}}, \quad (5)$$

gdzie $t_{(p, n-i-1)}$ jest wartością odczytaną z tablic wartości krytycznych rozkładu t -

Studenta o $n - i - 1$ stopniach swobody dla wartości $p = \frac{\alpha}{n-i+1}$.

Spełnienie nierówności $G_i > G_{\alpha i}$ świadczy o istnieniu wartości nietypowej, którą pomijamy w kolejnym kroku. Testowanie hipotez kontynuujemy, aż do momentu otrzymania decyzji o braku podstaw do odrzucenia hipotezy zerowej. Jeżeli nastąpiło to w r krokach, to wyznaczonych zostało $r-1$ obserwacji nietypowych.

TEST DIXONA

Test Dixona jest kolejnym testem wykorzystującym wartości ekstremalne i weryfikującym hipotezy o istnieniu w próbie losowej X_1, X_2, \dots, X_n obserwacji nietypowych, przy założeniu, że pochodzi ona z populacji o rozkładzie normalnym [Dixon 1950, Chromiński, Tkacz 2010].

Dla hipotezy alternatywnej H_1' statystyka testu ma postać:

$$Q = \frac{X_{(2)}^{(n)} - X_{(1)}^{(n)}}{X_{(n)}^{(n)} - X_{(1)}^{(n)}}, \quad (6)$$

natomiast dla hipotezy alternatywnej H_1'' :

$$Q = \frac{X_{(n)}^{(n)} - X_{(n-1)}^{(n)}}{X_{(n)}^{(n)} - X_{(1)}^{(n)}}, \quad (7)$$

gdzie $X_{(i)}^{(n)}$ oznacza i -tą statystykę pozycyjną ($i = 1, 2, n-1, n$).

Wartości krytyczne dla testu Dixona są stabilizowane m.in. w pracy Verma, Quiroz-Ruiz [2006] dla $n \leq 100$.

TESTY OPARTE NA ASYMPTOTYCZNYCH ROZKŁADACH STATYSTYK EKSTREMALNYCH

Do konstrukcji testów weryfikujących hipotezy o istnieniu wartości nietypowych można wykorzystać informacje o granicznych rozkładach statystyk maksimum i minimum.

Niech X będzie populacją o rozkładzie normalnym $N(\mu, \sigma)$ oraz X_1, X_2, \dots, X_n próbą losową pochodzącą z tej populacji. Ponadto, niech:

$$Z_n = \max_{i \leq n} \left(\frac{X_i - \mu}{\sigma} \right), \quad (8)$$

$$W_n = \min_{i \leq n} \left(\frac{X_i - \mu}{\sigma} \right). \quad (9)$$

Standaryzowane zmienne Z_n i W_n mają, odpowiednio, rozkład Gumbela o dystrybuancie $H(z) = \exp(-\exp(-z))$ dla $z \in R$ i odwrócony rozkład Gumbela o dystrybuancie $L(z) = 1 - \exp(-\exp z)$, czyli

$$\lim_{n \rightarrow \infty} P \left\{ \frac{Z_n - a_n}{b_n} < z \right\} = H(z), \quad (10)$$

oraz

$$\lim_{n \rightarrow \infty} P \left\{ \frac{W_n - c_n}{b_n} < z \right\} = L(z), \quad (11)$$

gdzie $a_n = (2 \ln n)^{1/2} - \frac{\ln \ln n + \ln 4\pi}{2(2 \ln n)^{1/2}}$, $b_n = (2 \ln n)^{-1/2}$ i $c_n = -a_n$ [Czekała 2001].

Zatem statystyka testu weryfikującego hipotezę H_0 wobec H_1 '' wyraża się wzorem:

$$T_M = \frac{Z_n - a_n}{b_n}, \quad (12)$$

a wartość krytyczna wyznaczona z równania $\exp(-\exp(-z)) = 1 - \alpha$ ma postać:

$$t_{M,\alpha} = -\ln \left(\ln \frac{1}{1 - \alpha} \right). \quad (13)$$

Jeśli wartość t_M statystyki T_M spełnia nierówność $t_M \geq t_{M,\alpha}$, to na poziomie istotności α odrzucamy hipotezę H_0 na korzyść hipotezy alternatywnej mówiącej, że w próbie X_1, X_2, \dots, X_n istnieje obserwacja nietypowa o zbyt dużej wartości.

Dla hipotez H_0 i H_1 ' statystyka testu wyrażona jest wzorem:

$$T_m = \frac{\hat{W}_n + a_n}{b_n}. \quad (14)$$

Wartość t_m statystyki T_m porównujemy z wartością krytyczną $t_{m,\alpha}$:

$$t_{m,\alpha} = \ln \left(\ln \frac{1}{1 - \alpha} \right). \quad (15)$$

Jeśli $t_m \geq t_{m,\alpha}$, to nie ma podstaw do odrzucenia hipotezy H_0 , zaś jeśli $t_m < t_{m,\alpha}$, to odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej mówiącej, że istnieje wartość nietypowa w próbie o zbyt małej wartości.

Analogicznie skonstruować można testy weryfikujące hipotezy o istnieniu wartości odstających w próbie wylosowanej z populacji X o rozkładzie logarytmiczno-normalnym i logistycznym, gdyż odpowiednio wystandaryzowane statystyki maksimum i minimum mają również rozkład Gumbela lub odwrócony Gumbela. W przypadku rozkładu gamma, w szczególności wykładniczego oraz rozkładu Pareto statystyka maksimum ma rozkład Gumbela, a minimum Weibulla, natomiast dla rozkładu Cauchy'ego, t -Studenta odpowiednio standaryzowane statystyki ekstremalne mają rozkład Frecheta i odwrócony rozkład Frecheta [Castillo i in. 2004].

WYKRYWANIE I ELIMINACJA OBSERWACJI NIETYPOWYCH W PRZYPADKU MODELI EKONOMETRYCZNYCH

Rozważmy model ekonometryczny postaci:

$$Y_t = f(\mathbf{X}_t, \beta_0) + \varepsilon_t, \quad (t = 1, 2, \dots, n) \quad (16)$$

gdzie: (Y_t, \mathbf{X}_t) jest $(k+1)$ -elementowym wektorem losowym,

Y_t jest zmienną objaśnianą,

$\mathbf{X}_t = (X_{1t}, X_{2t}, \dots, X_{kt})$ jest wektorem zmiennych objaśniających,

β_0 jest nieznanym wektorem parametrów strukturalnych,

ε_t jest składnikiem losowym, $\varepsilon_t \sim N(0, \sigma)$.

Analiza wielkości składnika losowego pozwala na wyznaczenie wartości nietypowych przy ustalaniu linii regresji i ewentualne ich pominięcie przy budowie modelu. Ze względu na często zakładaną normalność rozkładu zmiennej ε_t korzysta się z faktu, że standaryzowane statystyki maksimum i minimum składnika losowego mają rozkłady graniczne będące, odpowiednio, rozkładem Gumbela i odwróconym Gumbela.

Oszacowaniami statystyk $Z_n = \max_{t \leq n} \left(\frac{\varepsilon_t}{\sigma} \right)$ oraz $W_n = \min_{t \leq n} \left(\frac{\varepsilon_t}{\sigma} \right)$ na podstawie ciągu obserwacji $(y_1, x_{1,1}, x_{2,1}, \dots, x_{k,1}), \dots, (y_n, x_{1,n}, x_{2,n}, \dots, x_{k,n})$ są statystyki $\hat{Z}_n = \max_{t \leq n} \left(\frac{e_t}{\sigma} \right)$, $\hat{W}_n = \min_{t \leq n} \left(\frac{e_t}{\sigma} \right)$, gdzie $e_t = y_t - f(\mathbf{x}_t, \beta_0)$.

W praktyce, w wielu przypadkach wariancja składnika losowego σ^2 jest nieznaną i szacuje się ją za pomocą estymatora:

$$s_n^2 = \frac{1}{n} \sum_{i=1}^n (y_i - f(\mathbf{x}_i, \hat{\beta}_0))^2. \quad (17)$$

Zastosowanie testów ze statystykami (12) i (14) sprowadza się do procedury sprawdzenia dla obserwacji (x_t, y_t) nierówności:

$$\frac{\frac{e_t - a_n}{s_n}}{b_n} > -\ln \left(\ln \frac{1}{1 - \alpha} \right) \quad \text{lub} \quad \frac{\frac{e_t + a_n}{s_n}}{b_n} < \ln \left(\ln \frac{1}{1 - \alpha} \right).$$

Problematykę wyznaczania wartości nietypowych można zilustrować na przykładzie konstrukcji modelu opisującego zależność dziennych stóp zwrotu

z inwestycji w akcje mBanku i PKO BP od stóp zwrotu z indeksu WIG 30 wprowadzonego na GPW w Warszawie dnia 23.09.2013 r. Rozważano następujący okres 23.09.2013 – 9.06.2014 r. Obserwacje z tego okresu stanowiły próbę 173 elementową.

Dla inwestycji w akcje mBanku otrzymano następującą funkcję (przy założeniu niezależności i normalności stóp zwrotu) :

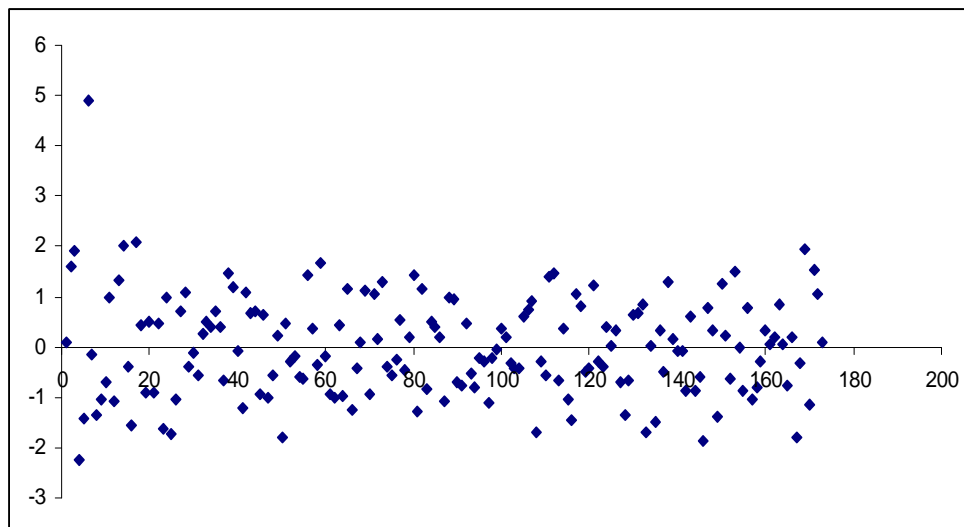
$$\hat{y}_t = 1,2966x_t + 0,0009, \quad (18)$$

gdzie y_t - stopa zwrotu z inwestycji w akcje mBanku,

x_t - stopa zwrotu z indeksu WIG 30.

Zastosowanie testu opartego na granicznych rozkładach statystyk ekstremalnych związane jest z następującymi obliczeniami: $\hat{Z}_n = 4,8942$, $\hat{W}_n = -2,2279$, $a_n = 2,5608$, $b_n = 0,3115$. Statystyki (12) i (14) są równe: $t_M = 7,4909$, $t_m = 1,0690$, a wartości krytyczne wyznaczone dla poziomu istotności $\alpha = 0,05$: $t_{M,\alpha} = 2,9702$, $t_{m,\alpha} = -2,9702$. Ponieważ $t_M > t_{M,\alpha}$ parę wartości (0,0127; 0,0974), która pojawiła się 1.10.2013 r., należy uznać za nietypową przy konstrukcji modelu ekonometrycznego opisującego zależność stopy zwrotu z inwestycji w akcje mBanku od stopy zwrotu z indeksu giełdowego WIG 30 (p -value=0,0006). Na rysunku 1 przedstawiono standaryzowane reszty modelu opisującego zależność między rozważanymi wielkościami..

Rysunek 1. Standaryzowane reszty modelu (18)



Źródło: opracowanie własne

Ze względu na normalność rozkładu składnika losowego możliwe jest również zastosowanie testu Grubbsa. Związane jest to z obliczeniami: $e_{(1)}^{(n)} = \min\{e_{t_1}, \dots, e_{t_n}\} = -0,0364$, $e_{(n)}^{(n)} = \max\{e_{t_1}, \dots, e_{t_n}\} = 0,08$, $\bar{e}_t = 0$, $S_{e_t} = 0,0164$ i $G_M = 4,8947$, $G_m = 2,2273$ oraz $G_\alpha = 3,3878$ dla $\alpha = 0,05$. Na podstawie otrzymanych wyników wnioskuje się, że wartość 0,08 (standaryzowana reszta 4,8942) jest zbyt duża w porównaniu z pozostałymi co oznacza, że para wartości (0,0127; 0,0974) jest nietypowa, a to potwierdza decyzję podjętą na podstawie wcześniej stosowanego testu.

Analizując zależność wielkości stóp zwrotu z akcji PKO BP od stóp zwrotu z indeksu WIG 30 otrzymano następujący model:

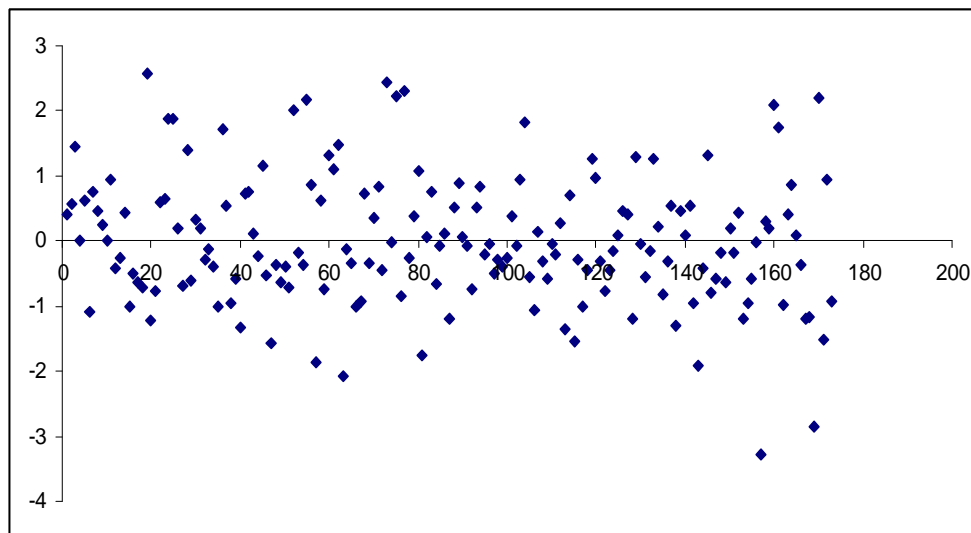
$$\hat{y}_t = 0,9143x_t + 0,0005, \quad (19)$$

gdzie y_t - stopa zwrotu z inwestycji w akcje PKO BP,

x_t - stopa zwrotu z indeksu WIG 30.

W tym przypadku: $\hat{Z}_n = 2,5827$, $\hat{W}_n = -3,2780$, $a_n = 2,5608$, $b_n = 0,3115$. Stąd $t_M = 0,0703$, $t_m = -2,3022$ oraz $t_{M,\alpha} = 2,9702$, $t_{m,\alpha} = -2,9702$ dla $\alpha = 0,05$. Zatem na poziomie istotności 0,05 brak jest podstaw do odrzucenia hipotezy zerowej. Ponieważ $p\text{-value} = 0,0952$, więc przyjmując $\alpha = 0,1$ parę wartości (0,0005; -0,0276) pochodzącą z dn. 16.05.2014 r. można uznać za nietypową ($t_{M,\alpha} = 2,2504$, $t_{m,\alpha} = -2,2504$). Graficzna prezentacja standaryzowanych reszt dla tego modelu przedstawiona jest na rysunku 2.

Rysunek 2. Standaryzowane reszty modelu (19)



Źródło: opracowanie własne

Przy zastosowaniu testu Grubbsa otrzymujemy:
 $e_{(1)}^{(n)} = \min\{e_{t_1}, \dots, e_{t_n}\} = -0,0285$, $e_{(n)}^{(n)} = \max\{e_{t_1}, \dots, e_{t_n}\} = 0,0224$, $\bar{e}_t = 0$,
 $S_{e_t} = 0,0087$ i $G_M = 2,5827$, $G_m = 3,2779$, $G_\alpha = 3,3878$ dla $\alpha = 0,05$ oraz
 $G_\alpha = 3,2045$ dla $\alpha = 0,1$. Stąd wnioskować można, na poziomie istotności 0,1, że wartość -0,0285 (standaryzowana reszta -3,278) jest zbyt mała, czyli para wartości (0,0005; -0,0276) jest nietypowa dla skonstruowanego modelu.

UWAGI KOŃCOWE

Wyznaczanie wartości nietypowych możliwe jest w analizach zmiennych losowych jednowymiarowych jak i wielowymiarowych, przy konstrukcji różnego rodzaju modeli ekonometrycznych. Zastosowanie testów Grubbsa, czy Dixona wiąże się z koniecznością sprawdzenia, czy zmienna losowa ma rozkład normalny. Testy wykorzystujące graniczne rozkłady statystyk ekstremalnych mają szersze zastosowanie. Znajomość rozkładu badanej zmiennej losowej, standaryzacja statystyk ekstremalnych i wykorzystanie twierdzeń mówiących o klasie ich rozkładów pozwalają konstruować tego typu testy.

W pracy przedstawiono testy statystyczne dla wartości odstających oparte na statystykach ekstremalnych. Inne testy np. Humpela, kwartyłowy wykorzystujące statystyki pozycyjne takie jak mediana, czy kwartyle można znaleźć w pracy Budka i in. [2013].

Rozważane testy znajdują zastosowanie w badaniach zależności między dwiema lub większą liczbą zmiennych. Założenie o normalności rozkładu składnika losowego sprawia, że rozkładami granicznymi statystyk ekstremalnych są rozkład Gumbela i odwrócony Gumbela. Ich wykorzystanie, zarówno przy określaniu zależności wielkości stopy zwrotu akcji mBanku od stopy zwrotu indeksu WIG 30, jak i przy wyznaczaniu funkcji opisującej zależność wielkości stopy zwrotu akcji PKO BP od stopy zwrotu z indeksu WIG 30 prowadziło do uzyskania rezultatów identycznych z rezultatami testu Grubbsa. Porównywalne wyniki zachęcają do analiz testów opartych na granicznych rozkładach statystyk ekstremalnych dla zmiennych o innych rozkładach niż normalny, dla których zastosowanie testu Grubbsa, czy Dixona nie jest możliwe.

BIBLIOGRAFIA

- Budka A., Kayzer D., Pietruczuk K., Szoszkiewicz K. [2013], Zastosowanie wybranych procedur do wykrywania obserwacji nietypowych w ocenie jakości rzek, PAN nr3/II, Warszawa.
- Castillo E., Hadi A. S., Balakrishnan N., Sarabia J. M. [2004], Extreme value and related models with application in engineering and science, Wiley Interscience, A. John Wiley & Sons, Inc. New Jersey.
- Chromiński K., Tkacz M. [2010], Comparison of outlier detection methods in biometric data, Journal of Medical Informatics and Technologies, 16, s. 89 – 94.
- Czekała M. [2001], Statystyki pozycyjne w modelowaniu ekonometrycznym. Wybrane problemy Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.
- Dixon W. J. [1950], Analysis of extreme values, Annals of Mathematical Statistics, 4, 488-506.
- Grubbs F. [1969], Procedures for detecting outlying observations in samples, Technometrics 11, s. 1-21.
- Verma S.P., Quiroz-Ruiz A. [2006], Critical values for six Dixon tests for outliers in normal samples up to sizes 100 and applications in science and engineering. Revista Mexicana de Ciencias Geológicas, 23(2), s.133 – 161.

CHOSEN STATISTICAL TESTS FOR OUTLIERS AND THEIR APPLICATION IN ECONOMETRIC ANALYSIS

Abstract: The problem of the existence of outliers in the sample is an important issue in statistical surveys. One of the methods of outliers detection is the application of statistical tests based on extreme statistics. Grubbs test and its generalization, Dixon test and tests based on asymptotic distributions of minimum and maximum (Gumbel, Frechet, Weibull distributions) belong to group of these tests. In the paper, besides the theoretical considerations the application of selected tests, used to verify the hypothesis of outliers in the construction of econometric models, is presented.

Keywords: extreme statistics, minimum, maximum, outlier, Gumbel distribution

OCENA POTENCJAŁU PAŃSTW UNII EUROPEJSKIEJ DO GENEROWANIA INNOWACJI Z ZASTOSOWANIEM ANALIZY SKUPIEŃ

Elżbieta Roszko–Wójtowicz
Katedra Statystyki Ekonomicznej i Społecznej
Uniwersytet Łódzki
e-mail: eroszko33@gmail.com

Streszczenie: W ramach wielowymiarowej analizy porównawczej badacze najczęściej stają przed problemem dotyczącym wyboru odpowiedniego sposobu łączenia wielu obiektów w grupy, charakteryzujące się podobnymi właściwościami pod względem wybranych zmiennych diagnostycznych. Niniejsza praca stanowi propozycję zastosowania hierarchicznych metod aglomeracji do poszukiwania związków pomiędzy poszczególnymi państwami członkowskimi UE ze względu na możliwości generowania innowacji. W opisywanym przypadku najkorzystniejszą do poszukiwania związków pomiędzy obiektami, odnośnie zdefiniowanych w pracy zdolności do generowania innowacji, wydaje się być metoda pełnego połączenia i metoda Warda dające podział zestawu obiektów na trzy skupienia.

Słowa kluczowe: analiza skupień, potencjał innowacyjny, metody aglomeracyjne, kryterium Mojena, kryterium Wisharta, kraje członkowskie UE

WSTĘP

Przedmiotem szczególnej uwagi ekonomistów na przestrzeni ostatniej dekady są czynniki kształtujące rozwój innowacyjny z punktu widzenia ich wpływu i znaczenia dla wzrostu gospodarczego. Co więcej, aktywność innowacyjna kraju nie jest działaniem autonomicznym, a zależy od wielu czynników związanych z obszarem edukacji, wiedzy oraz działalności badawczo-rozwojowej [Fagerberg 1988].

Tempo wzrostu innowacyjności UE spada, a opóźnienie w stosunku do światowych liderów (USA, Japonia, Korea Południowa) utrzymuje się.¹ W związku z tym strategia na rzecz wzrostu gospodarczego i zatrudnienia „Europa 2020” koncentruje się na stymulowaniu działań innowacyjnych i usuwaniu barier, które ograniczają ich wejście na rynek.²

Wsparcie innowacyjnego rozwoju regionów UE wymaga szczegółowego badania aktywności innowacyjnej poszczególnych gospodarek. Metodologia OECD [Podręcznik OSLO 2005] podaje, że badając procesy innowacyjne należy przede wszystkim zwrócić uwagę na aktywność innowacyjną przedsiębiorstw. Dla stymulowania innowacyjności jednak nie bez znaczenia pozostają również relacje z otoczeniem, zatem z innymi firmami i instytucjami otoczenia biznesu.

Czynniki, które mają największy wpływ na potencjał innowacyjny kraju związane są w głównej mierze z wykształceniem społeczeństwa, wyposażeniem w infrastrukturę teleinformatyczną, ogólną sytuacją na rynku pracy czy też aktywnością podmiotów w zakresie badań i rozwoju. W związku z tym do analizy państw Unii Europejskiej ze względu na możliwości do generowania innowacji wybrany został zestaw 15 zmiennych sklasyfikowanych w 4 obszary.

Pomiar innowacyjności obejmuje szeroką grupę wskaźników z zakresu nauki, techniki i innowacji (N + T + I) i charakteryzuje się następującymi własnościami [Markowska 2012]:

- Przedmiot pomiaru
 - Ludzie (specjaliści, badacze itp.)
 - Zasoby finansowe (źródła finansowania, rodzaje inwestycji)
 - Wiedza skodyfikowana (publikacje, patenty, kwalifikacje)
 - Wiedza ucieleśniona (urządzenia, komponenty, dobra trwałego użytku)
- Przestrzeń, w której dokonywany jest pomiar (organizacja, przestrzeń w sensie geograficznym)
- Typ działalności (dyscyplina, pole badawcze, dziedzina technologii, branża przemysłu itp.)
- Skala pomiaru (mikro – organizacja, mezo – dyscyplina, branża, gospodarka regionu, makro – gospodarka kraju i większe terytorium)
- Typ pomiaru (parametr zasobu – mierzy rozmiary lub parametr relacji – mierzy zależność lub przepływ)

Mając na uwadze wielowymiarowość zagadnienia celem głównym artykułu jest klasyfikacja państw Unii Europejskiej ze względu na czynniki kształtujące potencjał gospodarki do tworzenia innowacji, ze szczególnym uwzględnieniem pozycji Polski.

¹ http://europa.eu/rapid/press-release_IP-12-102_pl.htm

² http://ec.europa.eu/news/science/120208_pl.htm

Cel szczegółowy 1 – wybór, prezentacja i omówienie zmiennych pozwalających na ocenę innowacyjności gospodarek UE.

Cel szczegółowy 2 – omówienie wybranych metod grupowania obiektów oraz przedstawienie praktycznego ich zastosowania.

ANALIZA SKUPIEŃ – ZAŁOŻENIA TEORETYCZNE

Wielowymiarowa analiza porównawcza [Kisielińska, Stańko 2009] zajmuje się badaniem pewnego zbioru obiektów, które zostały opisane za pomocą wielu cech. WAP obejmuje szereg metod statystycznych dzięki którym możliwe jest dokonanie jednoczesnej analizy poziomu przynajmniej dwóch zmiennych opisujących każdą badaną jednostkę. Informacje o obiektach umieszczane są w macierzy obserwacji, której wiersze odpowiadają obiektom, zaś kolumny zebranych cechom.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

gdzie: n – liczba obiektów, p – liczba cech.

Wśród wielu metod wielowymiarowej analizy danych w badaniach ekonomiczno-społecznych szerokie zastosowanie znajdują:

- Bezwzorcowa klasyfikacja obiektów;
- Wzorcowa klasyfikacja obiektów;
- Liniowe porządkowanie obiektów;
- Regresja wieloraka;
- Analiza czynnikowa;
- Metoda składowych głównych.

Przedmiotem zainteresowania Autorki w niniejszym opracowaniu będzie zastosowanie analizy skupień do oceny państw członkowskich UE z punktu widzenia ich możliwości do generowania innowacji, co znalazło odzwierciedlenie w sformułowanych celach pracy.

Analiza skupień to jedna z najbardziej znanych metod eksploracji danych. Celem analizy skupień jest dokonanie grupowania obiektów w oparciu o pewien zestaw cech, które te obiekty charakteryzują. Istotą analizy skupień jest taki podział zbioru wyjściowego, aby jednostki jednej grupy były jak najbardziej do siebie podobne ze względu na wyróżniony zestaw zmiennych, a jednocześnie jak najbardziej odmienne od obiektów, które znajdują się w pozostałych skupieniach [Gatnar, Walesiak 2004]. Metoda nie dostarcza narzędzi do optymalnego doboru zmiennych opisujących badane jednostki. Wybór zmiennych i utworzenie

ostatecznego zbioru cech diagnostycznych spoczywa na badaczu i jest ważnym elementem analizy skupień gdyż wywiera istotny wpływ na finalne rozwiązanie.

- W literaturze przedmiotu istnieje ogromna liczba metod grupowania i zarazem mnogość propozycji odnoszących się do ich klasyfikacji. Ogół metod taksonomii numerycznej można podzielić na [Frątczak 2009]:
 - Metody hierarchiczne;
 - Metody optymalizacyjno-iteracyjne;
 - Metody obszarowe.

Przedmiotem zainteresowania Autorki w niniejszym opracowaniu są jedynie metody hierarchiczne, pozwalające na utworzenie pełnej hierarchii skupień, dla której współczynnik podobieństwa wzrasta monotonicznie. Skupienia powstające w kolejnych etapach (wyższego rzędu) zawierają w sobie rozłączne skupienia powstałe na etapach wcześniejszych (niższego rzędu). Struktura skupień tworzona jest w formie wykresu drzewkowego, zwanego dendrogramem. Wyróżnia się dwie grupy metod hierarchicznych [Grabiński, Wydymus, Zeliaś 1989]:

- procedury aglomeracyjne (sekwencyjne, indukcyjne) (ang. *agglomerative*)
- procedury deglomeracyjne (podziałowe) (ang. *divisive*)

Metody aglomeracyjne – początkowo przyjmuje się, że każdy z obiektów stanowi odrębną grupę, w kolejnych etapach łączy się w podzbiory podgrupy najbardziej do siebie podobne, aż do uzyskania finalnego rozwiązania w postaci jednego skupienia, które zawiera wszystkie elementy. W ten sposób powstaje uporządkowane zestawienie podziałów na segmenty. Wśród najczęściej stosowanych metod aglomeracyjnych znajdują się następujące metody: najbliższego sąsiedztwa, najdalszego sąsiedztwa, średnich połączeń oraz metoda Warda. Każdy z wymienionych algorytmów posiada odmienny sposób określania najmniejszej odległości pomiędzy łączonymi skupieniami.

W metodach podziałowych algorytm postępowania jest odwrotny do tego, który występuje w metodach aglomeracyjnych. Zatem punktem wyjścia jest zbiór obiektów, który traktowany jest jako jedna grupa, a w kolejnych krokach dzieli się jedno ze skupień na dwie części, a liczba skupień ulega zwiększeniu o jeden. W konsekwencji po $n-1$ krokach uzyskuje się n skupień jednoelementowych. Wyniki uzyskane za pomocą metod podziałowych mają analogiczną strukturę do tych uzyskanych za pomocą metod aglomeracyjnych.

Ogólny schemat postępowania w analizie skupień składa się z kilku następujących po sobie kroków. Początek stanowi dobór jednostek do badania, po którym następuje dobór cech diagnostycznych i ocena ich przydatności z punktu widzenia prowadzonej analizy z wykorzystaniem odpowiednich narzędzi statystycznych. Etapy końcowe to przeprowadzenie właściwej analizy skupień

w oparciu o wybrane algorytmy grupowania oraz ustalenie optymalnej liczby klas. [szerzej na ten temat w Milligan, Cooper 1987, Gordon 1999, Mikulec 2012].

WYBÓR ZMIENNYCH DIAGNOSTYCZNYCH

Obiektami poddanymi analizie w niniejszym opracowaniu będą poszczególne kraje członkowskie UE – 28 krajów, nazywane również jednostkami analizy. Cechy diagnostyczne natomiast, to zestaw 15 zmiennych wybranych z bazy Eurostat do opisu obiektów ze względu na ich zdolność do generowania innowacji. Początkowy zestaw składa się z następujących zmiennych diagnostycznych:

- X1 – Wydatki rządowe na działalność badawczo-rozwojową
- X 2 – Wskaźnik zatrudnienia (grupa wieku 20-64 lata)
- X 3 – Zatrudnienie w sektorze wysokich technologii i wiedzochłonnym
- X 4 – Ludność z wykształceniem wyższym
- X 5 – Poziom dostępu do Internetu w gospodarstwach domowych
- X6 – Studenci w grupie wieku 15-24 lat
- X 7 – Wydatki publiczne na edukację
- X 8 – Patenty przyznane przez Biuro Patentów i Znaków Towarowych Stanów Zjednoczonych
- X 9 – Zgłoszenia patentów do EUP
- X 10 – Korzystanie z Internetu przez mieszkańców
- X11 – Kompetencje w zakresie obsługi komputera
- X 12 – Przedsiębiorstwa zatrudniające specjalistów ICT
- X 13 – Wydatki przedsiębiorstw na B+R
- X 14 – Wydatki przedsiębiorstw na B+R (% PKB)
- X 15 – Kadra sektora B+R

W ocenie krajów członkowskich UE do generowania innowacji każdą z zaprezentowanych zmiennych diagnostycznych będziemy traktować jako stymulantę, oznacza to, że wzrost jej wartości ma pozytywny wpływ na badane zjawisko. W przypadku wystąpienia zmiennych określonych jako destymulanty, tzn. takich, których wzrost wartości ma niekorzystny wpływ na badane zjawisko, należy dokonać ich przekształcenia stosując odpowiednią formułę. Wybranych 15 zmiennych diagnostycznych podzielono na 4 obszary analizy. (Tabela 1)

Tabela 1. Obszary analizy pokryte wybranymi cechami diagnostycznymi – początkowy i ostateczny zestaw zmiennych diagnostycznych

Obszar	Początkowy zestaw zmiennych diagnostycznych	Ostateczny zestaw zmiennych diagnostycznych
1 – Finansowanie	X1 – Wydatki rządowe na działalność badawczo-rozwojową X 7 – Wydatki publiczne na edukację X 14 – Wydatki przedsiębiorstw na B+R (%PKB) X 13 – Wydatki przedsiębiorstw na B+R	X1 – Wydatki rządowe na działalność badawczo-rozwojową X 7 – Wydatki publiczne na edukację X 14 – Wydatki przedsiębiorstw na B+R (% PKB) X 13 – Wydatki przedsiębiorstw na B+R
2 – Rynek pracy	X 2 – Wskaźnik zatrudnienia (grupa wieku 20-64 lata) X 3 – Zatrudnienie w sektorze wysokich technologii i wiedzochłonnym X 12 – Przedsiębiorstwa zatrudniające specjalistów ICT X 15 – Kadra sektora B+R	X 2 – Wskaźnik zatrudnienia (grupa wieku 20-64 lata) X 3 – Zatrudnienie w sektorze wysokich technologii i wiedzochłonnym X 12 – Przedsiębiorstwa zatrudniające specjalistów ICT X 15 – Kadra sektora B+R
3 – Edukacja	X 4 – Ludność z wykształceniem wyższym X 5 – Poziom dostępu do Internetu w gospodarstwach domowych X6 – Studenci w grupie wieku 15-24 lata X 10 – Korzystanie z Internetu przez mieszkańców X11 – Kompetencje w zakresie obsługi komputera	X 4 – Ludność z wykształceniem wyższym X 5 – Poziom dostępu do Internetu w gospodarstwach domowych X6 – Studenci w grupie wieku 15-24 lata X 10 – Korzystanie z Internetu przez mieszkańców X11 – Kompetencje w zakresie obsługi komputera
4 – Patenty	X 8 – Patenty przyznane przez Biuro Patentów i Znaków Towarowych Stanów Zjednoczonych X 9 – Zgłoszenia patentów do EUP	X 8 – Patenty przyznane przez Biuro Patentów i Znaków Towarowych Stanów Zjednoczonych X 9 – Zgłoszenia patentów do EUP

Źródło: opracowanie własne

Ostateczny wybór cech ze wstępnie zaprezentowanej listy zmiennych (Tabela 1) opisujących gospodarki unijne pod względem zdolności do tworzenia innowacji dokonany został na podstawie następujących kryteriów [por. Ostasiewicz 1998, Malina, Zieliński 1996, Kunasz 2006]:

1. Kryterium uniwersalności – cechy diagnostyczne, opisujące badane zjawisko muszą być mierzalne, oraz powinny być źródłem istotnych informacji w obszarze poddawanym analizie.
2. Kryterium zmienności – analizowane cechy powinny dostatecznie różnicować badane obiekty, w tym celu dla analizowanych cech wyznacza się wartość współczynnika zmienności.
3. Kryterium stopnia skorelowania – oceny stopnia skorelowania zmiennych dokonuje się najczęściej za pomocą współczynnika korelacji liniowej Pearsona

lub współczynnika rang Spearmana. Wysoka wartość współczynnika korelacji świadczy o silnej zależności korelacyjnej między dwoma cechami diagnostycznymi i oznacza, że są one nośnikiem podobnych informacji.

4. Kryterium ważności – przyjmuje się, że cechy są ważne, jeśli trudno osiągnąć wysokie wartości. Celem sprawdzenia ważności cech i tym samym eliminacji cech nieważnych oblicza się wartości współczynnika asymetrii.

Weryfikacji kryterium zmienności dokonujemy na podstawie wyliczonych wartości współczynników zmienności dla poszczególnych cech diagnostycznych. Uzyskane wartości porównujemy z arbitralnie zadaną wartością krytyczną wynoszącą w tym przypadku $V_s = 10\%$. Z początkowego zestawu zmiennych eliminacji podlegają te, dla których wyznaczone wartości współczynnika zmienności są mniejsze od przyjętej wartości progowej. W niniejszym przypadku współczynnik zmienności dla 14 spośród 15 analizowanych zmiennych jest wyższy niż 10%, co oznacza, iż wybrane cechy diagnostyczne mają dużą zdolność do różnicowania krajów członkowskich UE ze względu na potencjał do generowania innowacji. Wyjątek stanowi wartość wskaźnika zatrudnienia w grupie wieku 20-64 lata, gdzie współczynnik zmienności osiąga poziom 9,5%. Niemniej jednak, ze względu na przydatność merytoryczną tej zmiennej z punktu widzenia omawianej tematyki nie została ta cecha wyeliminowana ze zbioru cech diagnostycznych. Rozkład żadnej z wybranych zmiennych nie charakteryzuje się silną asymetrią lewostronną (wartość współczynnika asymetrii nie jest niższa od -0,7). Zatem, można przyjąć, iż każda z wybranych cech diagnostycznych dostatecznie różnicuje obiekty, nie osiągając tym samym wysokiego stopnia nasycenia.

Celem dokonania oceny stopnia skorelowania wybranego zestawu zmiennych najczęściej oblicza się współczynnik korelacji liniowej Pearsona lub współczynnik rang Spearmana. W niniejszym opracowaniu do oceny stopnia skorelowania zmiennych wybrano współczynnik korelacji liniowej Pearsona, przyjmując, że cechy charakteryzują się relatywnie dużą zmiennością, jeżeli współczynnik korelacji przyjmuje wartości wyższe niż 0,7. W niniejszym opracowaniu przyjęto arbitralnie zadaną wartość współczynnika korelacji i eliminacji cech dokonano na podstawie przesłanek merytorycznych. Niemniej jednak, oprócz arbitralnie przyjętej wartości współczynnika korelacji, w literaturze przedmiotu opisywane są również inne sposoby wskazania wartości progowej współczynnika korelacji, powyżej której dokonuje się eliminacji cech będących nośnikiem podobnych informacji, i tak w podejściu formalnym stosuje się najczęściej metodę minimaksową lub procedurę weryfikacji istotności współczynnika korelacji z populacji generalnej na podstawie próby [por. Grabiński, Wydymus, Zeliaś 1989, Panek 2009].

W analizie skupień wysoki stopień skorelowania zmiennych może powodować dominację tych wymiarów w tworzeniu klastrów i tym samym nieistotność pozostałych zmiennych. Dlatego też odrzucono zmienne, których korelacje z pozostałymi przekraczały wartość 0,7 [Nowak 1990]. Wyjątek tutaj

stanowi zależność korelacyjną między zmiennymi X9 i X14, $r_{yx} = 0,86$. Ze względu na duże znaczenie merytoryczne i przydatność społeczno-ekonomiczną tych zmiennych dla omawianej problematyki oraz fakt, iż zmienne należą do odrębnych obszarów analizy (X9 – obszar patenty, X14 – obszar finansowanie) cechy zostały uwzględnione w ostatecznym zbiorze. Na podstawie współczynników korelacji między zmiennymi i sum ich wartości bezwzględnych, wykluczone zostały następujące zmienne: X5, X8, X10, X13, X15. (zob. tab. 1)

Na podstawie przeprowadzonej analizy w zakresie doboru cech diagnostycznych można stwierdzić, iż w ostatecznym zestawie znalazły się zmienne, charakteryzujące się wysoką zdolnością do dyskryminacji jednostek w analizowanym obszarze. Oznacza to ich dużą zmienność w przestrzeni oraz niski stopień skorelowania między sobą. Tym samym usunięte zostały zmienne będące nośnikiem podobnych informacji.

ZAŁOŻENIA I SCHEMAT POSTĘPOWANIA EMPIRYCZNEGO

W literaturze przedmiotu opisanych jest wiele miar „podobieństwa” i „niepodobieństwa” obiektów oraz szeroka gama metod tworzenia skupień, o czym będzie mowa w tej części opracowania.

Empiryczna analiza danych została przeprowadzona w następującym układzie:

1. Liczba obiektów poddanych badaniu – 28 krajów członkowskich UE;
2. Liczba zmiennych diagnostycznych użytych do opisu potencjału innowacyjnego gospodarek UE – 10 wskaźników;
3. Normalizacja zmiennych;
4. Wybór miar odległości;
5. Wybór metod aglomeracyjnych.

Przebieg etapu 1 oraz 2, które zamieszczono powyżej, zostały już scharakteryzowane w poprzednim rozdziale. Kolejno, dla wybranego zestawu zmiennych diagnostycznych została przeprowadzona procedura normalizacji w formie unitaryzacji zerowej oraz standaryzacji. Ostatecznie jednak zdecydowano o przeprowadzeniu analizy skupień dla wybranych obiektów (N = 28 państw członkowskich) bez poddawania zmiennych procedurze normalizacji. Poddanie zmiennych procedurze normalizacji wprowadziło istotne zakłócenia w formułowaniu ostatecznych wniosków dotyczących podziału zbioru na klastry. Wartości indeksu sylwetkowego dla skupień uzyskanych za pomocą zmiennych poddanych procedurze normalizacji są znacznie niższe aniżeli w przypadku wyodrębnienia skupień na zmiennych rzeczywistych. Ze względu na mierzalny charakter zmiennych diagnostycznych analiza skupień została przeprowadzona dla dwóch miar odległości, tj. dla odległości euklidesowej oraz dla kwadratu odległości euklidesowej [StatSoft – Electronic Statistics Textbook].

W literaturze przedmiotu prezentowany jest szereg algorytmów analizy skupień dokonywanej przy pomocy metod aglomeracyjnych, wśród których w zależności od techniki grupowania rozróżnia się procedury kombinatoryczne i niekombinatoryczne. Poszczególne algorytmy różnią się sposobem wyznaczania odległości pomiędzy skupieniami. W niniejszym opracowaniu Autorka przeprowadziła grupowanie państw UE ze względu na ich zdolność do generowania innowacji stosując dziewięć algorytmów analizy skupień hierarchicznych metod aglomeracyjnych [por. Wishart 2006, Frątczak 2009]:

- Pojedynczego połączenia (najbliższego sąsiedztwa);
- Całkowitego połączenia (najdalszego sąsiedztwa);
- Średniego połączenia (średniej odległości między skupieniami);
- Ważona średniego połączenia (ważonej średniej odległości między skupieniami);
- Środka ciężkości;
- Minimalnej wariancji Warda;
- Sumy kwadratów;
- Średniego sąsiedztwa.

W artykule zaprezentowane zostały trzy spośród licznie dostępnych procedur wyboru liczby skupień, dwa kryteria Mojeny oraz kryterium Wisharta. Są to jedne z nielicznych procedur dedykowanych metodom klasyfikacji hierarchicznej stąd znalazły one zastosowanie empiryczne w niniejszym opracowaniu [Mikulec 2013]. (zob. Tabela 2)

Tabela 2. Kryteria wyboru wraz z opisem

KRYTERIUM	Formuła, przedział zmienności, kryterium wyboru
Mojeny I – Górnego obszaru odrzucenia (eng. <i>upper tail rule</i>)	$d_{i+1} > \bar{d} + k * S(d)$ <p>KRYTERIUM WYBORU</p> <p>Klasyfikacja P_x, aby odpowiadający jej krok $i: i = 2, 3, \dots, n - 1$ pierwszy spełniał nierówność;</p> <p>d_i = długość i-tego wiązania (i-tej gałęzi drzewa)</p> <p>d_{i+1} = odległość połączenia grup w kroku $i+1$</p> <p>$\bar{d}, S(d)$ – średni poziom (średnia arytmetyczna) i odchylenie standardowe długości wiązań (odległości połączenia grup)</p> <p>k – stała, której wartość według R. Mojeny powinna zawierać się w przedziale [2,5; 3,5].</p>

KRYTERIUM	Formuła, przedział zmienności, kryterium wyboru
Mojeny II – Średniej ruchomej (eng. <i>moving average quality control rule</i>)	$\alpha_{i+1} > \bar{\alpha}_i + L_i + b_i + k \cdot s_i$ <p>KRYTERIUM WYBORU</p> <p>Klasyfikacja P_i, aby odpowiadający jej krok $i : i = y, y + 1, \dots, n - 2$ pierwszy spełniał nierówność</p> <p>y – liczba wartości poziomu (odległości) połączenia klas α w danym kroku wykorzystana do wyznaczenia średniej ruchomej; $\bar{\alpha}_i$ – średnia ruchoma wartości parametru α obliczona w kroku i; L_i – korekta dla opóźnionego „trendu” poziomu (odległości) połączenia klas obliczona w kroku i; b_i – „ruchome” średniokwadratowe nachylenie linii trendu poziomu połączenia klas w kroku i; s_i – „ruchome” odchylenie standardowe wartości parametru α (odległości).</p>
Wisharta – Losowości podziału obiektów na wykresie drzewa (eng. <i>tree validation</i>)	<p>Porównywanie wyników ciągu klasyfikacji uzyskanych metodami aglomeracyjnymi z rodziną drzew generowanych na podstawie losowej permutacji zbioru danych.</p> <p>KRYTERIUM WYBORU</p> <p>H_0 mówiąca o tym, że struktura grupowania obiektów w postaci danego drzewa jest losowa (brak struktury), $H_1 : \sim H_0$</p>

Źródło: na podstawie [Mikulec 2013]

ANALIZA SKUPIEŃ – WYNIKI

Stosując 9 algorytmów aglomeracyjnych przy trzech wybranych kryteriach podziału zbioru na skupienia uzyskano 54 rozwiązania opisane poniżej. (zob. Tabela 3) Rozwiązania A wygenerowano dla kwadratu odległości euklidesowej a Rozwiązania B dla odległości euklidesowej. Każde rozwiązanie zostało poparte stosowną ilustracją graficzną w postaci dendrogramu. Mając na uwadze przesłanki natury merytorycznej i statystycznej Autorka zdecydowała o ostatecznym podziale zbioru na 3 skupienia. Po pierwsze, wyłonienie większej aniżeli 3 liczby skupień, nie dawało lepszych rezultatów, tzn. kolejne skupienia były wyłonione ze skupień najmniej licznych, a nie ze skupienia 15-elementowego, czyli najbardziej licznego w przypadku podziału na 3 grupy (sprawdzony szczegółowo został podział na 4 i 5 skupień). Po drugie, tworzenie większej liczby skupień nie poprawiało jakości grupowania mierzonej wartością indeksu sylwetkowego³.

³ Ocenie jakości grupowania wyrażonej za pomocą indeksu sylwetkowego oraz skorygowanego indeksu Randa będzie poświęcone odrębne opracowania.

Tabela 3. Rozwiązanie uzyskane dla wybranych algorytmów i kryteriów wyboru liczby skupień

Lp.	R algorytmu	Kryterium	Skupienia	Lp.	R algorytmu	Kryterium	Skupienia
Rozw_A_1	Ward	UT	3	Rozw_B_1	Ward	UT	3
Rozw_A_2	Ward	MA	6	Rozw_B_2	Ward	MA	5
Rozw_A_3	Ward	TREE	6	Rozw_B_3	Ward	TREE	3
Rozw_A_4	Pojedynczego połączenia	UT	4	Rozw_B_4	Pojedynczego połączenia	UT	7
Rozw_A_5	Pojedynczego połączenia	MA	4	Rozw_B_5	Pojedynczego połączenia	MA	2
Rozw_A_6	Pojedynczego połączenia	TREE	1	Rozw_B_6	Pojedynczego połączenia	TREE	2
Rozw_A_7	Pełnego połączenia	UT	3	Rozw_B_7	Pełnego połączenia	UT	5
Rozw_A_8	Pełnego połączenia	MA	5	Rozw_B_8	Pełnego połączenia	MA	3
Rozw_A_9	Pełnego połączenia	TREE	3	Rozw_B_9	Pełnego połączenia	TREE	4
Rozw_A_10	Średniego połączenia	UT	3	Rozw_B_10	Średniego połączenia	UT	5
Rozw_A_11	Średniego połączenia	MA	5	Rozw_B_11	Średniego połączenia	MA	5
Rozw_A_12	Średniego połączenia	TREE	3	Rozw_B_12	Średniego połączenia	TREE	3
Rozw_A_13	Średniego ważonego połączenia	UT	3	Rozw_B_13	Średniego ważonego połączenia	UT	5
Rozw_A_14	Średniego ważonego połączenia	MA	5	Rozw_B_14	Średniego ważonego połączenia	MA	2
Rozw_A_15	Średniego ważonego połączenia	TREE	2	Rozw_B_15	Średniego ważonego połączenia	TREE	2
Rozw_A_16	Środków ciężkości	UT	3	Rozw_B_16	Środków ciężkości	UT	4
Rozw_A_17	Środków ciężkości	MA	5	Rozw_B_17	Środków ciężkości	MA	4
Rozw_A_18	Środków ciężkości	TREE	5	Rozw_B_18	Środków ciężkości	TREE	5
Rozw_A_19	Mediany	UT	4	Rozw_B_19	Mediany	UT	4
Rozw_A_20	Mediany	MA	5	Rozw_B_20	Mediany	MA	4
Rozw_A_21	Mediany	TREE	4	Rozw_B_21	Mediany	TREE	4
Rozw_A_22	Sumy kwadratów	UT	3	Rozw_B_22	Sumy kwadratów	UT	3
Rozw_A_23	Sumy kwadratów	MA	5	Rozw_B_23	Sumy kwadratów	MA	3
Rozw_A_24	Sumy kwadratów	TREE	5	Rozw_B_24	Sumy kwadratów	TREE	5
Rozw_A_25	Średniego sąsiedztwa	UT	4	Rozw_B_25	Średniego sąsiedztwa	UT	6
Rozw_A_26	Średniego sąsiedztwa	MA	5	Rozw_B_26	Średniego sąsiedztwa	MA	2
Rozw_A_27	Średniego sąsiedztwa	TREE	3	Rozw_B_27	Średniego sąsiedztwa	TREE	3

Źródło: opracowanie własne na podstawie danych statystycznych z bazy Eurostat, obliczenia wykonano w Clastan Graphics

Bazując na wartości indeksu sylwetkowego (0,791) Autorka zdecydowała o ostatecznym podziale zbioru na 3 grupy. Identyfikacyjnych rozwiązań dostarczyły

dwie metody tj., metoda Warda przy I kryterium Mojeny oraz metoda pełnego połączenia stosując I kryterium Mojeny oraz kryterium Wisharta.

Tabela 4. Przyporządkowanie krajów do poszczególnych grup

Numer klastra	Państwa
Klaster 1	Rumunia Bułgaria Grecja Chorwacja Słowacja Portugalia Republika Czeska Węgry Malta Litwa Łotwa Polska Cypr Hiszpania Estonia
Klaster 2	Włochy Słowenia Irlandia Wielka Brytania Luksemburg Belgia Francja
Klaster 3	Austria Niderlandy Dania Finlandia Szwecja Niemcy

Źródło: opracowanie własne na podstawie obliczeń wykonanych w ClustanGraphics

W klastrze 3, dla każdej ze zmiennych diagnostycznych, wartości średnich wewnątrzgrupowych są znacznie wyższe aniżeli średnie ogólne, jest to jednocześnie grupa najlepsza z punktu widzenia potencjału do generowania innowacji. (zob. tab. 5) W grupie tej znajdują się wszystkie kraje skandynawskie oraz Austria, Niemcy i Niderlandy. Skupienie trzecie jest również najbardziej jednorodne, współczynniki zmienności wewnątrzgrupowe stojące przy każdej ze zmiennych diagnostycznych są zdecydowanie niższe aniżeli wartości współczynników zmienności wyznaczone dla całej badanej zbiorowości. Największa różnica w poziomie zmienności występuje dla cechy patenty zgłoszone do Europejskiego Urzędu Patentowego. Tutaj współczynnik zmienności ogólny jest ponad siedmiokrotnie wyższy aniżeli wewnątrzgrupowy. W skupieniu trzecim znalazły się jednostki, dla których odnotowano najwyższą wartość, ze wszystkich możliwych wartości zarejestrowanych dla całej populacji, w odniesieniu do następujących cech: wskaźnik zatrudnienia, wydatki publiczne na edukację, zgłoszenia patentów do EUP, kompetencje w zakresie obsługi komputera oraz wydatki przedsiębiorstw na B+R. Analizując wartości minimalne dla poszczególnych zmiennych, należy stwierdzić, iż w niniejszym klastrze nie wystąpiły państwa, dla których zaobserwowano najniższą z możliwych wartości.

W przypadku grupy 1, do której należy również Polska, wraz ze wszystkimi państwami które dołączyły do UE nie wcześniej jak w 2004, średnie wewnątrzgrupowe dla każdej z 10 analizowanych cech są niższe aniżeli średnie ogólne, jest to jednocześnie najłabsza grupa ze względu na potencjał do generowania innowacji. (zob. tab. 5) Grupa ta, na tle dwóch pozostałych, jest najmniej jednorodna. Potwierdzają to wewnątrzgrupowe wartości współczynników zmienności, które aż dla 7 spośród 10 zmiennych diagnostycznych są wyższe aniżeli wartości ogólne tej miary. Ponadto, poddając analizie wartości maksymalne i minimalne stojące przy wybranych zmiennych diagnostycznych, widzimy, że w skład skupienia pierwszego weszły kraje, w których dla wybranych cech odnotowano wartość najwyższą bądź najniższą spośród występujących w całej populacji. Taka sytuacja ma miejsce w odniesieniu do następujących cech diagnostycznych: wydatki rządowe na działalność badawczo-rozwojową, studenci w grupie wieku 15-24 lata, przedsiębiorstwa zatrudniające specjalistów ICT.

Polska w zakresie potencjału innowacyjnego ma poziom zbliżony do następujących krajów: Rumunia, Bułgaria, Grecja, Chorwacja, Słowacja, Portugalia, Republika Czeska, Węgry, Malta, Litwa, Łotwa, Cypr, Hiszpania, Estonia.

Obecnie (dane z roku 2012) wydaje się, że jedynie Słowenia ma szansę dorównać do poziomu państw byłej „piętnastki”. Grupa 2 to grupa złożona w przeważającej mierze z krajów dawnej „15”, które w znaczący sposób odczuły skutki kryzysu gospodarczego. (zob. tab. 5) Słowenia jest w tej grupie jedynym reprezentantem „nowych” krajów członkowskich. Porównując tę grupę z grupą pierwszą należy stwierdzić, iż jest ona bardziej jednorodna. Dla większości spośród badanych cech współczynniki zmienności wewnątrzgrupowe są niższe aniżeli współczynniki ogólne obliczone dla całej populacji. W niniejszym klastrze znalazły się państwa, dla których zaobserwowano najniższą wartość, ze wszystkich możliwych w populacji, jedynie w odniesieniu do dwóch zmiennych: ludność z wykształceniem wyższym oraz wydatki publiczne na edukację.

Tabela 5. Wybrane miary statystyczne dla poszczególnych klastrów

Klaster					Klaster				
	Ogółem	1	2	3		Ogółem	1	2	3
Wielkość	28	15	7	6	Wielkość	28	15	7	6
Wydatki rządowe na działalność badawczo-rozwojową					Wydatki publiczne na edukację (% PKB)				
\bar{X}	1,25	1,08	1,22	1,74	\bar{X}	5,49	5,06	5,49	6,60
S_x	0,47	0,52	0,14	0,19	S_x	1,30	1,14	1,24	1,28
Min	0,4	0,4	1,09	1,54	Min	3,15	3,53	3,15	5,08
Max	2,12	2,12	1,49	2,02	Max	8,8	7,92	6,57	8,8
V_s	37,5%	48,2%	11,7%	10,8%	V_s	23,7%	22,5%	22,7%	19,5%
Wskaźnik zatrudnienia (grupa wieku 20-64 lata)					Zgłoszenia patentów do EUP				
\bar{X}	67,94	64,58	67,89	76,38	\bar{X}	79,04	13,57	91,53	228,14
S_x	6,42	5,19	4,47	1,85	S_x	88,48	12,70	26,77	34,74
Min	55,3	55,3	61	74	Min	1,53	1,53	63,75	193,95
Max	79,4	72,1	74,2	79,4	Max	272,25	44,25	132,55	272,25
V_s	9,4%	8,0%	6,6%	2,4%	V_s	111,9%	93,6%	29,2%	15,2%
Zatrudnienie w sektorze technologii i wiedzochłonnym					Kompetencje w zakresie obsługi komputera				
\bar{X}	3,95	3,32	4,70	4,63	\bar{X}	27,79	24,73	29,29	33,67
S_x	1,30	1,16	1,32	0,87	S_x	7,79	7,44	5,77	7,74
Min	2,2	2,2	3,3	3,5	Min	8	8	23	21
Max	7,5	6,1	7,5	5,7	Max	42	35	40	42
V_s	33,0%	34,8%	28,1%	18,8%	V_s	28,0%	30,1%	19,7%	23,0%
Ludność z wykształceniem wyższym (grupa wieku 30-34 lata)					Przedsiębiorstwa zatrudniające specjalistów ICT				
\bar{X}	36,41	32,41	42,31	39,53	\bar{X}	23,14	21,93	24,00	25,17
S_x	10,02	9,32	9,93	8,49	S_x	7,17	7,71	8,17	4,54
Min	21,7	21,8	21,7	26,3	Min	5	5	14	20
Max	51,1	49,9	51,1	47,9	Max	36	36	33	31
V_s	27,5%	28,8%	23,5%	21,5%	V_s	31,0%	35,2%	34,0%	18,0%

Tabela 5 cd.

Klaster					Klaster				
	Ogółem	1	2	3		Ogółem	1	2	3
Studenci w grupie wieku 15-24 lata					Wydatki przedsiębiorstw na B&R (% PKB)				
\bar{x}	61,66	60,38	60,36	66,37	\bar{x}	1,01	0,48	1,30	1,97
S_x	8,63	9,10	8,95	6,34	S_x	0,73	0,35	0,47	0,42
Min	41,3	41,3	48,5	55,5	Min	0,06	0,06	0,69	1,22
Max	74,8	74,8	72,4	72,7	Max	2,44	1,25	2,16	2,44
V_s	14,0%	15,1%	14,8%	9,6%	V_s	72,3%	72,6%	36,1%	21,5%

Źródło: opracowanie własne na podstawie obliczeń wykonanych w ClustanGraphics

PODSUMOWANIE I WNIOSKI

Wybrane wskaźniki do oceny potencjału innowacyjnego europejskich gospodarek wykazały, iż poziom zróżnicowania poszczególnych krajów w wybranych obszarach – finansowanie, B+R i nauka, edukacja, rynek pracy, patenty – jest bardzo duży. Wyróżnienie państw, które znacznie odbiegają od czołówki państw członkowskich pozwoli na efektywniejsze wdrażanie polityki unijnej poprzez tworzenia rozwiązań adekwatnych do poziomu zaawansowania innowacyjnego w każdym z analizowanych krajów. Wyróżnienie grup państw o podobnym potencjale jest również korzystne z punktu widzenia prowadzenia analiz porównawczych i tworzenia oraz implementowania wspólnych rozwiązań dla wybranych grup krajów.

Zaproponowany w analizie zestaw cech diagnostycznych zawierał 15 zmiennych, z którego po wstępnej weryfikacji merytorycznej i statystycznej zostało wyeliminowanych 5 zmiennych, ostatecznie zestaw 10 zmiennych diagnostycznych stanowił podstawę klasyfikacji państw. W zależności od algorytmu grupowania wyróżniony podział państw zawierał od 1 do 7 skupień.

W wyniku przeprowadzonej analizy skupień sprawdzono 54 rozwiązania stosując 9 algorytmów grupowania oraz 3 kryteria wyboru optymalnej liczby klas. Ostatecznie obliczony dla wszystkich 54 wariantów indeks sylwetkowy wskazał na podział zbioru na 3 klastry.

Najkorzystniejszą do poszukiwania związków pomiędzy obiektami, odnośnie zdefiniowanych w pracy zdolności do generowania innowacji, wydają się być metoda pełnego połączenia i metoda Warda dające podział zestawu obiektów na trzy skupienia. Skład grup uzyskany tymi metodami jest jednakowy, co wskazuje na możliwość wykorzystania tych metod w badaniach podobnego typu.

Polska na tle pozostałych krajów członkowskich, które dołączyły do struktur unijnych nie wcześniej niż w 2004 roku wypada nie najgorzej, niemniej jednak jej przynależność do najsłabszej z wyróżnionych grup potwierdza, iż nakłady jakie muszą jeszcze zostać poniesione na rozwój nauki, edukacji w tym sferę badawczo-rozwojową są znaczące. Dodatkowo analiza skupień pokazała, że Rumunia

najbardziej odstaje od pozostałych państw członkowskich UE i jednocześnie różni się najbardziej od Finlandii i Szwecji. Jednocześnie Finlandia, Dania i Szwecja wyróżniają się istotnie na tle państw UE-28. Przy czym to Dania i Finlandia są do siebie najbardziej podobne ze względu na wybrany zbiór danych.

Warto podkreślić, iż w analizie skupień badacz przed zaproponowaniem ostatecznego rozwiązania powinien przetestować wiele różnych możliwych grupowań, ponieważ zastosowanie różnych algorytmów grupowania oraz odmiennych kryteriów wyboru optymalnej liczby klastrów prowadzi do różnych rezultatów. Niemniej jednak w przypadku metody pełnego połączenia zarówno dla I kryterium Mojeny jak i kryterium Wisharta otrzymujemy te same wyniki.

BIBLIOGRAFIA

- Fagerberg J. (1998) International Competitiveness, *Economic Journal*, vol. 98, No. 2, s. 102-104.
- Frątczak E. (red.) (2009) Wielowymiarowa analiza statystyczna. Teoria – przykłady zastosowań z systemem SAS, Szkoła Główna Handlowa w Warszawie, Warszawa, s. 120-122.
- Gatnar E., Walesiak M. (2004) Metody statystycznej analizy wielowymiarowej w badaniach marketingowych, Wydawnictwo Akademii Ekonomicznej, Wrocław, s. 310-317.
- Gordon A. D. (1999) Classification, (2nd edition), Chapman & Hall/CRC, Boca Raton, s. 6-10.
- Grabiński T., Wydymus S., Zeliaś A. (1989) Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych, PWN, Warszawa, s. 50.
- Kisielewska J., Stańko S. (2009) Istota analizy wielowymiarowej, *Roczniki Nauk Rolniczych, Seria G, T. 96, z. 2*, Warszawa, s. 64.
- Kunasz M. (2006) Przykład zastosowania metod WAP do analizy procesów gospodarowania zasobami ludzkimi w przedsiębiorstwie, [w:] *Kapitał ludzki w gospodarce opartej na wiedzy*, D. Kopycińska (red.), Katedra Mikroekonomii Uniwersytetu Szczecińskiego, Szczecin, s. 131-139
http://mikroekonomia.net/system/publication_files/904/original/11.pdf?1315223800.
- Malina A., Zeliaś A. (1996) Taksonomiczna analiza przestrzennego zróżnicowania jakości życia ludności w Polsce w 1994 r., [w:] *Ekonometryczne modelowanie danych finansowo-księgowych*, Nowak E., Urbaniak M. (red.), UMCS, Lublin, s. 85-89.
- Markowska M. (2012) Dynamiczna taksonomia innowacyjności regionów, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław, s. 73.
- Mikulec A. (2013) Kryterium Mojeny i Wisharta w analizie skupień – przypadek skupień o różnych macierzach kowariancji [w:] *Taksonomia 20. Klasyfikacja i analiza danych – teoria i zastosowania*, Jajuga K., Walesiak M. (red.), *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, Nr 278, Wrocław, ISSN 1899-3192, s. 206-208.
- Mikulec A. (2012) Analiza skupień z wykorzystaniem programu komputerowego Clustangraphics [w:] *Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne*, Zieliński Z. E. (red.), Kielce, ISBN 978-83-89274-75-5, Tom. II, s. 214-216.

- Milligan G.W., Cooper M.C. (1987) Methodology review: clustering methods, *Applied Psychological Measurement*, Vol. 11, No 4, 1987, s. 329-331.
- Nowak E. (1990) Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych, PWN, Warszawa, [za:] Wojnar J., Zastosowanie metod taksonomicznych do klasyfikacji państw Unii Europejskiej, *Zeszyty Naukowe, Świętokrzyskie Centrum Edukacji na Odległość*, s. 396,
http://www.sceno.edu.pl/cms_tmp/2438_Jolanta%20Wojnar%20-%20III%20SCENO.pdf.
- Ostasiewicz W. (red.) (1998) Statystyczne metody analizy danych, Akademia Ekonomiczna we Wrocławiu, Wrocław, s. 118.
- Panek T. (2009) Statystyczne metody wielowymiarowej analizy porównawczej, Szkoła Główna Handlowa w Warszawie, Warszawa, s. 20-23.
- Podręcznik OSLO (2005) Zasady gromadzenia i interpretacji danych dotyczących innowacji, OECD, Eurostat, Paryż, s. 20-23,
<http://www.uwm.edu.pl/ciitt/wp-content/uploads/2013/10/Podrecznik-OSLO-MANUAL1.pdf>
http://europa.eu/rapid/press-release_IP-12-102_pl.htm
http://ec.europa.eu/news/science/120208_pl.htm

COMPARATIVE STUDIES ON THE INNOVATIVE POTENTIAL OF THE EUROPEAN UNION MEMBER COUNTRIES BASED ON THE CLUSTER ANALYSIS

Abstract: Under multidimensional data analysis researchers face the problem of grouping methods selection. The aim of the paper is to find the inner homogenic groups created by 28 EU member countries in relation to the selected characteristics describing the potential of economies to create innovations. The issue of clustering objects – countries was conducted with the use of ClustanGraphics software. Ten variables were selected from the Eurostat database in order to describe the potential of the EU member countries to create innovations. These were tested using nine hierarchical clustering methods. The *Increase in Sum of Squares* method and the *Complete Linkage* method seem to be the most appropriate for further discussion regarding cross-group analysis. The rest of tested methods do not seem suitable for deep analysis between the groups. This is mainly due to the tendency to form rather unclear clusters with the structure of long chains.

Keywords: cluster analysis, agglomerative methods, Mojena criteria, Wishart criteria, innovative potential, EU member countries

EFEKTYWNOŚĆ I PRODUKTYWNOŚĆ ROLNICTWA W POLSCE – ANALIZA Z WYKORZYSTANIEM INDEKSÓW TFP HICKSA-MOORSTEENA

Robert Rusielik

Katedra Zarządzania Przedsiębiorstwami
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: robert.rusielik@zut.edu.pl

Streszczenie: Głównym celem badań była próba wykorzystania indeksów produktywności TFP (Total Factor Productivity) Hicksa-Moorstena do badania produktywności i efektywności rolnictwa w skali regionalnej oraz przyczyn ich zmian. Badania obejmują lata 2010 – 2012. Obliczenia wykonano w podziale na województwa. Dla każdego województwa obliczono indeksy produktywności TFP Hicksa-Moorstena. Dekompozycja obliczonych indeksów na kilka alternatywnych miar efektywności pozwoliła również na określenie przyczyn zmian pomiędzy poszczególnymi okresami badań.

Słowa kluczowe: produktywność, efektywność, rolnictwo, TFP, DEA, indeksy Hicksa-Moorsteena.

WSTĘP

Do badania produktywności całkowitej (TFP) w sytuacji wielowymiarowych nakładów i efektów najpowszechniej wykorzystywane są indeksy produktywności Malmquista. Jest wiele badań, które dotyczyły grup przedsiębiorstw i całych sektorów w tym rolnictwa w Polsce wykorzystujących tę technologię np. [Breummer i in. 2002], [Latruffe i in. 2004], [Zawaliska 2004], [Balcombe i in. 2005], [Rusielik 2009] jak i badań porównawczych pomiędzy różnymi krajami. Färe [Färe i in. 1992] przedstawił technikę dekompozycji indeksów TFP na komponenty zmian efektywności i technologii a następnie koncepcję tę rozwinął w publikacji [Färe i in. 1994], gdzie wykazał jak zmiany efektywności można dekomponować na zmiany czystej efektywności technicznej i zmiany efektywności skali. Przedstawiona w tej publikacji postać indeksu Malmquista stała się jednym z najbardziej popularnych narzędzi pomiaru w badaniu zmian produktywności. Jednak w związku z tym, że

metoda zakłada przyjęcie stałych efektów skali (CRS) wywołuje dyskusję na temat możliwych błędów i niewiarygodnych wyników. O'Donnell [2010, 2012a, 2012b] i [Hoang 2011] a w Polsce [Rusielik 2014] w badaniach produktywności rolnictwa wskazywali na większą przydatność indeksów TFP Hicksa-Moorsteena do pomiaru produktywności rolnictwa, co między innymi wynikało z dopuszczenia założeń zmiennych efektów skali (VRS) i lepszego dopasowania modelu do słabszych technologicznie obiektów. W prezentowanych badaniach obliczono indeksy produktywności całkowitej (TFP) działalności rolniczej w Polsce za lata 2010-2012 a następnie przeprowadzono dekompozycję otrzymanych indeksów na kilka wybranych miar efektywności. Wyniki przedstawiono w układzie województw a także regionów rolniczych FADN. Ze względu na ograniczenia redakcyjne w artykule przedstawiono tylko wybrane wyniki badań.

MATERIAŁ I METODY

Celem badań było zbadanie możliwości wykorzystania indeksów produktywności Hicksa-Moorsteena do pomiaru efektywności rolnictwa w Polsce. Ze względu na wymogi redakcyjne w opracowaniu wykorzystano tylko wybrane syntetyczne wyniki badań. Dane wykorzystane w badaniach pochodzą z publikacji statystycznych tj.: Roczniki statystyczne województw i Roczniki statystyczne rolnictwa i obszarów wiejskich.

Mając technologię opartą na pojedynczym nakładzie i pojedynczym efekcie produktywności całkowitej (TFP total factor productivity) obiektu zazwyczaj definiujemy jako stosunek efektu do nakładu. W sytuacji wielowymiarowej TFP można z kolei zdefiniować jako stosunek zagregowanych efektów do zagregowanych nakładów. Kiedy nie są znane relacje cenowe obliczenia są problematyczne. O'Donnell [O'Donnell 2008] zdefiniował, w jaki sposób można ten problem rozwiązać bez znajomości relacji cenowych wykorzystując indeksy produktywności oparte na relacjach pomiędzy badanymi obiektami.

Niech $Q_t \equiv Q(q_t)$ i $X_t \equiv X(x_t)$ oznaczają zagregowane efekty i nakłady skojarzone z wektorami q_t i x_t to TFP dla obiektu t prezentuje równanie $TFP_t = Q_t / X_t$. Zmiany TFP pomiędzy dającym odniesienie obiektem 0 a obiektem t można z kolei przedstawić za pomocą indeksu (1) [O'Donnell 2008]:

$$TFP_{0t} = TFP_t / TFP_0 = Q_{0t} / X_{0t} \quad (1)$$

gdzie $Q_{0t} = Q_t / Q_0$ i $X_{0t} = X_t / X_0$ są indeksami mierzącymi zmiany zagregowanych efektów i nakładów.

Założenia te można wykorzystać do obliczeń indeksów zmian TFP pomiędzy dwoma punktami w czasie. W zależności od przyjętej postaci funkcji indeksy te mogą przyjmować różną postać. Do najczęściej wykorzystywanych indeksów należą

indeksy Laspeyresa, Paaschego i Fishera jednak wymagają one znajomości wektorów cen efektów i nakładów jako wag poszczególnych czynników. Można ten problem rozwiązać przez wykorzystanie różnych zagregowanych funkcji odległości, kalkulowanych na bazie dostępnych nakładów i efektów. Przykładami wykorzystania takich zagregowanych funkcji do kompozycji indeksów TFP są indeksy Malmquista, Hicksa-Moorsteena i Färe-Primonta, które można obliczyć wykorzystując programowanie liniowe (LP) i założenia metody Data Envelopment Analysis (DEA) [O'Donnel 2011].

Jeżeli przyjmiemy, że $x_{it} = (x_{1it}, \dots, x_{Kit})'$ i $q_{it} = (q_{1it}, \dots, q_{Jit})'$ to wektory nakładów i efektów to TFP obiektu i w okresie t to:

$$TFP_{it} \equiv \frac{Q_{it}}{X_{it}} \quad (2)$$

gdzie $Q_{it} = Q(q_{it})$ to zagregowany efekt, $X_{it} = X(x_{it})$ to zagregowany nakład natomiast $Q(\cdot)$ i $X(\cdot)$ są niemalejącymi, nieujemnymi, liniowo jednorodnymi funkcjami. Z kolei indeks produktywności, który mierzy TFP obiektu i w okresie t w relacji do TFP obiektu h w okresie s można przedstawić równaniem:

$$TFP_{hs,it} \equiv \frac{TFP_{it}}{TFP_{hs}} = \frac{Q_{it}/X_{it}}{Q_{hs}/X_{hs}} = \frac{Q_{hs,it}}{X_{hs,it}} \quad (3)$$

gdzie $Q_{hs,it} = Q_{it}/Q_{hs}$ to indeks wielkości efektów a $X_{hs,it} = X_{it}/X_{hs}$ to indeks wielkości nakładów. W tym kontekście, wymiarem zmian TFP będzie iloraz zmian efektów do zmian nakładów. Indeksy w postaci (3) O'Donell [2008, 2010, 2011] określił jako w pełni multiplikatywne (*multiplicatively-complete*). Wszystkie tego typu indeksy można dekomponować na miarę zmian technologii i kilka miar zmian efektywności.

Przyjmując, że q_0, x_0 to wektory efektów i nakładów, t_0 oznacza okres odniesienia w czasie, natomiast $D_0(\cdot), D_I(\cdot)$ to odpowiednio funkcje odległości efektów i nakładów, oraz że $Q(q) = [D_O(x_{hs}, q, s) D_O(x_{it}, q, t)]^{1/2}$ i $X(x) = [D_I(x, q_{hs}, s) D_I(x, q_{it}, t)]^{1/2}$ to indeks Hicksa-Moorsteena przedstawia równanie (4) [Diewert 1992]¹:

$$TFP_{hs,it}^{HM} = \left(\frac{D_O(x_{hs}, q_{it}, s) D_I(x_{hs}, q_{hs}, s) D_O(x_{it}, q_{it}, t) D_I(x_{hs}, q_{it}, t)}{D_O(x_{hs}, q_{hs}, s) D_I(x_{it}, q_{hs}, s) D_O(x_{it}, q_{hs}, t) D_I(x_{it}, q_{it}, t)} \right)^{1/2}. \quad (4)$$

¹ Diewert, W. E. (1992) Fisher Ideal Output, Input, and Productivity Indexes Revisited. Journal of Productivity Analysis, 3, s. 211-248.

Tak zdefiniowany indeks był zaproponowany przez Diewerta [1992] jako iloraz indeksów Malmquista zorientowanych na efekty i nakłady. Funkcje odległości efektów i nakładów zostały estymowane przy wykorzystaniu metody DEA przez rozwiązanie odpowiednich zadań programowania liniowego. Indeksy, które mogą być rozwijane z postaci (4) można dekomponować na miarę zmian technologii i kilka miar zmian efektywności. Przykładowo pomiar całkowitej efektywności produkcji danej firmy może być wyrażony jako stosunek obserwowanej TFP do maksymalnej TFP możliwej do osiągnięcia przy użyciu dostępnych technologii. Efektywność taką możemy oznaczyć jako efektywność TFPE. Matematycznie dla obiektu i w okresie t można to zapisać w postaci:²

$$TFPE_{it} = \frac{TFP_{it}}{TFP_t^*} = \frac{Q_{it} / X_{it}}{Q_{it}^* / X_{it}^*} \leq 1, \quad (5)$$

gdzie TFP_t^* oznacza maksymalną TFP możliwą do uzyskania w okresie t . Z kolei Q_t^* i X_t^* reprezentują kombinację zagregowanych efektów i nakładów maksymalizujących TFP.

Przyjmując te założenia można zdefiniować kilka miar efektywności stanowiących komponenty TFPE. Dla dekompozycji zmian TFP zorientowanej na nakłady przedstawiają równania (6), (7) i (8).

Efektywność techniczna (ITE):

$$ITE_{it} = \frac{Q_{it} / X_{it}}{Q_{it} / \bar{X}_{it}} = \frac{\bar{X}_{it}}{X_{it}} = D_t(x_{it}, q_{it}, t)^{-1} \leq 1. \quad (6)$$

Efektywność skali (ISE):

$$ISE_{it} = \frac{Q_{it} / \bar{X}_{it}}{\tilde{Q}_{it} / \tilde{X}_{it}} \leq 1. \quad (7)$$

Efektywność typu mix (IME):

$$IME_{it} = \frac{Q_{it} / \bar{X}_{it}}{Q_{it} / \hat{X}_{it}} = \frac{\hat{X}_{it}}{\bar{X}_{it}} \leq 1, \quad (8)$$

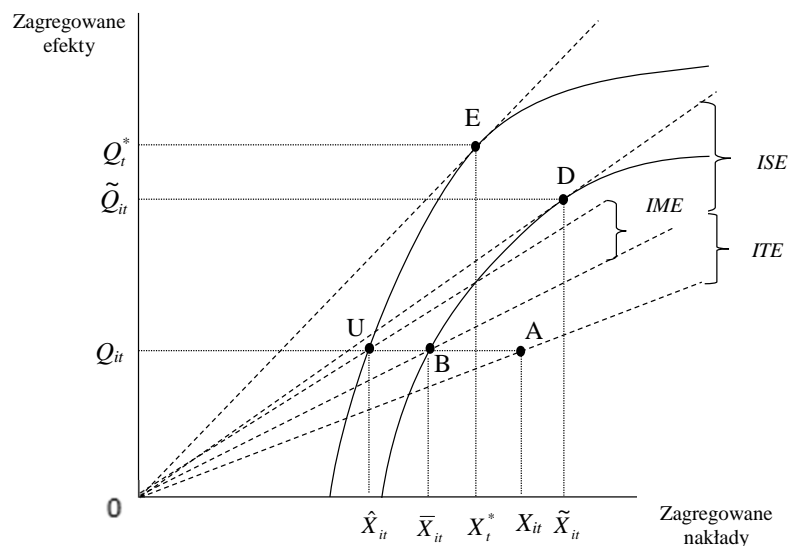
gdzie \bar{X}_{it} to minimalny zagregowany nakład, jaki jest możliwy do osiągnięcia, dla wytworzenia q_{it} przy wykorzystaniu wielokrotnionego skalarnie x_{it} ; natomiast \hat{X}_{it} to minimalny możliwy zagregowany nakład przy użyciu dowolnego

² C.J. O'Donnell. (2010) Measuring and decomposing agricultural productivity and profitability change. The Australian Journal of Agricultural and Resource Economics, 54, s. 533.

wektora nakładów do wytworzenia q_{it} , \tilde{Q}_{it} i \tilde{X}_{it} to zagregowane efekty i nakłady uzyskane przy maksymalnym TFP z ograniczeniem, że wektory efektów i nakładów (q_{it}, x_{it}) są odpowiednio zwielokrotnione skalarne. Przedstawiane miary efektywności są wybranymi do celów publikacji komponentami efektywności. Pełna dekompozycja obejmuje kilka dodatkowych indeksów które m.in. można znaleźć w publikacji [O'Donnell 2008].

Zależności pomiędzy poszczególnymi prezentowanymi miarami produktywności i efektywności można przedstawić obrazując wielowymiarową produkcję na zagregowanym obszarze efektów i nakładów w sposób pokazany na rysunku 1.

Rysunek 1. Miary efektywności dla modelu zorientowanego na nakłady



Źródło: opracowano na podstawie [O'Donnell 2011 s. 6]

Na rysunku 1 krzywa przechodząca przez punkty B i D reprezentuje granicę możliwości produkcyjnych dla obiektu operującego w punkcie A. Stanowi ona granicę technicznie możliwych kombinacji zagregowanych efektów i nakładów dla tego obiektu. Jeżeli ograniczenia zostaną złagodzone tj. możliwości produkcyjne rozszerzymy do maksymalnie możliwych w danym okresie to możliwe kombinacje zagregowanych efektów i nakładów będą reprezentowane przez krzywą przechodzącą przez punkty U i E. Pozwala to na zobrazowanie przedstawianych miar efektywności. Przykładowo ITE_{it} mierzy proporcjonalny wzrost TFP, czego odzwierciedleniem jest przesunięcie obiektu z punktu A do punktu B, który jest dla niego granicą możliwości produkcyjnych. Ograniczenie restrykcyj i otwarcie tej granicy daje możliwość przesunięcia punktu A do punktu U, czego miarą zmian

produktywności TFP jest wskaźnik IME_{it} . Efektywność skali ISE_{it} reprezentowana jest przez proporcję nachylenia prostych 0B/0D.

Wzorując się na publikacji [Coelli, Rao 2003], do modelu wykorzystano następujący zestaw zmiennych odzwierciedlających technologię produkcji w rolnictwie: efekt - $Y1$ – wartość skupu produktów rolnych (mln zł); nakłady – $X1$ – powierzchnia użytków rolnych (tys. ha), $X2$ – wartość brutto środków trwałych w rolnictwie (mln zł), $X3$ – pracujący w rolnictwie (tys.), $X4$ – pogłowie bydła (tys. szt.), $X5$ - pogłowie trzody (tys. szt.), $X6$ – nawożenie NPK (tys. t), $X7$ - nawożenie CaO (tys. t). W stosunku do oryginału zmodyfikowano dane wejściowe w ten sposób że wartość produkcji roślinnej i zwierzęcej reprezentowana jest przez wartość skupu produktów rolnych, natomiast ilość ciągników zastąpiono wartością środków trwałych w rolnictwie.

Tabela 1. Statystyki opisowe zmiennych za lata 2010-2012

Wyszczególnienie		Średnia	Min	Max	Odst.
Wartość skupu produktów rolnych (mln zł)	Y1	2381,2	739,5	8640,9	2163,4
Pow. UR (tys. ha)	X1	863,5	424,6	2033,0	448,8
Wartość brutto środków trwałych w rolnictwie i łowiectwie (mln zł)	X2	7139,9	2873,6	17603,6	4064,8
Pracujący w rolnictwie (tys.)	X3	115,9	32,6	305,8	92,0
Pogłowie bydła (tys. szt.)	X4	249,6	67,4	1042,0	309,5
Pogłowie trzody (tys. szt.)	X5	557,3	164,8	4277,4	974,0
NPK (tys. t)	X6	99,0	39,7	288,5	65,9
CaO (tys. t)	X7	25,4	3,6	85,9	22,2

Źródło: badania własne

Taki dobór zmiennych może budzić pewne obawy, co do zbyt dużego obciążenia wyników niewykorzystanym majątkiem trwałym, ale zostanie to zweryfikowane w trakcie kolejnych badań. Podstawowe statystyki opisowe zmiennych dla badanego okresu zamieszczono w Tabeli 1. Natomiast udział poszczególnych województw w wielkości wykorzystanych zmiennych zamieszczono w Tabeli 2.

Tabela 2. Średni udział województw w wielkości poszczególnych zmiennych w latach 2010-2012

	Y1	X1	X2	X3	X4	X5	X6	X7
Dolnośląskie	5,4	6,3	5,6	3,6	1,8	2,0	7,7	10,9
Kujawsko-pomorskie	8,4	6,9	6,9	4,5	8,0	11,9	9,2	9,9
Lubelskie	7,6	9,3	8,6	13,1	6,6	5,9	8,7	8,0
Lubuskie	2,3	3,1	2,3	1,4	1,2	1,2	2,7	3,1
Łódzkie	6,8	6,5	7,5	7,6	8,0	8,5	6,8	4,7
Małopolskie	2,1	4,2	5,1	11,6	3,3	2,1	2,3	1,2

	Y1	X1	X2	X3	X4	X5	X6	X7
Mazowieckie	15,3	13,3	13,5	12,8	18,1	9,2	11,0	7,5
Opolskie	4,0	3,4	3,7	2,1	2,1	4,2	5,5	8,6
Podkarpackie	1,5	4,3	4,6	11,0	1,9	1,7	2,1	1,4
Podlaskie	7,2	7,0	6,4	5,3	15,9	3,4	5,7	2,7
Pomorskie	5,2	5,0	3,9	2,7	3,4	6,0	5,4	6,7
Śląskie	2,5	2,8	4,0	4,3	2,1	2,3	2,6	2,6
Świętokrzyskie	2,1	3,4	3,8	6,4	3,0	2,3	2,5	0,7
Warmińsko-mazurskie	7,0	6,9	5,6	2,8	8,1	4,5	6,1	7,2
Wielkopolskie	17,9	11,7	13,8	8,9	14,8	32,0	15,4	15,7
Zachodniopomorskie	4,7	5,9	4,7	1,9	1,7	2,7	6,2	9,3
Razem	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Źródło: badania własne

WYNIKI BADAŃ

Dla wszystkich województw obliczono indeksy TFP zgodnie z równaniem (4). Następnie przeprowadzono dekompozycję otrzymanych indeksów na kilka wybranych miar efektywności (5), (6), (7) i (8). Wyniki obliczeń wskaźników efektywności dla modeli zorientowanych na nakłady z wykorzystaniem zagregowanych funkcji Hicksa-Moorsteena przedstawione są w tabelach 3. i 4.

Średni poziom efektywności TFPE w analizowanych latach wahał się w granicach od 0,556 do 0,642. Najwyższy średni poziom tego wskaźnika wystąpił w roku 2010 a najniższy w 2011. Można zauważyć znaczne zróżnicowanie pomiędzy poszczególnymi województwami. Zdecydowanie najniższy poziom TFPE wystąpił w województwie wielkopolskim i znacznie odbiegał od pozostałych województw. Niski poziom tego wskaźnika odnotowano również w województwach: podkarpackim, mazowieckim i śląskim. Analizując układ zmiennych w tych województwach można zauważyć, że w przypadku województw wielkopolskiego i mazowieckiego udział ich w strukturze poszczególnych zmiennych był największy. Z kolei w przypadku województw śląskiego i podkarpackiego udział ten był najmniejszy. Najwyższymi wskaźnikami TFPE wykazały się w analizowanych latach województwa dolnośląskie, zachodniopomorskie, pomorskie, kujawsko-pomorskie, i warmińsko-mazurskie. Województwa te charakteryzują się proporcjonalnie wysokim udziałem produkcji zwierzęcej i niskim udziałem nawożenia w stosunku do wartości skupu produktów rolnych. Analiza poszczególnych regionów rolniczych wykazuje, że najwyższy średni wskaźnik TFPE występuje w badanych latach w regionie Pomorze i Mazury. Średni poziom efektywności w poszczególnych latach wahał się tutaj w granicach od 0,702 do 0,825. Wysoki poziom tego wskaźnika wykazują również województwa regionu Wielkopolska i Śląsk, ale z uwagi na bardzo niski poziom wskaźnika w województwie Wielkopolskim średnia jest tutaj zaniżona. Najniższy średni

wskaźnik TFPE występuje w regionie Małopolska i Pogórze. Średni poziom efektywności w poszczególnych latach wahał się tutaj w granicach od 0,702 do 0,825.

Analiza wskaźnika efektywności technicznej ITE wykazuje, że większość województw jest efektywna technicznie. Najniższy poziom efektywności ITE wynosił w analizowanych latach 0,893. Do województw nieefektywnych technicznie można zaliczyć województwa łódzkie, lubelskie i kujawsko-pomorskie. Może to sugerować, że w województwach tych poziom skupu jest poniżej oczekiwanych możliwości, ale oczywiście przy danym zestawie zmiennych. Można też zauważyć, że wszystkie województwa o najniższym poziomie wskaźnika TFPE są efektywne technicznie.

Tabela 3. Efektywność TFPE i efektywność techniczna ITE rolnictwa w Polsce w latach 2010-2012 obliczona z wykorzystaniem indeksu Hicksa-Moorsteena

Wyszczególnienie	TFPE			ITE		
	Rok			Rok		
	2010	2011	2012	2010	2011	2012
Pomorzanie i Mazury						
Lubuskie	0,691	0,607	0,533	1,000	1,000	1,000
Pomorskie	0,859	0,775	0,813	1,000	1,000	1,000
Warmińsko-mazurskie	0,869	0,745	0,880	1,000	1,000	1,000
Zachodniopomorskie	0,901	0,693	0,803	1,000	1,000	1,000
Wielkopolska i Śląsk						
Dolnośląskie	0,915	0,694	0,885	1,000	1,000	1,000
Kujawsko-pomorskie	0,851	0,748	0,761	0,956	0,928	1,000
Opolskie	0,649	0,733	0,787	1,000	1,000	1,000
Wielkopolskie	0,251	0,175	0,197	1,000	1,000	1,000
Mazowsze i Podlasie						
Lubelskie	0,737	0,537	0,755	0,981	0,933	0,920
Łódzkie	0,725	0,702	0,668	0,893	0,904	0,957
Mazowieckie	0,562	0,406	0,663	1,000	1,000	1,000
Podlaskie	0,829	0,590	0,827	1,000	1,000	1,000
Małopolska i Pogórze						
Małopolskie	0,618	0,496	0,528	1,000	0,959	1,000
Podkarpackie	0,445	0,357	0,366	1,000	1,000	1,000
Śląskie	0,318	0,571	0,590	1,000	1,000	1,000
Świętokrzyskie	0,649	0,564	0,624	1,000	1,000	1,000
Min	0,251	0,175	0,197	0,893	0,904	0,920
Max	0,915	0,775	0,885	1,000	1,000	1,000
Średnia	0,642	0,556	0,631	0,989	0,982	0,992

Źródło: obliczenia własne

Średnia efektywność skali (ISE) w analizowanych latach kształtowała się na poziomie 0,953 do 0,961. Najniższa efektywność skali wystąpiła w województwie podkarpackim i kształtowała się na poziomie 0,692 do 0,890. Analizując dane w tym województwie można zauważyć, że o ile udział tego województwa w wartości skupu produktów rolnych kształtuje się na poziomie 1,5% natomiast w wykorzystywanych nakładach i zasobach jest nieproporcjonalnie duży. Do województw o niskim poziomie tego wskaźnika można zaliczyć jeszcze województwa małopolskie i śląskie a także łódzkie. Najwyższa średnia efektywność skali w analizowanych latach występowała w rejonie Pomorze i Mazury natomiast najmniejsza w rejonie Małopolska i Pogórze.

Tabela 4. Efektywność skali ISE i efektywność typu mix (ITE) rolnictwa w Polsce w latach 2010-2012 obliczona z wykorzystaniem indeksu Hicksa-Moorsteena

	ISE			IME		
	Rok			Rok		
	2010	2011	2012	2010	2011	2012
Pomorze i Mazury						
Lubuskie	1,000	1,000	0,958	1,000	1,000	1,000
Pomorskie	1,000	1,000	1,000	0,927	1,000	0,967
Warmińsko-mazurskie	1,000	1,000	1,000	1,000	1,000	1,000
Zachodniopomorskie	1,000	1,000	1,000	0,960	0,904	1,000
Wielkopolska i Śląsk						
Dolnośląskie	1,000	1,000	1,000	0,995	1,000	1,000
Kujawsko-pomorskie	0,990	0,994	0,995	0,988	0,961	0,945
Opolskie	1,000	1,000	1,000	1,000	1,000	0,999
Wielkopolskie	1,000	1,000	1,000	1,000	1,000	1,000
Mazowsze i Podlasie						
Lubelskie	0,953	0,969	1,000	0,972	0,916	0,958
Łódzkie	0,955	0,943	0,943	0,932	0,968	0,948
Mazowieckie	1,000	1,000	1,000	1,000	1,000	1,000
Podlaskie	1,000	1,000	1,000	1,000	0,951	1,000
Małopolska i Pogórze						
Małopolskie	0,914	0,759	0,745	0,859	0,922	0,995
Podkarpackie	0,730	0,692	0,890	0,941	1,000	1,000
Śląskie	0,874	0,961	0,887	0,840	1,000	1,000
Świętokrzyskie	1,000	1,000	1,000	0,968	1,000	0,985
Min	0,730	0,692	0,745	0,840	0,904	0,945
Max	1,000	1,000	1,000	1,000	1,000	1,000
Średnia	0,961	0,953	0,961	0,960	0,976	0,987

Źródło: obliczenia własne

Kolejnym analizowanym wskaźnikiem jest efektywność typu mix (IME). Najwyższy stopień nieefektywności tego typu występuje w województwach małopolskim, lubelskim, łódzkim i kujawsko-pomorskim. W województwach tych istnieją największe możliwości poprawy efektywności poprzez alokację wykorzystywanych zasobów i nakładów. Najwyższa średnia efektywność mix w analizowanych latach występuje w regionie Wielkopolska i Śląsk natomiast najniższa w regionie Małopolska i Pogórze.

PODSUMOWANIA I WNIOSKI

W prezentowanych badaniach podjęto próbę wykorzystania indeksów TFP Hicksa-Moorsteena do badania produktywności i efektywności rolnictwa w Polsce. Obliczono indeksy produktywności a następnie zdekomponowano je na kilka miar efektywności. Otrzymane wyniki porównano w układzie województw i w podziale na 4 regiony rolnicze FADN.

Wykazano duże zróżnicowanie w poziomie efektywności TFPE pomiędzy poszczególnymi województwami i regionami. Najniższy poziom efektywności TFPE wystąpił w województwach wielkopolskim, podkarpackim, mazowieckim i śląskim. Na tym poziomie szczegółowości badań nie stwierdzono jednoznacznie przyczyn nieefektywności. Najwyższy średni wskaźnik TFPE występuje w badanych latach w regionie Pomorze i Mazury natomiast najniższy średni w regionie Małopolska i Pogórze. Analiza przyczyn zróżnicowania tego wskaźnika pomiędzy poszczególnymi województwami wymaga pełnej dekompozycji.

Stwierdzono małe zróżnicowanie wskaźnika efektywności technicznej ITE. Zdecydowaną większość województw można uznać za efektywne technicznie. Można zauważyć, że wszystkie województwa o najniższym poziomie wskaźnika TFPE są efektywne technicznie.

Najwyższa średnia efektywność skali (ISE) w analizowanych latach występowała w rejonie Pomorze i Mazury natomiast najmniejsza w rejonie Małopolska i Pogórze. Najniższa efektywność skali wystąpiła w województwie podkarpackim. Do województw o niskim poziomie tego wskaźnika można zaliczyć jeszcze województwa małopolskie i śląskie a także łódzkie. Można zauważyć, że województwa podkarpackie i śląskie, które miały jedne z najniższych wskaźników efektywności TFPE mają również nieefektywną skalę natomiast są efektywne technicznie. W tym przypadku trudno jest również jednoznacznie stwierdzić jakie są przyczyny nieefektywności. Konieczne są bardziej pogłębione badania.

Największe możliwości poprawy efektywności poprzez alokację wykorzystywanych zasobów i nakładów występują w województwach małopolskim, lubelskim, łódzkim i kujawsko-pomorskim. Najwyższa średnia efektywność mix w analizowanych latach występuje w regionie Wielkopolska i Śląsk natomiast najniższa w regionie Małopolska i Pogórze.

Badania wykazały, że na tym poziomie szczegółowości badań wykorzystanie otrzymanych wskaźników jest ograniczone. Można stwierdzić zróżnicowanie

pomiędzy województwami w poziomie poszczególnych wskaźników natomiast znalezienie przyczyn wymaga pogłębienia badań i analiz.

BIBLIOGRAFIA

- Balcombe K., Davidova S., Latruffe L. (2005) Productivity change in polish agriculture: An application of a bootstrap procedure to Malmquist indicies, *Materiały z konferencji: The Future of Rural Europe in the Global Agri-Food System*. Copenhagen, Denmark, August 24-27.
- Brümmer B., Glauben T., Thijssen G. (2002) Decomposition of productivity growth using distance functions: The case of dairy farms in three European countries, *American Journal of Agricultural Economics*, 84(3), 628-644.
- Diewert W.E. (1992) Fisher ideal output, input, and productivity indexes revisited. *Journal of Productivity Analysis* 3, 211–248.
- Coelli T.J., Rao D.S.P. (2003) Total factor productivity growth in agriculture: a Malmquist index analysis of 93 countries, 1980-2000. *Agricultural Economics* 32(s1), 115-134.
- Färe R., Grosskopf S., Lindgren, B., Roos P. (1992) Productivity changes in Swedish pharmacies 1980–1989: a nonparametric Malmquist approach, *Journal of Productivity Analysis* 3, 85–101.
- Färe R., Grosskopf S., Norris M., Zhang Z. (1994) Productivity growth, technical progress, and efficiency change in industrialized countries, *American Economic Review* 84, 66–83.
- Hoang V.N. (2011) Measuring and decomposing changes in agricultural productivity, nitrogen use efficiency and cumulative exergy efficiency: application to OECD agriculture, *Ecological Modelling* 222, 164–175.
- Latruffe L., Balcombe K., Davidova S., Zawalinska K. (2004) Determinants of technical efficiency of crop and livestock farms in Poland, *Applied Economics*, 36(12), 1255-1263.
- O'Donnell C. J. (2008) An Aggregate Quantity-Price Framework for Measuring and Decomposing Productivity and Profitability Change, *Centre for Efficiency and Productivity Analysis Working Papers WP07/2008*, University of Queensland.
- O'Donnell C.J. (2010) Measuring and decomposing agricultural productivity and profitability change, *Australian Journal of Agricultural and Resource Economics* 54, 527–560.
- O'Donnell C.J. (2011) DPIN version 3.0: a program for decomposing productivity index numbers, *Centre for Efficiency and Productivity Analysis*, University of Queensland, Brisbane.
- O'Donnell C.J. (2012a) An aggregate quantity framework for measuring and decomposing productivity change, *Journal of Productivity Analysis* 38 (3), 255–272.
- O'Donnell C.J. (2012b) Nonparametric estimates of the components of productivity and profitability change in U.S. agriculture, *American Journal of Agricultural Economics* 94, 873–890.
- Rusielik R., Świtłyk M. (2009) Zmiany efektywności technicznej rolnictwa w Polsce w latach 1998-2006, *Rocz. Nauk Roln., seria G*, t. 96, z. 3. 20-27.

- Rusielik R. (2014) Zmiany produktywności rolnictwa Polski po wstąpieniu do Unii Europejskiej – analiza z wykorzystaniem indeksów TFP Hicksa-Moorsteena., Rocz. Nauk SERiA, t. XVI, z. 4. 246-252.
- Zawalinska K. (2004) The Competitiveness of Polish Agriculture in the Context of Integration with the European Union, Praca doktorska, WNE UW, Warszawa.

**THE EFFICIENCY AND PRODUCTIVITY
IN POLISH AGRICULTURE
- ANALYSIS OF USING HICKS-MOORSTEEN TFP INDEX**

Abstract: The main purpose of the research discussed in this paper is used to calculate productivity and efficiency in agriculture with Hicks-Moorsteen Total Factor Productivity (TFP) indexes. Productivity and efficiency have been identified at the regional level, indicating the reasons for the changes. The calculation period covers 3 years (2010-2012). Hicks-Moorsteen (TFP) indexes was calculated for each Polish province. The TFP index decomposition analysis into measure of efficiency change allowed to identification determinants between periods of calculation.

Keywords: productivity, efficiency, agriculture, TFP, DEA, Hicks-Moorsteen index.

WYKORZYSTANIE MIAR POZYCYJNYCH W OCENIE POZIOMU ROZWOJU SPOŁECZNEGO W POLSCE

Agnieszka Sompolska-Rzechuła

Katedra Zastosowań Matematyki w Ekonomii
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: asompolska@zut.edu.pl

Streszczenie: Zróżnicowanie poziomu rozwoju społecznego w Polsce na poziomie lokalnym jest znacznie większe niż pomiędzy województwami. Właściwy pomiar wymaga wykorzystania odpowiednich mierników. Jednym z nich jest Wskaźnik Rozwoju Społecznego (HDI), a na poziomie lokalnym Lokalny Wskaźnik Rozwoju Społecznego (LHDI). Celem pracy jest ocena stopnia rozwoju społecznego na poziomie lokalnym. Wykorzystano miarę opartą na wskaźnikach dotyczących: zdrowia, edukacji i zamożności oraz wskaźnikach nakładów na: zdrowie, edukację oraz wydatki na poziomie lokalnym. Budowę miary syntetycznej oparto na medianie Webera.

Słowa kluczowe: poziom rozwoju społecznego, zmienna syntetyczna, mediana Webera

WPROWADZENIE

Dzięki działaniom Ministerstwa Rozwoju Regionalnego Polska została objęta nowatorskim badaniem dotyczącym mierzenia rozwoju społecznego kraju na poziomie lokalnym. Projekt na zlecenie MRR został realizowany przez ekspertów Programu Narodów Zjednoczonych ds. Rozwoju (UNDP) działającego przy ONZ.

Nowatorskość projektu polega przede wszystkim na tym, że badanie zostało oparte na Lokalnym Wskaźniku Rozwoju Społecznego (LHDI – *Local Human Development Index*), który został opracowany w celu mierzenia lokalnego i regionalnego poziomu rozwoju społecznego. Opiera się on na istniejącym już wskaźniku *Human Development Index* (HDI), służącym do badania tego zjawiska na poziomie krajowym. Dla jak najlepszego uwzględnienia wielowymiarowego charakteru rozwoju, metodologia opracowanego wskaźnika obejmowała trzy komponenty: zdrowie, edukację oraz zamożność.

Dokonując przeglądu badań poziomu rozwoju społecznego można zauważyć, że dotyczą one głównie krajów oraz większych miastach. Natomiast zróżnicowanie poziomu rozwoju społecznego w Polsce na poziomie lokalnym jest znacznie większe niż pomiędzy województwami i wymaga systematycznej oceny. Dzięki temu możliwe jest lepsze powiązanie celów strategii rozwoju województw oraz strategii lokalnych z celami krajowymi, przy zachowaniu odpowiedniej hierarchii. W pracy podjęto próbę wypełnienia luki w badaniach poziomu rozwoju społecznego na poziomie lokalnym, proponując wykorzystanie metod WAP w analizie badanego zjawiska.

Celem pracy jest ocena stopnia rozwoju społecznego na poziomie lokalnym poprzez wskazanie elementów dotyczących aspektów zdrowia, edukacji i zamożności, które miały decydujący wpływ na osiągnięcie określonego poziomu rozwoju. Do realizacji celu wybrano metodę liniowego porządkowania obiektów opartą na miarach pozycyjnych, w szczególności na medianie Webera. Badanie przeprowadzono w powiatach Polski, tworząc liniowe porządkowania obiektów w grupie powiatów ziemskich i grodzkich.

POMIAR POZIOMU ROZWOJU SPOŁECZNEGO Z WYKORZYSTANIEM WSKAŹNIKA ROZWOJU SPOŁECZNEGO

Od wielu lat toczy się na świecie debata dotycząca zagadnień obejmujących rozwój społeczny oraz takie kategorie, jak: nierówności społeczne, zrównoważony rozwój, dobrobyt, dobrostan oraz pomiar poziomu rozwoju w sposób inny niż tylko przez wymiar ekonomiczny. Rozwój społeczny powinno się zatem analizować na podstawie szerokiego zbioru czynników, opisujących różne sfery działalności życia człowieka¹: ekonomiczne, społeczne, demograficzne i przyrodnicze.

Przez długi czas za najważniejszy wskaźnik rozwoju społecznego uważany był produkt krajowy brutto (PKB). Jednak coraz częściej podważane jest stosowanie wartości PKB jako głównego wskaźnika rozwoju społecznego². Prace nad modyfikacją PKB w celu dostosowania go do pomiaru rozwoju społecznego polegały między innymi na eliminacji tych jego elementów, które nie przyczyniają się bezpośrednio lub pośrednio, do zaspokojenia szeroko pojmowanych potrzeb konsumpcyjnych (na przykład nakłady na zbrojenia, wydatki gospodarstw domowych na dojazdy do pracy itd.), przy jednoczesnym włączeniu do obliczeń szeregu dodatkowych elementów mających istotny wpływ na poziom i jakość życia (na przykład wartość czasu wolnego, wartość dóbr i usług nierejestrowanych przez rynek). W latach 90-tych XX wieku zainteresowanie alternatywą lub uzupełnieniem PKB stopniowo wzrastało czego rezultatem było opracowanie Wskaźnika Rozwoju Społecznego (HDI).

¹ Więcej informacji można znaleźć w pracy [Berbeka 2006].

² Takie podejście jest reprezentowane m.in. przez komisję Stiglitz i Fitoussiego [Szukielojć-Bieńkuńska, Walczak 2011].

Podstawowym założeniem, leżącym u podstaw metodologii badawczej Wskaźnika Rozwoju Społecznego (HDI) jest rozumienie rozwoju społecznego jako zjawiska niezwykle złożonego, którego nie da się oddać tylko za pomocą rankingu dochodów. Uwzględnia on te czynniki, które mają wpływ na jakość życia i poczucie szczęścia, a których nie da się w prosty sposób przeliczyć na kategorie materialne.

Oryginalna metodologia HDI³ (stosowana w tworzeniu wskaźnika HDI na poziomie krajów) sugeruje, że pomiar rozwoju społecznego powinien koncentrować się na trzech istotnych elementach ludzkiego życia: długości życia (zdrowie), wiedzy (edukacja) oraz takiego standardu, który pozwala na godne życie (reprezentowanego przez poziom dochodu - zamożności). Miernik ten zdefiniowano jako średnią geometryczną trzech wskaźników: długowieczności (reprezentowanej przez przeciętne dalsze trwanie życia w momencie narodzin), wykształcenia (opisanego przez wskaźnik umiejętności pisania i czytania ze zrozumieniem wśród dorosłych (bez względu na płeć) oraz ogólny wskaźnik skolaryzacji brutto) i standardu życia (życia mierzonego wielkością PKB per capita w dolarach amerykańskich USD).

LHDI przyjmuje podobną formułę i wykorzystuje następujące dane⁴:

- zdrowie: oczekiwana dalsza długość trwania życia noworodka, zagregowany współczynnik zgonów na nowotwory i choroby serca,
- edukacja: odsetek dzieci w edukacji przedszkolnej (przedział wiekowy 3-4 lata), średnia z wyników egzaminu gimnazjalnego (tylko dla części matematyczno-przyrodniczej),
- zamożność: średni poziom zamożności mieszkańców.

Podstawową jednostką analizy rozwoju lokalnego, ze względu na bliskość usług publicznych zdrowotnych, rynku pracy oraz edukacji na poziomie średnim i wyższym, jest powiat.

Konstrukcja Lokalnego Wskaźnika Rozwoju Społecznego na poziomie powiatu jest następująca:

$$LHDI_i = \sqrt[3]{HI_i \cdot EI_i \cdot WI_i} \quad i = 1, 2, \dots, n, \quad (1)$$

gdzie:

$LHDI_i$ – wartość wskaźnika rozwoju społecznego dla i -tego powiatu,

HI_i – wskaźnik zdrowia w i -tym powiecie,

EI_i – wskaźnik edukacji w i -tym powiecie,

WI_i – wskaźnik zamożności w i -tym powiecie,

n – liczba powiatów.

³ Więcej informacji można znaleźć na stronie hdr.undp.org/en.

⁴ Szczegółowe informacje na temat HDI i LHDI oraz sposobów liczenia wskaźnika i jego składowych znajdują się w Krajowym Raporcie o Rozwoju Społecznym dostępnym na www.mir.gov.pl. W Raporcie zawarte są także koncepcje rozwoju społecznego ze szczególnym uwzględnieniem podejścia reprezentowanego przez A. Sena.

Wskaźnik zdrowia jest średnią geometryczną dwóch wskaźników cząstkowych: wskaźnika przeciętnego trwania życia (*LEI*) i zagregowanego współczynnika zgonów na nowotwory i choroby serca (*CDRI*):

$$HI_i = \sqrt{LEI_i \cdot CDRI_i} \quad i = 1, 2, \dots, n, \quad (2)$$

Wskaźnik edukacji składa się ze średniej geometrycznej dwóch wskaźników cząstkowych: wskaźnika edukacji przedszkolnej (*PEI*) i wskaźnika wyników egzaminu gimnazjalnego (*PLSSEI*):

$$EI_i = \sqrt{PEI_i \cdot PLSSEI_i} \quad i = 1, 2, \dots, n. \quad (3)$$

Wskaźnik zamożności to suma dochodu podatników ogółem przed opodatkowaniem (zeznania podatkowe: PIT-36, PIT-36L, PIT-37) plus dochód z rolnictwa na bazie hektarów przeliczeniowych (wpływy gmin z podatku rolnego podzielone przez wartość podatku z jednego hektara przeliczeniowego pomnożone przez średni dochód z 1 ha przeliczeniowego z działalności indywidualnych gospodarstw rolniczych, które stanowią podstawę podatku rolniczego dla gruntów rolnych)⁵ oraz suma wydatków na pomoc społeczną oraz inne zadania polityki społecznej: suma świadczeń społecznych oraz wydatków na politykę rodzinną w budżetach lokalnych władz samorządowych (gminnych oraz powiatowych) z wyłączeniem świadczeń z tytułu klęsk żywiołowych podzielona przez liczbę mieszkańców powiatu. Inaczej ujmując jest to średni poziom zamożności mieszkańców.

Wyznaczenie wartości wskaźnika zamożności WI_i dla powiatów zostało dokonane na podstawie wzoru:

$$WI_i = 1 + 99 \cdot \frac{WI_{ei} - WI_{min}}{WI_{max} - WI_{min}} \quad i = 1, 2, \dots, n, \quad (4)$$

gdzie:

WI_{ei} – średni poziom zamożności mieszkańców w i -tym powiecie,

WI_{min} , WI_{max} – wartości odpowiednio minimalne (zaobserwowane w badanych powiatach w latach 2007-2010) i maksymalne (zaobserwowane w badanych powiatach w latach 2007-2010) średniej zamożności mieszkańca: $min=663,5436$ (powiat kolneński 2007), $max=2687,3660$ (Warszawa 2008).

Elementem wyróżniającym prezentowane podejście jest próba skonstruowania Wskaźnika Rozwoju Społecznego, który pokazywałby różne nakłady polityki publicznej na rozwój społeczny.

Wzór na wskaźnik nakładów polityk publicznych z uwzględnieniem wymiarów LHDI jest następujący:

$$LHDI_{PI_i} = \sqrt[3]{HI_{PI_i} \cdot EL_{PI_i} \cdot LEI_{PI_i}} \quad i = 1, 2, \dots, n, \quad (5)$$

⁵ Wartość jednego hektara przeliczeniowego ustalana jest na podstawie powierzchni, typu i klasy użytków rolnych na podstawie ksiąg wieczystych, a także na podstawie dodatku do podatków powiatu.

gdzie:

$LHDI_{pi}$ – nakłady publiczne na wymiary rozwoju społecznego dla i -tego powiatu,

HI_{pi} – wskaźnik nakładów zdrowotnych w i -tym powiecie,

EI_{pi} – wskaźnik nakładów edukacyjnych w i -tym powiecie,

LEI_{pi} – wskaźnik wydatków lokalnych w i -tym powiecie.

W obliczeniach miary nakładów na sferę zdrowia wykorzystane zostały następujące wskaźniki:

- liczba lekarzy i lekarzy dentyków według podstawowego miejsca pracy na 100 tys. osób ($MDCI_i$)
- liczba pielęgniarek i położnych podstawowej opieki zdrowotnej według podstawowego miejsca pracy na 100 tys. ($NMCI_i$).

Wskaźniki te pokazują stopień zaangażowania zasobów ludzkich w świadczenie opieki zdrowotnej. Podczas gdy liczba lekarzy podstawowej opieki medycznej pokazuje ogólną dostępność opieki zdrowotnej, liczba pielęgniarek i położnych w ważny sposób różnicuje jakość opieki zapewnianej przez placówki medyczne w danym regionie. Wskaźnik nakładów zdrowotnych przyjmuje następującą postać:

$$HI_{pi} = \sqrt{MDCI_i \cdot NMCI_i} \quad i = 1, 2, \dots, n, \quad (6)$$

natomiast wskaźnik nakładów edukacyjnych obliczany jest według następującego wzoru:

$$EI_{pi} = \sqrt{EEI_i \cdot STRI_i} \quad i = 1, 2, \dots, n, \quad (7)$$

gdzie:

EEI_i – wskaźnik wydatków na edukację wyznaczany na podstawie sumy wydatków samorządów (gmin i powiatów) na edukację podzieloną przez liczbę uczniów (przedszkola, szkoły podstawowe, gimnazja, szkoły średnie) w danym powiecie,

$STRI_i$ – wskaźnik liczby uczniów na nauczyciela wyznaczany na podstawie zsumowanej liczby uczniów szkół podstawowych i gimnazjów podzielonej przez zsumowaną liczbę nauczycieli w tych szkołach w danym i -powiecie.

Wskaźnik wydatków lokalnych przyjmuje postać sumy nakładów finansowych samorządu terytorialnego per capita. Punktem wyjścia wyznaczenia wartości wskaźnika dla badanych powiatów jest następujący wzór:

$$LEI_{pi} = 1 + 99 \cdot \frac{LEI_{pei} - LEI_{pmin}}{LEI_{pmax} - LEI_{pmin}} \quad i = 1, 2, \dots, n, \quad (8)$$

gdzie:

LEI_{pei} – wydatki samorządów na mieszkańca w i -tym powiecie,

LEI_{pmin} , LEI_{pmax} – wartości odpowiednio minimalne (zaobserwowane w badanych powiatach w latach 2007-2010) i maksymalne (zaobserwowane w badanych powiatach w latach 2007-2010) wydatków samorządów w przeliczeniu na mieszkańca: $min=2049,7501$ (Piekary Śląskie 2007), $max=9454,3677$ (Sopot 2007).

METODA BADAWCZA

Poziom rozwoju społecznego jako zjawisko ekonomiczne ma charakter złożony i określany jest przez wiele cech pochodzących z różnych dziedzin życia. Do analizy poziomu rozwoju społecznego można zatem wykorzystać metody WAP. Wśród nich znajdują się metody liniowego porządkowania obiektów, które stanowią narzędzie do konstrukcji rankingu obiektów. W budowie tzw. syntetycznej miary rozwoju wykorzystywane są dwa podejścia:

- bezwzorcowe oraz
- oparte na konstrukcji wzorca rozwoju.

W drugim podejściu wykorzystywane są miary pozycyjne a szczególną koncepcją jest uogólniona wersja mediany – mediana Webera. Stosowanie miar pozycyjnych jest uzasadnione w przypadkach gdy cecha ma rozkład asymetryczny. Mediana jako podstawowy parametr pozycyjny charakteryzuje się wysoką odpornością na występowanie obserwacji odstających i jest stosowana na etapie normalizacji cech, w przekształceniu standaryzacyjnym. Standaryzacja cech z wykorzystaniem miar pozycyjnych jest przeprowadzana według następującego wzoru⁶:

$$z_{ij} = \frac{x_{ij} - \text{med}(X_j)}{1,4826 \cdot \text{mad}(X_j)} \quad (9)$$

gdzie: $\text{med}(X_j)$ – mediana cechy X_j , $\text{mad}(X_j) = \text{med}_{i=1, \dots, n} |x_{ij} - \text{med}(X_j)|$ – medianowe odchylenie bezwzględne, n – liczba obiektów, 1,4826 – stała ustalona na drodze badań empirycznych.

Natomiast normalizacja cech z zastosowaniem wielowymiarowej mediany Webera polega na zastąpieniu we wzorze (9) mediany cechy $\text{med}(X_j)$ medianą Webera⁷. Wzór na normalizację ma wówczas następująca postać:

$$z_{ij} = \frac{x_{ij} - \text{m}\ddot{\text{e}}\text{d}(X_j)}{1,4826 \cdot \text{m}\ddot{\text{a}}\text{d}(X_j)} \quad (10)$$

gdzie: $\text{m}\ddot{\text{e}}\text{d}(X_j)$ – mediana Webera, $\text{m}\ddot{\text{a}}\text{d}(X_j) = \text{med}_{i=1, \dots, n} |x_{ij} - \text{m}\ddot{\text{e}}\text{d}(X_j)|$ – medianowe odchylenie bezwzględne badające dystanse cech do mediany Webera.

W kolejnym kroku buduje się miernik rozwoju o postaci:

$$\mu_i = 1 - \frac{d_i}{d_{\max}} \quad (11)$$

gdzie:

⁶ Por. [Lira J. i in. 2002].

⁷ Mediana Webera stanowi wielowymiarowe uogólnienie tego powszechnie znanego pojęcia. Chodzi tu o wektor, który minimalizuje sumę euklidesowych odległości od danych punktów reprezentujących rozpatrywane obiekty, a więc znajduje się niejako „pośrodku” nich, ale jest jednocześnie uodporniony na występowanie obserwacji odstających [Młodak 2010].

$$d_i = \sqrt{\sum_{j=1}^J (z_{ij} - z_{0j})^2} \quad (i = 1, 2, \dots, n), \quad (12)$$

natomiast z_{ij} są zestandaryzowanymi wartościami cech według wzoru (10), z_{0j} są współrzędnymi wzorca rozwoju, gdzie, jeśli cecha jest destymulantą to $z_{0j} = \min_i \{z_{ij}\}$ a gdy cecha jest stymulantą to $z_{0j} = \max_i \{z_{ij}\}$, ($j = 1, 2, \dots, J$), J – liczba cech. Miernik agregatowy oblicza się według wzoru (11), w którym:

$$d_- = \text{med}(\mathbf{d}) + 2,5 \cdot \text{mad}(\mathbf{d}), \quad (13)$$

gdzie:

$\mathbf{d} = (d_1, \dots, d_n)$ jest wektorem odległości wyznaczanym według wzoru:

$$d_i = \text{med}_{j=1, \dots, J} |z_{ij} - \varphi_j|, \quad \varphi_j = \min_{i=1, \dots, n} z_{ij} \text{ – dla destymulant lub } \varphi_j = \max_{i=1, \dots, n} z_{ij} \text{ – dla stymulant, } (i = 1, 2, \dots, n).$$

Im wyższa wartość miernika wyznaczonego za pomocą wzoru (11) tym wyższy poziom rozwoju obiektu.

Zaprezentowana metoda została wykorzystana do oceny poziomu rozwoju społecznego na poziomie lokalnym, na przykładzie powiatów Polski.

WYNIKI

Do oceny poziomu rozwoju społecznego na poziomie lokalnym, wykorzystano powiaty Polski. Zbudowano ranking dla dwóch grup powiatów: ziemskich i grodzkich, których jest w Polsce odpowiednio 314 i 65. Rolę cech diagnostycznych pełniły następujące wskaźniki dotyczące 2010 roku, zawarte w Krajowym Raportie o Rozwoju Społecznym, Polska 2012, Rozwój Regionalny i Lokalny: *HI* – Wskaźnik Zdrowia, *EI* – Wskaźnik Edukacji, *WI* – Wskaźnik Zamożności, *HIFI* – Wskaźnik Nakładów Zdrowotnych, *EIPI* – Wskaźnik Nakładów Edukacyjnych, *LEIPI* – Wskaźnik Wydatków Lokalnych.

Wartości wskaźników poddano analizie statystycznej, obliczając wartości podstawowych parametrów opisowych. Wszystkie wskaźniki charakteryzowały się silną zmiennością, wyrażoną za pomocą współczynnika zmienności, który wyniósł dla każdej cechy powyżej 20%. W kolejnym kroku sprawdzono siłę korelacji pomiędzy cechami i metodą parametryczną wyeliminowano te cechy, które są silnie ze sobą powiązane, oddzielnie dla powiatów ziemskich i grodzkich.

W efekcie otrzymano następujące zbiory cech diagnostycznych dla powiatów:

- ziemskich: *EI*, *EIPI*, *HI*, *HIFI*,
- grodzkich: *EI*, *EIPI*, *HI*,

dla których współczynniki korelacji były na poziomie niższym niż 0,5.

Otrzymane cechy charakteryzują się silną asymetrią, np. współczynnik skośności dla wskaźnika *EIPI* w przypadku powiatów ziemskich wyniósł 1,17, a w przypadku powiatów grodzkich – 4,12, dlatego do liniowego porządkowania powiatów wykorzystano metodę opartą na mierze pozycyjnej – medianie Webera.

Otrzymane zbiory cech wykorzystano do konstrukcji rankingu powiatów ziemskich i grodzkich. Ze względu na bardzo dużą liczbę powiatów ziemskich w tabeli 1 zaprezentowano fragment rankingu tych powiatów.

Tabela 1. Fragment liniowego porządkowania powiatów ziemskich

Lp.	Powiat	Wartość miernika	Lp.	Powiat	Wartość miernika
1.	Otwocki	0,7056	190.	Drawski	0,1949
2.	Warszawski Zachodni	0,6506	191.	Żniński	0,1948
3.	Oświęcimski	0,6145	192.	Świdwiński	0,1944
4.	Cieszyński	0,5913	193.	Wschowski	0,1940
5.	Stalowowolski	0,5847	194.	Ostrowski	0,1933
....
45.	Kołobrzeski	0,4158	250.	Tczewski	0,1312
46.	Olecki	0,4075	251.	Wrocławski	0,1307
47.	Krapkowicki	0,4065	252.	Makowski	0,1307
...
100.	Bieruńsko-Lędziński	0,3114	278.	Bydgoski	0,0826
101.	Pabianicki	0,3098	279.	Płocki	0,0772
102.	Słubicki	0,3086	280.	Wągrowiecki	0,0763
...
154.	Piski	0,2381	312.	Chełmski	-0,1184
155.	Lidzbarski	0,2379	313.	Koniński	-0,1197
156.	Tomaszowski	0,2374	314.	Łomżyński	-0,1555

Źródło: opracowanie własne

Największą wartość miernika rozwoju wśród powiatów ziemskich posiadał powiat otwocki. Na drugim miejscu znajdował się powiat warszawski zachodni. Pozycja powiatu otwockiego wynika z korzystnych wartości w przypadku takich wskaźników, jak: zdrowia, edukacji i nakładów zdrowotnych i ich składowych, które swoimi wartościami znacznie przewyższają średnie dla wszystkich powiatów ziemskich i są, w przypadku, wskaźników edukacji i nakładów zdrowotnych jednymi z najwyższych.

Dla dwunastu powiatów ziemskich miernik jest wartością ujemną, a najniższą wartość przyjął dla powiatu łomżyńskiego. Tak niska pozycja związana jest z najniższą wartością wskaźnika nakładów zdrowotnych wśród wszystkich powiatów ziemskich a co za tym idzie niskimi wartościami liczby lekarzy i lekarzy dentyistów według podstawowego miejsca pracy na 100 tys. osób oraz liczbą pielęgniarek i położnych podstawowej opieki zdrowotnej na 100 tys. i jedną z najniższych wartości wskaźnika edukacji.

W Tabeli 2 przedstawiono fragment liniowego porządkowania powiatów grodzkich.

Tabela 2. Liniowe porządkowanie powiatów grodzkich

Lp.	Powiat	Wartość miernika	Lp.	Powiat	Wartość miernika
1.	Sopot	0,8474	34.	Tychy	0,4173
2.	Białystok	0,7763	35.	Katowice	0,4144
3.	Krosno	0,7646	36.	Płock	0,4122
4.	Rzeszów	0,7636	37.	Konin	0,3953
5.	Opole	0,7198	38.	Legnica	0,3813
....	42.	Ostrołęka	0,3151
10.	Siedlce	0,6756	43.	Skierniewice	0,2936
11.	Gdynia	0,6706	44.	Świnoujście	0,2892
12.	Kielce	0,6341	45.	Chorzów	0,2658
....
20.	Łomża	0,5290	53.	Mysłowice	0,2169
21.	Bielsko-Biała	0,5169	54.	Rybnik	0,2140
30.	Gdańsk	0,4525	63.	Bytom	0,0056
31.	Przemysł	0,4399	64.	Siemianowice Śląskie	-0,0511
32.	Toruń	0,4389	65.	Świętochłowice	-0,1791
33.	Bydgoszcz	0,4356			

Źródło: opracowanie własne

Największą wartość miernika rozwoju wśród powiatów grodzkich posiada Sopot. Na drugim miejscu znajduje się Białystok. Pozycja Sopotu wynika z najwyższej wartości dla wskaźnika nakładów edukacyjnych oraz wysokiej wartości wskaźników: edukacji i zdrowia. Natomiast ostatnia pozycja Świętochłowic jest związana z najniższymi wartościami takich wskaźników, jak: edukacji i nakładów na edukację.

PODSUMOWANIE

Celem pracy była ocena stopnia rozwoju społecznego na poziomie lokalnym. Utworzono dwa liniowe porządki powiatów: ziemskich i grodzkich. Podział ten ma kluczowe znaczenie, ponieważ miasta na prawach powiatu mają inne charakterystyki niż powiaty ziemskie. Do realizacji celu wykorzystano Lokalny Wskaźnik Rozwoju Społecznego (LHDI) i jego składowe, który został opracowany do oceny lokalnego i regionalnego poziomu rozwoju społecznego. Przygotowanie i wykorzystanie wskaźnika HDI na poziomie lokalnym przez ekspertów opracowujących Krajowy Raport o Rozwoju Społecznym jest przedsięwzięciem nowatorskim na skalę międzynarodową. Wpisuje się też w globalne trendy w zakresie nowych sposobów mierzenia rozwoju, w szczególności w inicjatywy Komisji Europejskiej i OECD. Dla jak najlepszego uwzględnienia wielowymiarowego charakteru rozwoju, metodologia opracowanego wskaźnika obejmowała trzy komponenty: zdrowie, edukację i zamożność. Jest to zatem miara pozwalająca na pełniejszą ocenę poziomu rozwoju społecznego w porównaniu do PKB na jednego mieszkańca, ponieważ łączy wymiar społeczny i ekonomiczny.

Ze względu na złożony charakter badanego zjawiska, do badania wykorzystano liniowe porządkowanie obiektów. Większość cech przyjętych w badaniu miała asymetryczny rozkład, zatem wykorzystano miary pozycyjne, w tym wielowymiarowe uogólnienie mediany – medianę Webera, dzięki czemu podejście zaprezentowane w pracy można uznać za nowatorskie. Powiaty w Polsce są silnie zróżnicowane przestrzennie. Na podstawie przeprowadzonego badania stwierdzono, że powiaty ziemskie o najwyższych wartościach miernika znajdują się przede wszystkim na południu Polski oraz wokół dużych miast, natomiast powiaty ziemskie, dla których wartości miernika były najniższe tworzą skupienia w różnych częściach kraju. Powiaty grodzkie o najwyższej wartości zmiennej agregatowej to przede wszystkim duże miasta, a najwięcej powiatów grodzkich o najniższym poziomie miernika znajduje się na Górnym Śląsku.

Metody WAP okazały się użytecznym narzędziem w ocenie stopnia rozwoju społecznego na poziomie lokalnym.

BIBLIOGRAFIA

- Berbeka J. (2006) Poziom życia ludności a wzrost gospodarczy w krajach Unii Europejskiej, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Lira J., Wagner W., Wysocki F. (2002) Mediana w zagadnieniach porządkowania obiektów wielocechowych, w: Statystyka regionalna w służbie samorządu lokalnego i biznesu, Internetowa Oficyna Wydawnicza Centrum Statystyki Regionalnej, Akademia Ekonomiczna w Poznaniu, Poznań, str. 87 - 99.
- Młodak A. (2006) Analiza taksonomiczna w statystyce regionalnej, Difin, Warszawa, str. 7-23.
- Szukielójć-Bieńkuńska A., Walczak T. (2011) Statystyczny pomiar postępu społeczno-gospodarczego w zmieniającym się świecie, Wiadomości Statystyczne 7-8, str. 9-29.
www.hdr.undp.org/sites/default/files/hdr_2013_en_technotes.pdf.
www.mir.gov.pl/aktualnosci/polityka_rozwoju/Documents/raport_undp_2012_www.pdf.

THE USE OF POSITIONAL MEASURES IN THE ASSESSMENT OF LEVEL OF SOCIAL DEVELOPMENT IN POLAND

Abstract: Diversity of level of social development in Poland on the local level is much higher than between regions. Correct measurement requires the use of suitable measurers. One of the most common measures used to assess the level of social development is *Human Development Index* and *Local Human Development Index*. The aim of the paper is the assessment of level of social development on the local level. There was used the synthetic variable, with indicators for: health, education and wealth as well as health expenditures index, educational expenditures index and local expense index. The construction of synthetic measure based on the median Weber.

Keywords: level of social development, synthetic variable, Weber's median

RELACJA MIĘDZY EKSPORTEM ROLNYM I EKSPORTEM OGÓŁEM KRAJÓW UE

Jacek Strojny

Katedra Statystyki Matematycznej
Uniwersytet Rolniczy w Krakowie
e-mail: rrstrojn@cyf-kr.edu.pl

Streszczenie: Opracowanie ma na celu badanie relacji między eksportem produktów rolno-żywnościowych oraz eksportem ogółem krajów UE. Podjęto próbę zweryfikowania hipotezy o współzależności między tymi procesami. Do oceny stopnia współzależności między zmiennymi wykorzystano metody analizy szeregów czasowych.

Słowa kluczowe: eksport rolny, eksport ogółem, analiza kointegracji, dynamiczny model panelowy

WPROWADZENIE

Międzynarodowa wymiana ekonomiczna znajduje istotne odzwierciedlenie w gospodarce każdego kraju (Johnson 2013]. Znacznym zmianom otoczenia międzynarodowego, jak postępujące procesy liberalizacji, towarzyszył wzrost handlu światowego [Farmer, Schelnast 2013]. Jednocześnie, kryzys finansowy z końca lat 90-tych ubiegłego wieku i wynikłe spowolnienie gospodarcze odcisnęły swoje piętno na międzynarodowych relacjach ekonomicznych. Wyraźne spadki tempa wzrostu gospodarczego stworzyły poważne implikacje dla dostawców na rynek międzynarodowy – skutkowały znacznymi ograniczeniami eksportu. Zakres redukcji potoków handlu międzynarodowego był zróżnicowany w zależności od grupy towarowej.

Dla oceny wpływu handlu międzynarodowego na rozwój gospodarki narodowej ważniejsze, od skutków zaburzeń incydentalnych, są tendencje długookresowe [Van den Berg, Lewer 2007]. W ostatnich latach większość krajów członkowskich UE systematycznie zwiększała wartość swego eksportu. Niektóre z tych państw odnotowują systematyczne wzrosty wartości eksportu towarów sektora rolnego-spożywczego. Na ogół, wobec pogarszających się warunków

handlu międzynarodowego obserwowano korzystniejsze tendencje kształtowania się eksportu rolnego niż eksportu ogółem.

Opracowanie ma na celu badanie charakteru zależności między eksportem produktów rolno-żywnościowych oraz eksportem ogółem oraz weryfikację hipotezy o współzależności tych procesów. Do określenia stopnia współzależności eksportu rolnego i eksportu ogółem wykorzystano metody analizy szeregów czasowych. Badania oparto o dane statystyczne w ujęciu wartościowych, pozyskane z bazy EUROSTAT [2014]. Dane mają charakter roczny, a analiza obejmuje lata 1994-2012.

METODA

Własności predykcyjne prostych metod szeregów czasowych nie ustępują złożonym modelom strukturalnym lub nawet je przewyższają [Cooper 1972]. Modele szeregów czasowych w miejsce opisu statycznych relacji przyczynowo-skutkowych podejmują problem procesów stochastycznych będących u podstaw kształtowania się zmiennych [Box, Jenkins 1970]. Słabością metod analizy szeregów czasowych jest niezadawalający stopień możliwości objaśniania badanych relacji.

Większość zmiennych wykorzystywanych w modelach ekonomicznych charakteryzuje się niestacjonarnymi przebiegami, co może prowadzić do ustalania pozornych relacji, ale także oddziałuje na własności estymatorów. Engle i Granger [1987] pomogli w usunięciu głównej przeszkody w modelowaniu niestacjonarnych zmiennych przez wprowadzenie idei kointegracji oraz modeli korekty błędem (ECM). Przełomu w metodologii modelowania szeregów czasowych dokonał Sims [1980] proponując alternatywę dla modelowania wielorównaniowego – model wektorowo autoregresyjny (VAR).

O ile tradycyjna ekonometria pomijała na ogół niestacjonarność szeregów czasowych (w szczególności niestacjonarność w wariancji) w modelach VAR badanie stacjonarności jest punktem wyjścia procesu modelowania. Phillips [1986] wykazał, że wyniki estymacji modeli dla niestacjonarnych zmiennych na ogół budzą wątpliwości i prowadzą do identyfikacji pozornych relacji. Modele VAR umożliwiają nie tylko tworzenie dynamicznych prognoz, ale są platformą do badania zależności między zmiennymi i testowania zjawiska kointegracji.

Modele VAR mają charakter wielorównaniowy. Każda ze zmiennych jest objaśniana przez opóźnienia (własne oraz innych zmiennych). W podstawowej formie model VAR przyjmuje postać [Kusideł 2000]:

$$x_t = A_0 D_t + A_1 x_{t-1} + A_2 x_{t-2} + \dots + A_k x_{t-k} + e_t, \quad (1)$$

Natomiast w formie zwartej równanie można zapisać:

$$x_t = A_0 D_t + \sum_{i=1}^k A_i x_{t-i} + e_t, \quad t = 1, 2, \dots, T; \quad (2)$$

gdzie:

- x_t – wektor obserwacji na wartościach n zmiennych modelu $x_t = [x_{1t} \ x_{2t} \ \dots \ x_{nt}]'$,
- D_t – wektor deterministycznych składników równań (wyraz wolny, zmienna czasowa, itp.),
- A_0 – macierz parametrów przy zmiennych wektora D_t ,
- A_i – macierze parametrów przy opóźnionych zmiennych wektora x_t (nie zawierają elementów zerowych),
- e_t – wektory stacjonarnych zakłóceń losowych: $e_t = [e_{1t} \ e_{2t} \ \dots \ e_{nt}]'$.

Modele VAR wykazują cenne własności w zakresie opisu związków i prognozowania, jednak wymagają użycia stacjonarnych zmiennych. Różnicowanie jako przekształcenie danych celem osiągnięcia stacjonarności nie pozwala na poszukiwanie tendencji długookresowych. Uzupełnienie równań tworzonych na przyrostach zmiennych o składnik korekty błędem umożliwia modelowanie relacji długookresowych dla zmiennych niestacjonarnych. Jednak przed podjęciem prób opisu należy stwierdzić czy między niestacjonarnymi zmiennymi występuje relacja długookresowa – czy zmienne są skointegrowane [Engel, Granger 1987].

Niniejsze badanie obejmuje takie etapy, jak:

1. Badanie stacjonarności zmiennych.
2. Analizę kointegracji, która pozwala na określenie rodzaju relacji między zmiennymi.
3. Analizę przyczynowości – dynamiczny model panelowy.

Box i Jenkins [1970] określają stacjonarność szeregu czasowego jako pewien stan statycznej równowagi. Wiedza odnośnie stopnia stacjonarności jest istotna ponieważ większość testów dla modeli VAR nie ma pożądanych rozkładów dla niestacjonarnych zmiennych. W badaniu stacjonarności odwołano się do hipotezy pierwiastków jednostkowych, która zakłada, że niestacjonarność wynika z niestajej wariacji procesów w czasie. Analizę taką przeprowadzono z wykorzystaniem testu Kwiatkowskiego–Phillipsa–Schmidta–Shina (KPSS) [Kwiatkowski i in. 1992].

Mimo niewątpliwych zalet modele VAR można wykorzystywać jedynie dla stacjonarnych zmiennych. Różnicowanie celem doprowadzenia do stacjonarności eliminuje z danych tendencje długookresowe. Warunkowo, model zbudowany na przyrostach zmiennych można uzupełnić o składnik korekty błędem, który reprezentuje relacje długookresowe między niestacjonarnymi zmiennymi. Modele zawierające składnik korekty błędem (określane terminem ECM) umożliwiają rozdzielenie tendencji krótko- i długookresowych.

Zasadniczym problemem w zastosowaniu modeli VAR jest zatem ocena czy między niestacjonarnymi zmiennymi występuje relacja długookresowa. Jeżeli takowy związek zachodzi to określa się go terminem kointegracji. Powszechnie znaną definicją kointegracji jest koncepcja Engle'a i Grangera [1987]. W ekonometrii najbardziej interesująca może okazać się sytuacja, w której szereg

czasowy transformowany z udziałem wektora kointegrującego staje się stacjonarny. Wtedy współczynniki wektora kointegrującego można kojarzyć z długookresowymi relacjami między zmiennymi.

Metoda testowania kointegracji Engle'a i Grangera mimo iż prosta spotyka się szeroką krytyką z uwagi na pewne niedostatki [Kusideł 2000]. Przełomu w badaniach kointegracji dokonały prace Johansena [Johansen 1995]. Autor po ustaleniu, że zmienne są zintegrowane rzędu pierwszego wykorzystał do dalszego badania współzależności między zmiennymi model korekty błędem. Rozwinięcie tej koncepcji do badania związków długookresowych, które bazuje na modelach VAR [Johansen 1998] eliminuje większość zastrzeżeń wysuwanych pod adresem metody Engle'a i Grangera.

Przed zastosowaniem modelu VAR do badania relacji długookresowej konieczne jest jego przekształcenie do postaci modelu korekty błędem (VECM). Od obu stron równania (1) odejmowany jest wyraz x_{t-1} . Zatem, otrzymujemy:

$$\Delta x_t = A_0 D_t + (A_1 - I)A_1 x_{t-1} + A_2 x_{t-2} + \dots + A_k x_{t-k} + e_t, \quad (3)$$

Dalsze przekształcenia polegające na odejmowaniu i dodawaniu $(A_1 - I) x_{t-2}$, $(A_2 + A_1 - I) x_{t-3}$, ... , prowadzą do modelu VECM w postaci [Johansen 1995]:

$$\Delta x_t = \Psi_0 D_t + \Pi x_{t-1} + \sum_{i=1}^{k-1} \Pi_i \Delta x_{t-i} + \epsilon_t, \quad (4)$$

gdzie: $\Pi = \sum_{i=1}^k A_i - I$, $\Pi_i = - \sum_{j=i+1}^k A_j$.

Johansen dowiódł, że bazując na równaniu (4) do identyfikacji kointegracji można wykorzystać rząd macierzy Π ponieważ jest on równy liczbie niezależnych wektorów kointegrujących. Jeżeli zmienne nie są skointegrowane to rząd macierzy Π wyniesie zero.

Analizę przyczynowości dla relacji eksport produktów żywnościowych – eksport ogółem krajów UE wykonano na podstawie danych panelowych. Baltagi [2008] rozróżnia mikro-panele (najczęściej krótkie szeregi składające się z wielu obiektów) oraz makro-panele (dane o mniejszej liczbie obiektów i dłuższym horyzoncie czasowym). Rozróżnienie to ma uzasadnienie z uwagi na odmienne podejścia metodologiczne do modelowania obu rodzajów danych. Dynamiczny model panelowy można przedstawić jako wyrażenie:

$$y_{it} = \gamma_{i,t-1} + x_{it}^T \beta + (a_i + \epsilon_{it}), \quad i = 1, \dots, N, \quad t = 1, \dots, T, \quad (5)$$

gdzie:

x_{it} – wektor zmiennych objaśniających o k współrzędnych – $x_{it} = [x_{kit}]_{K \times 1}$,

β – wektor parametrów (jednakowych dla każdego i oraz t),

γ – parametr strukturalny,

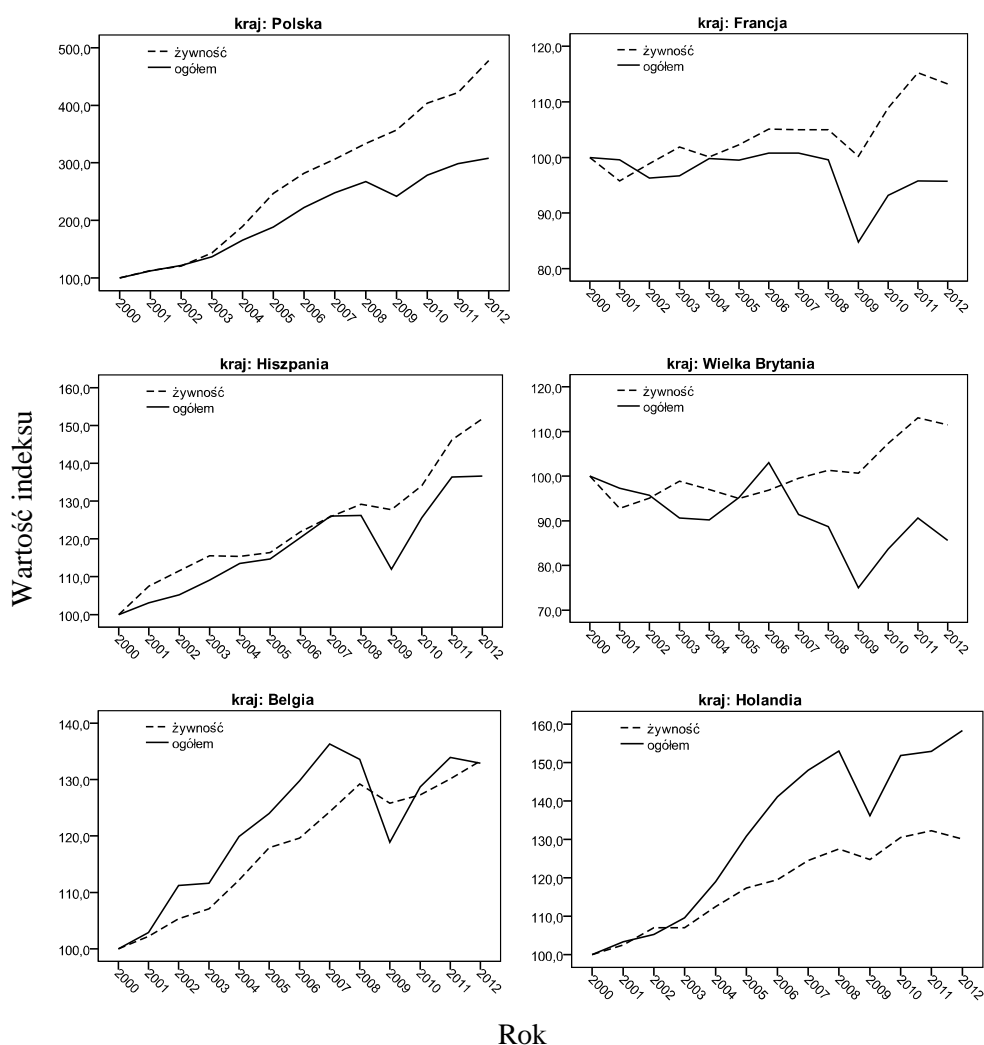
a_i – efekt grupowy,

ϵ_{it} – składnik losowy – $\epsilon_{it} \sim N(0, \delta_\alpha^2)$,

EKSPORT OGÓŁEM A EKSPORT ROLNY KRAJÓW UE

W zbiorze krajów UE współzależności między eksportem ogółem i eksportem żywności przyjmują różnoraki charakter oraz podlegają pewnej ewolucji w czasie. Kształty tych relacji w oparciu o indeksy obu zmiennych dla wybranych krajów (reprezentujących typy występujących zależności) przedstawia Rysunek 1.

Rysunek 1. Relacja między indeksami eksportu ogółem i eksportu produktów żywnościowych dla wybranych krajów UE (2000 = 100)



Źródło: obliczenia własne

Tabela 1. Badanie stacjonarności zmiennych eksport produktów żywnościowych oraz eksport ogółem – test KPSS*

Lp	Kraj	Liczba obserwacji	Eksport produktów żywnościowych		Eksport ogółem	
			Statystyka testu	p^{**}	Statystyka testu	p^{**}
1	Austria	T = 19	0,100230	$p > 0,10$	0,112142	$p > 0,10$
2	Belgia	T = 19	0,154483	$p < 0,05$	0,070970	$p > 0,10$
3	Bułgaria	T = 14	0,202848	$p < 0,01$	0,092814	$p > 0,10$
4	Cypr	T = 14	0,113084	$p > 0,10$	0,077383	$p > 0,10$
5	Czechy	T = 14	0,112057	$p > 0,10$	0,061771	$p > 0,10$
6	Dania	T = 19	0,042960	$p > 0,10$	0,109776	$p > 0,10$
7	Estonia	T = 14	0,122042	$p > 0,10$	0,071975	$p > 0,10$
8	Finlandia	T = 19	0,178509	$p < 0,05$	0,146143	$p > 0,05$
9	Francja	T = 19	0,172865	$p < 0,05$	0,168281	$p < 0,05$
10	Grecja	T = 19	0,199818	$p < 0,05$	0,181318	$p < 0,05$
11	Hiszpania	T = 19	0,129147	$p > 0,05$	0,040560	$p > 0,10$
12	Holandia	T = 19	0,209395	$p > 0,10$	0,090826	$p > 0,10$
13	Irlandia	T = 19	0,054758	$p > 0,10$	0,230666	$p < 0,01$
14	Litwa	T = 14	0,158507	$p < 0,05$	0,104398	$p > 0,10$
15	Luksemburg	T = 14	0,122034	$p > 0,10$	0,178773	$p < 0,05$
16	Łotwa	T = 14	0,174836	$p < 0,05$	0,129859	$p > 0,05$
17	Malta	T = 14	0,078037	$p > 0,10$	0,145035	$p > 0,05$
18	Niemcy	T = 19	0,237602	$p < 0,01$	0,045007	$p > 0,10$
19	Polska	T = 19	0,161649	$p < 0,05$	0,073882	$p > 0,10$
20	Portugalia	T = 19	0,239502	$p < 0,01$	0,040052	$p > 0,10$
21	Rumunia	T = 14	0,204697	$p < 0,01$	0,107269	$p > 0,10$
22	Słowacja	T = 14	0,119629	$p > 0,10$	0,095989	$p > 0,10$
23	Słowenia	T = 14	0,160058	$p < 0,05$	0,072283	$p > 0,10$
24	Szwecja	T = 19	0,234729	$p < 0,01$	0,048912	$p > 0,10$
25	Węgry	T = 14	0,159786	$p < 0,05$	0,071154	$p > 0,10$
26	Wielka Brytania	T = 19	0,124849	$p = 0,10$	0,118516	$p > 0,10$
27	Włochy	T = 19	0,202862	$p < 0,05$	0,055544	$p > 0,10$

*test KPSS dla zmiennej (z trendem), parametr rzędu opóźnienia = 1

**test KPSS - hipoteza zerowa: proces stacjonarny

Źródło: obliczenia własne

W Tabeli 2 zaprezentowano analogiczny charakter informacji, jak na Rysunku 1 – jednak uwzględniono w niej relacje we wszystkich badanych krajach. Opierając się na relacji indeksów (rok bazowy 2000 = 100) eksport ogółem/eksport rolny można wyróżnić trzy ogólne podgrupy krajów UE ze względu na kształtowanie się zjawiska: kraje z wyraźnie szybszym wzrostem eksportu rolnego, kraje o zbliżonych tempach wzrostu obu procesów oraz państwa, w których znacząco szybciej wzrastał eksport ogółem.

Tabela 2. Relacja indeksów eksport ogółem / eksport rolny (rok 2012)

Relacja eksport ogółem / eksport rolny	Kraj
Wyraźnie wyższy indeks eksportu rolnego	Austria, Bułgaria, Czechy, Estonia, Finlandia, Francja, Hiszpania, Irlandia, Litwa, Niemcy, Polska, Portugalia, Słowacja, Słowenia, Szwecja, Wielka Brytania, Włochy
Zbliżone wartości obu indeksów	Belgia, Dania
Wyraźnie wyższy indeks eksportu ogółem	Cypr, Grecja, Holandia, Luksemburg, Węgry

Źródło: obliczenia własne

Baza danych nie udostępniła informacji dla: Łotwy, Malty, Rumunii

W teście KPSS hipoteza zerowa zakłada stacjonarność szeregu czasowego. W przeprowadzonym badaniu wartość statystyki testu KPSS w większości przypadków przewyższa wartości krytyczne dla powszechnie przyjmowanych poziomów istotności (Tabela 1). Nakazuje to odrzucanie hipotezy zerowej o stacjonarności w poszczególnych przypadkach, co sugeruje niestacjonarność zmiennych [Welfe 2009]. Podejście ostrożnościowe nakazuje potraktowanie wszystkich zmiennych jako niestacjonarnych.

Tabela 3. Badanie kointegracji zmiennych eksport produktów żywnościowych oraz eksport ogółem – test Johannesa

Lp	Kraj	Rząd macierzy	Wartość własna	Test kointegracji			
				λ_{trace}	p	λ_{max}	p
1	Austria	0	0,83169	38,5530	0,0000	32,0760	0,0000
		1	0,30224	6,4779	0,1619	6,4779	0,1617
2	Belgia	0	0,86212	39,5300	0,0003	35,6650	0,0000
		1	0,19322	3,8647	0,7592	3,8647	0,7609
3	Bułgaria	0	0,86189	32,9980	0,0044	23,7570	0,0086
		1	0,53704	9,2413	0,1699	9,2413	0,1698
4	Cypr	0	0,84169	21,8550	0,0143	20,2750	0,0149
		1	0,13374	1,5792	0,2089	1,5792	0,2089
5	Czechy	0	0,74742	25,2160	0,0583	17,8880	0,0803
		1	0,43088	7,3277	0,3210	7,3277	0,3214
6	Dania	0	0,65526	18,1630	0,0523	17,0390	0,0493
		1	0,06782	1,1237	0,2891	1,1237	0,2891
7	Estonia	0	0,91054	34,2250	0,0028	28,9680	0,0009
		1	0,35472	5,2569	0,5682	5,2569	0,5696
8	Finlandia	0	0,64442	18,0520	0,0542	17,5780	0,0407
		1	0,02749	0,4739	0,4912	0,4739	0,4912
9	Francja	0	0,97225	50,6910	0,0000	50,1840	0,0000
		1	0,03562	0,5077	0,4761	0,5077	0,4761

Lp	Kraj	Rząd macierzy	Wartość własna	Test kointegracji			
				λ_{trace}	p	λ_{max}	p
10	Grecja	0	0,72852	24,305	0,0760	20,8620	0,0273
		1	0,19360	3,4427	0,8135	3,4427	0,8152
11	Hiszpania	0	0,64616	24,121	0,0124	18,7000	0,0153
		1	0,26003	5,4207	0,2492	5,4207	0,2487
12	Holandia	0	0,60222	19,1810	0,0027	16,593	0,0043
		1	0,13393	2,5882	0,1259	2,5882	0,1275
13	Irlandia	0	0,78237	35,0690	0,0020	25,9240	0,0034
		1	0,41603	9,1443	0,1759	9,1443	0,1758
14	Litwa	0	0,72346	17,6110	0,0053	16,7100	0,0041
		1	0,06697	0,9012	0,3977	0,9012	0,3936
15	Luksemburg	0	0,96359	37,0380	0,0000	36,4430	0,0000
		1	0,05263	0,5947	0,4406	0,5947	0,4406
16	Łotwa	0	0,63618	13,1670	0,0351	13,1440	0,0211
		1	0,00178	0,0232	0,9243	0,0232	0,9179
17	Malta	0	0,81910	18,8100	0,0138	18,8080	0,0075
		1	0,00016	0,0017	0,9668	0,0017	0,9668
18	Niemcy	0	0,74219	24,932	0,0002	24,4000	0,0001
		1	0,02913	0,5322	0,5329	0,5322	0,5259
19	Polska	0	0,84911	24,5990	0,0002	24,5860	0,0001
		1	0,00096	0,0125	0,9482	0,0125	0,9430
20	Portugalia	0	0,71378	22,5210	0,0006	22,5180	0,0002
		1	0,00017	0,0030	0,9785	0,0030	0,9755
21	Rumunia	0	0,70670	16,1550	0,0381	15,9450	0,0248
		1	0,01602	0,2010	0,6468	0,2010	0,6468
22	Słowacja	0	0,76500	17,856	0,0200	17,3780	0,0138
		1	0,03904	0,4779	0,4894	0,4779	0,4894
23	Słowenia	0	0,63117	12,9670	0,0380	12,9670	0,0229
		1	0,00005	0,0007	0,9911	0,0007	0,9896
24	Szwecja	0	0,74021	24,2640	0,0003	24,2620	0,0001
		1	0,00014	0,0025	0,9807	0,0025	0,9779
25	Węgry	0	0,92934	31,9270	0,0002	31,7990	0,0001
		1	0,01068	0,1288	0,7197	0,1288	0,7196
26	Wielka Brytania	0	0,62421	15,7190	0,0447	15,6600	0,0278
		1	0,00368	0,0589	0,8082	0,0589	0,8082
27	Włochy	0	0,66919	18,8270	0,0418	18,8060	0,0260
		1	0,00126	0,0214	0,8837	0,0214	0,8837

Źródło: obliczenia własne

Opierając się na wynikach testu kointegracji (Tabela 3), mimo różnego kształtu zależności, można potwierdzić hipotezę o współzależności przebiegu w czasie procesów: eksport produktów żywnościowych oraz eksport ogółem. Test

kointegracji odwołujący się do równania 4, wykorzystujący rząd macierzy Π do identyfikacji liczby wektorów kointegrujących, pozwala na odrzucenie w poszczególnych przypadkach hipotezy o niewystępowaniu wektorów kointegrujących (odrzuca się hipotezę, że rząd macierzy Π jest równy zero). Dla żadnego z obiektów nie ma natomiast podstaw do odrzucenia hipotezy, że macierz Π jest pierwszego rzędu, co należy interpretować jako występowanie dokładnie jednego wektora kointegrującego. Zatem, szeregi czasowe są skointegrowane, a czynnikiem korekty błędem jest wyrażenie Πx_{t-1} . W niektórych przypadkach prawdopodobieństwo błędu I rodzaju jest większe od 0,05, ale mniejsze od 0,10 dla testu, w którym hipoteza zerowa mówi o braku kointegracji.

Tabela 4. Relacja między eksportem produktów żywnościowych a eksportem ogółem krajów UE – dynamiczny model panelowy

zmienna	współczynnik	błąd standardowy	z	p
e_F(-1)	0,9455	0,0218	43,420	0,0000
stała	36,3268	14,7631	2,461	0,0139
e_T	0,0329	0,0057	5,797	0,0000
e_T(-1)	-0,0207	0,0069	-3,002	0,0027
e_T(-2)	-0,0051	0,0019	-2,769	0,0056

Źródło: obliczenia własne

Przedstawiony w Tabeli 4 model dynamiczny dla danych panelowych ma za zadanie próbę opisu związków między eksportem ogółem i eksportem żywności. Zmienną zależną w modelu jest wielkość eksportu żywności. Stosownie do wyników estymacji, w wyjaśnianiu badanych współzależności istotne są pierwsze opóźnienia zmiennej eksport żywności (e_F) oraz wartości od bieżących do drugich opóźnień zmiennej eksport ogółem (e_T).

PODSUMOWANIE

W opracowaniu oceniano długookresowe relacje między eksportem produktów żywnościowych a eksportem ogółem krajów UE. Opierając się na wynikach testu kointegracji metodą Johannaesa, mimo różnego kształtu zależności, potwierdzono hipotezę o współzależności szeregów czasowych rozważanych zmiennych w okresie badania.

Eksport towarów rolno-żywnościowych jest podkategorią eksportu ogółem, jednak rozwój wymiany międzynarodowej w tej grupie towarowej kształtują inne czynniki niż ma to miejsce w pozostałych grupach towarowych. Uzasadnia to testowanie wzajemnej relacji w czasie między przebiegiem poszczególnych procesów.

Analiza przyczynowości wskazała na istotność w wyjaśnianiu badanych procesów opóźnień zmiennej eksport żywności, jednak jeszcze trwalsze

oddziaływanie na kształt badanych związków wykazuje eksport ogółem reprezentujący kompleks otoczenia zewnętrznego.

BIBLIOGRAFIA

- Baltagi B. (2008) *Econometric analysis of panel data*, Wiley, Chichester.
- Box G.E.P., Jenkins G.M. (1970) *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.
- Cooper R. L. (1972) *The Predictive Performance of Quarterly Econometric Models of the United States*. [W:] B. G. Hickman (red.), *Econometric Models of Cyclical Behaviour*, Columbia University Press, New York.
- Engle R.F., Granger C.W.J., (1987) *Co-integration and error correction: representation, estimation and testing*, *Econometrica*, vol. 55, s. 251 – 276.
- Farmer K., Schelnast M. (2013), *Growth and International Trade: An Introduction to the Overlapping Generations Approach*, Springer Verlag, Heidelberg.
- Johansen S., (1995) *Likelihood-Based Inference in Cointegrated Vector Auto-regressive Models*, Oxford University Press, New York.
- Johnson H.G. (2013), *International Trade and Economic Growth (Collected Works of Harry Johnson) Studies in Pure Theory*, Routledge, New York.
- Johansen S. (1988) *Statistical Analysis of Cointegration Vectors*, *Journal of Economic Dynamics and Control*, vol. 12 , s. 231-254.
- Kusideł E. (2000) *Modele wektorowo-autoregresyjne VAR. Metodologia i zastosowania*, Absolwent, Łódź.
- Kwiatkowski D., Phillips P.C.B., Schmidt P., Shin Y. (1992) *Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root*, *Journal of Econometrics*, vol. 54, s. 159–178.
- Phillips P.C.B. (1986) *Understanding Spurious Regressions in Econometrics*, *Journal of Econometrics*, vol. 33, s. 311-340.
- Sims C.A. (1980) *Macroeconomics and Reality*, *Econometrica*, vol. 48. s. 1-48.
- Van den Berg H., Lewer J.J. (2007) *International Trade and Economic Growth*, M.E. Sharpe, New York.
- Welfe A. (2009) *Ekonometria. Metody i ich zastosowanie*, PWE, Warszawa.

CONNECTIONS BETWEEN AGRICULTURAL EXPORTS AND TOTAL EXPORTS OF EU COUNTRIES

Abstract: The paper attempts to investigate the relationship between the agro-food exports and total exports of EU countries. An effort was done to test the hypothesis of interdependence between these processes. In the sake of accessing the degree of the mutual interdependency between this variables the methods of time series analysis were employed.

Keywords: agricultural exports, total exports, cointegration analysis, dynamic panel model

PRZYCZYNOWOŚĆ W SENSIE GRANGERA – WYBRANE METODY

Ewa M. Syczewska

Instytut Ekonometrii, Kolegium Analiz Ekonomicznych
Szkola Główna Handlowa w Warszawie
e-mail: ewams@sgh.waw.pl

Streszczenie: Definicja przyczynowości w sensie Grangera stanowi kompromis umożliwiający statystyczne testowanie tej zależności. Liniowa przyczynowość to zawężenie tego pojęcia, w istocie chodziło o zależności procesów generujących dane. Istnieją testy dla zmiennych niestacjonarnych, testy przyczynowości nieliniowej itd. Omawiamy niektóre z nich. W części empirycznej stosujemy nieliniowy test Diksa-Panchenki do szeregów finansowych w celu sprawdzenia, czy wyniki różnią się od wyników testu Grangera.

Słowa kluczowe: przyczynowość w sensie Grangera, przyczynowość nieliniowa, test Hiemstry-Jonesa, test Diksa-Panchenki, kurs walutowy, wpływ kryzysu

WPROWADZENIE

Pojęcie przyczynowości w ekonomii jest szeroko dyskutowane z punktu widzenia filozofii i metod opisu aksjomatycznego oraz sposobów analizy (Pearl [2000] stosuje metody teorii grafów oparte na podstawach probabilistycznych w ujęciu bayesowskim; Hoover [2001] podkreśla konieczność analizy struktury przyczynowej i posługuje się narzędziami wielorównaniowych modeli ekonometrycznych; Cartwright [2001] sięga do podstaw filozoficznych i krytykuje podejście Pearla i Hoovera; zob. też dyskusję pomiędzy Pearlem, Steelem i Cartwright [2010] nt. jej książki i recenzję Hoovera [2003] monografii Pearla [2000]). Praca [Syczewska, Struzik 2014] zawiera przegląd wybranych metod testowania przyczynowości: oprócz metod stosowanych w ekonometrii, istnieją narzędzia związane z teorią informacji, a w szczególności skonstruowane na podstawie miar entropii. [Barnett, Barret, Seth 2009] oraz [Hlaváčková-

Schindler 2011] udowodnili równoważność metod tzw. transferu entropii i liniowego testu Grangera dla zmiennych o rozkładzie normalnym.

Przedmiotem tej pracy nie jest włączanie się w szeroką dyskusję nt. koncepcji przyczynowości (obszerną analizę liniowych i nieliniowych testów przyczynowości zawiera monografia Osińskiej [2008], krótszy przegląd – [Osińska 2011]), lecz praktyczne zastosowanie liniowego testu Grangera i testu Diksa i Panchenki do dziennych notowań bilateralnego kursu walutowego i odpowiednich indeksów giełdowych. Diks i Panchenko [2005] stwierdzili, że test Hiemstry-Jonesa [1994] (dalej HJ) może odrzucać hipotezę o braku przyczynowości nawet gdy liniowy test Grangera zdecydowanie na nią wskazuje; zaproponowali więc skorygowaną wersję. Test T_n [Diks, Panchenko 2006] można stosować do wykrywania zależności przyczynowych niekoniecznie liniowych.

Część empiryczna (punkt 2 i 3) poprzedzona jest omówieniem wybranych metod definiowania i testowania przyczynowości. We wcześniejszych pracach stwierdzono poprawę prognoz bilateralnych kursów walutowych (USDPLN oraz EURPLN) dzięki zastosowaniu indeksów giełdowych jako zmiennych objaśniających (co wynika z pracy [Bauwens, Rime, Succarat 2008]). W pracy [Syczewska 2013] wykazano, że istnieje stacjonarna kombinacja liniowa kursu USDPLN i dwu indeksów dla Polski i dla Stanów Zjednoczonych, a w [Syczewska 2010a,b] porównano zachowanie kursów w okresie kryzysu i poza nim. Podobny zestaw zmiennych badamy w niniejszym opracowaniu. W literaturze omawiany jest wpływ kryzysu lat 2008-09 na zmiany powiązań zwrotów z instrumentów finansowych, dlatego porównujemy wyniki dla kryzysu i w okresie późniejszym.

PRZYCZYNOWOŚĆ W SENSIE GRANGERA

Pojęcie przyczynowości w sensie Grangera nie ogranicza się do wyszukiwania zmiennej, której obecność w modelu może poprawić dokładność prognoz. Zwracał na to uwagę sam Granger, podkreślając, że w [Granger 1980] chodziło mu o to, że „dla szeregu czasowego, przyczyna poprzedza skutek a szereg będący przyczyną (*causal series*) zawiera informacje o zmiennej będącej skutkiem, które nie były zawarte w innych szeregach [...] G-przyczynowość dotyczy procesów generujących dane (DGP), a testy – modeli” (por. [Granger 2003]).

Liniowa wersja testu Grangera jest stosowana do wykrywania przyczynowości w średniej. Monika Krawiec [2012] analizuje notowania wybranych polskich funduszy towarowych oraz światowe ceny najważniejszych dóbr, dla okresu 01.2009–12.2011 testuje liniową przyczynowość Grangera stóp zwrotu w ujęciu logarytmicznym, $r_t = \ln P_t - \ln P_{t-1}$, i stwierdza występowanie powiązań przyczynowych od indeksu S&P500 do niektórych ale nie wszystkich funduszy surowcowych. [Misiuk, Zajkowska 2010] na podstawie danych dziennych od 01. 2004 do 03.2010 wykazały, że WIG20 wpływa na DAX, BUX, ATX i S&P500, i jest pod wpływem CAC, PX (indeks dla Czech), BUX (indeks węgierski), ATX (austriacki), FTSE, FTSE MIB (włoski). Zastosowania

w ekonometrii finansowej często dotyczą przyczynowości w wariancji [Cheung, Ng 1996] i ryzyku. [Fałdziński, Osińska, Zdanowicz 2012] przypominają, że pierwotna idea przyczynowości Grangera w ryzyku zakłada wykorzystanie Value at Risk jako miary ryzyka, i poszerzają zastosowania testu o Expected Shortfall i miary spektralne. Przyczynowość w ryzyku testują dla okresu przed kryzysem (1.02.2006-31.07.2008) oraz podczas i po kryzysie (1.08.2008-18.02.2011).

Nieliniowymi aspektami zależności przyczynowych zajmował się już Granger (zob. np. [Granger Teräsvirta 1993], [Granger Maassoumi Racine 2004]). [Bruzda 2004, 2007] omawia wprowadzone w tej ostatniej pracy warunki, które musi spełniać idealna miara zależności funkcyjnej dwu procesów stochastycznych: ma być dobrze zdefiniowana dla zmiennych ciągłych jak i dyskretnych; równa zeru dla zmiennych niezależnych; unormowana do $\langle 0,1 \rangle$ lub $\langle -1,1 \rangle$; ma spełniać warunki odległości, a dla zmiennych o rozkładzie normalnym być równa współczynnikowi korelacji lub pewnej jego prostej funkcji. Warunki te spełniają względna entropia (MI, *mutual information* czyli *relative entropy*) dwu zmiennych X, Y (określona wzorem $I(X, Y) = H(X) + H(Y) - H(X, Y)$ gdzie $H(X, Y)$ – entropia Shannona), a także współczynnik wzajemnej informacji (*mutual information coefficient*, MIC): $R(X, Y) = 1 - [-2I(X, Y)]^{0,5}$ (por. [Granger Teräsvirta 1993]). Mogą być stosowane np. do wykrywania nieliniowości i rozróżniania jej typu ([Bruzda 2004]), wykrywania zależności zmiennej i jej opóźnień ([Orzeszko 2012], [Orzeszko 2008]).

Interesuje nas tu kwestia wpływu kryzysu na ewentualne zmiany kierunku zależności przyczynowej. [Olbryś, Majewska 2014] badają sposób wyznaczania początku i końca kryzysu dla Stanów Zjednoczonych i wybranych krajów europejskich. Zwracają uwagę na brak zgodności wśród badaczy, np. [Mishkin 2011] dzieli ten okres na dwa: 08.2007– 08.2008 i kryzys globalny od połowy września 2008; inni autorzy wyróżniają trzy podokresy. Olbryś i Majewska na podstawie punktów zwrotnych szeregów czasowych głównych indeksów giełdowych badanych krajów stwierdzają, że 10.2007 – 02.2009 to wspólny dla nich okres globalnego kryzysu¹.

DANE I TESTY NIESTACJONARNOŚCI

Badanie empiryczne dotyczy dziennych notowań bilateralnego kursu USDPLN, i dwu indeksów: S&P500 i WIG20, od 04.01.2000 r. do 31.10.2014. Dla kursu walutowego i WIG20 są to notowania zamknięcia, dla S&P500 reprezentującego indeksy amerykańskie². – notowania otwarcia. Dni, dla których brakuje obserwacji jednej lub kilku zmiennych, zostały z próby usunięte. Źródłem danych jest portal stooq.pl.

¹ Dla WIG20 główny okres kryzysu to czas od czerwca 2007 do lutego 2009.

² Przed zamknięciem GPW otwarta jest giełda nowojorska, więc inwestorzy mogą śledzić tamtejsze kursy otwarcia.

Tabela 1. Wybrane statystyki dla całej próby 2000/01/04 - 2014/10/31

Zmienna	Średnia	Mediana	Odchylenie standardowe	Skośność	Nadwyżka kurtozy	Test ADF dla:	
						Zmiennej	przyrostów
WIG20	2202,3	2278,4	660,94	0,2492	-0,3799	-0,287 [0,58]	-14,292
USDPLN	3,3456	3,2033	0,5589	0,2270	-0,5819	-0,757 [0,39]	-12,144
S&P500	1272,4	1254,0	258,68	0,6279	0,4406	0,740 [0,87]	-13,0456

Źródło: stooq.pl. W nawiasach kwadratowych podano wartość p odpowiedniej statystyki.

Zastosowano najpierw test ADF w celu zbadania stopnia integracji, oraz testy kointegracji Johansena (kointegracja implikuje przyczynowość w sensie Grangera, ale nie odwrotnie). Wszystkie trzy szeregi są I(1). Procedura Engle'a-Grangera nie wykrywa kointegracji – reszty regresji USDPLN względem dwu pozostałych zmiennych są niestacjonarne (ADF = -0,813 [0,93]). Tabela 2 zawiera wyniki testu Johansena. Liczba wektorów kointegrujących, równa jest rzędowi r macierzy Π pomocniczego równania testu. Nie ma podstaw do odrzucenia hipotezy zerowej o braku wektorów kointegrujących: $r = 0$, przeciw hipotezie alternatywnej o występowaniu kointegracji ($r = 1$ lub $r \geq 1$): zarówno statystyka testu śladu L_{trace} jak i L_{max} mają wartości p bliskie 95%.

Tabela 2. Testy Johansena dla USDPLN, WIG20 i S&P500

Rząd	Wartość własna	Test śladu	Test Lmax
0	0,001404	7,6912 [0,9544]	5,0771 [0,9415]
1	0,000652	2,6141 [0,8798]	2,3584 [0,8691]
2	7,0739e-005	0,2557 [0,6856]	0,2557 [0,6764]

Źródło: obliczenia własne na podstawie danych stooq.pl.

W nawiasach kwadratowych podano wartość p dla odpowiedniej statystyki.

LINIOWY TEST GRANGERA

Test liniowy przyczynowości Grangera par zmiennych przeprowadzamy w następujący sposób. Szacujemy równania modelu VAR o jednakowej liczbie opóźnień dla obu zmiennych, k , i stosujemy test łącznej istotności opóźnień danej zmiennej w równaniu objaśniającym drugą zmienną:

$$y_t = \alpha_{10} + \sum_{j=1}^k \alpha_{1j} y_{t-j} + \sum_{j=1}^k \beta_{1j} x_{t-j} + \varepsilon_{1t} \quad (3a)$$

$$x_t = \alpha_{20} + \sum_{j=1}^k \alpha_{2j} x_{t-j} + \sum_{j=1}^k \beta_{2j} y_{t-j} + \varepsilon_{2t} \quad (3b)$$

W równaniu (3a), $H_0: \beta_{11} = \beta_{12} = \dots = \beta_{1k} = 0$ oznacza brak zależności przyczynowej w sensie Grangera zmiennej X do Y . W równaniu (3b), $H_0: \beta_{21} = \beta_{22} = \dots = \beta_{2k} = 0$ oznacza brak (liniowej) przyczynowości od Y do X .

Testowanie przyczynowości w sensie Grangera w modelu VAR z nadwyżką parametrów

[Toda, Yamamoto 1995] oraz [Bauer, Maynard 2012] stwierdzają, że w przypadku zmiennych niestacjonarnych wskazane jest wybranie optymalnej liczby k opóźnień modelu VAR, a następnie zwiększenie jej o liczbę równą rzędowi integracji zmiennych, d :

$$y_t = \alpha_{10} + \sum_{j=1}^{k+d} \alpha_{1j} y_{t-j} + \sum_{j=1}^{k+d} \beta_{1j} x_{t-j} + \varepsilon_{1t} \quad (4)$$

i testowanie $H_0: \beta_{11} = \beta_{12} = \dots = \beta_{1k} = 0$ dla pierwotnej liczby opóźnień. Toda i Yamamoto wykazali, że statystyka testu ma lepsze własności niż w wersji (3a-b), w szczególności graniczny rozkład odpowiada teoretycznemu. Zgodnie z tą zasadą, wybieramy liczbę opóźnień na podstawie kryterium Akaike (AIC), zwiększamy ją o 1 (czyli stopień integracji zmiennych). Model szacujemy dla zwrotów logarytmicznych (wg procedury opisanej m.in. przez Osińską [2008]) i stosujemy test łącznej istotności do pierwszych k opóźnień, odpowiadający hipotezie zerowej. Dodatkowo sprawdzamy normalność składników losowych testem Doornika-Hansena. Tabela 3 ilustruje wybór liczby opóźnień modelu VAR, dla całej próby, na podstawie kryterium Akaike, AIC, bayesowskiego kryterium Schwarz, BIC, oraz Hannana-Quinna, HQC. Ich wartości wyznaczono dla 10 opóźnień, ale optymalna liczba nie przekracza 4, więc przytaczamy tylko początkowe wartości.

Tabela 3. Wybór liczby opóźnień modelu VAR dla par zmiennych

Liczba opóźnień	USDPLN i WIG20			USDPLN i SP500			SP500 i WIG20		
	AIC	BIC	HQC	AIC	BIC	HQC	AIC	BIC	HQC
1	5,6753	5,6856*	5,6789	4,0137	4,0240	4,0174	18,0676	18,0779	18,0713
2	5,6706*	5,6877	5,6767*	3,9553	3,9724*	3,9614	17,9824	17,9996*	17,9885
3	5,6708	5,6948	5,6794	3,9493	3,9733	3,9579*	17,9782*	18,00232	17,9868*
4	5,6715	5,7023	5,6825	3,9481*	3,9790	3,9592	17,9789	18,0097	17,9899

Źródło: obliczenia własne w gretl

Zwroty USDPLN i WIG20, cała próba

Optymalna wg. AIC liczba opóźnień modelu VAR dla USDPLN i WIG20 jest równa $k=2$, stopień integracji $d = 1$, więc w regresji testu przyczynowości obie zmienne mają po $k+d=3$ opóźnienia i zgodnie z zasadą Tody i Yamamoto, test łącznej istotności stosujemy do $k = 2$ opóźnień zmiennej objaśniającej. Pełne wyniki estymacji podano w [Syczewska 2014]. Wnioski są następujące: WIG20 \rightarrow USDPLN: Łączny test istotności pierwszych dwu opóźnień WIG20 $F(2, 3608) = 0,6893$ [0,502], więc WIG20 nie jest przyczyną USDPLN w sensie Grangera. Odrzucamy H_0 o normalności reszt pierwszego równania: $\chi^2(2) = 1053,79$. USDPLN \rightarrow WIG20: Statystyka testu istotności 1–3 opóźnień USDPLN $F(2, 3608) = 14,618$ [0,000], więc hipotezę o braku przyczynowości odrzucamy. Reszty nie mają rozkładu normalnego: ($\chi^2(2) = 1037,12$ i $495,72$ [0,000]).

USDPLN i S&P500, cała próba

AIC wskazuje na 4 opóźnienia, więc przyjmujemy $k + 1 = 5$. Odrzucamy S&P500 \rightarrow USDPLN: $F(4, 3602) = 8,715$ [5,347e-007] oraz H_0 o normalności reszt: $\chi^2(2) = 1003,08$ [1,526e-218]. Odrzucamy USDPLN \rightarrow S&P500: $F(4, 3602) = 53,76$ [4,52e-044], a także H_0 o normalności reszt: $\chi^2(2) = 2310,84$ [0.0000].

S&P500 i WIG20, cała próba

Optymalna liczba opóźnień VAR dla zmiennych S&P500 i WIG20 jest równa 3, więc równanie zawiera $k + d = 3 + 1$ opóźnień każdej zmiennej, a testujemy istotność pierwszych trzech. Odrzucamy H_0 : WIG20 \rightarrow S&P500: $F(3, 3605) = 98,66$ [2,25e-61]. Nie odrzucamy hipotezy S&P500 \rightarrow WIG20: $F(3, 3605) = 2,675$ [0,046]. Reszty nie mają rozkładu normalnego: $\chi^2(2) = 2457,91$ [0,000] dla pierwszego równania oraz 487,97 [0,000] dla drugiego.

Testy liniowej przyczynowości dla okresu kryzysu i pokryzysowego

Zgodnie ze wskazaniem Olbryś i Majewskiej [2014], wybrano 1.10.2007 – 27.02.2009 jako okres kryzysu i dla porównania równoliczny (347 obserwacji) przedział po kryzysie (02.03.2009–26.07.2010). W Tabeli 4. kolumny 2. i 3. zawierają wartości statystyki testu Grangera, gwiazdka oznacza odrzucenie hipotezy o braku przyczynowości. Kolumny 4 i 5 to statystyki testu Doornika-Hansena dla reszt odpowiednich równań. Dla kryzysu test Grangera sugeruje występowanie dwustronnej przyczynowości obu indeksów i dla pary {USDPLN, S&P500}. Kurs USDPLN wpływa na WIG20, ale nie odwrotnie. W próbie po kryzysie, S&P500 nie jest już przyczyną w sensie Grangera dla WIG20 ani dla USDPLN. Wartość statystyki testu dla hipotezy WIG20 \rightarrow USDPLN jest na granicy odrzucenia (wartość p testu w wersji $\chi^2(2)$ jest nieco mniejsza niż 0,10). Rozkłady reszt nie są normalne.

Tabela 4. Test liniowej przyczynowości w sensie Grangera i test normalności reszt

Hipoteza zerowa	Test Grangera		Test Doornika-Hansena	
	Okres kryzysu	Po kryzysie	Okres kryzysu	Po kryzysie
SP500 \rightarrow WIG20	3,068 [0,028]*	1,24 [0,296]	17,833 [0,0001]	11,522 [0,003]
USDPLN \rightarrow WIG20	2,8397 [0,0598]*	6,816 [0,0013]*	14,976 [0,0006]	11,844 [0,0027]
WIG20 \rightarrow SP500	9,717 [3,62e-6]*	32,48 [1,80e-18]*	103,08 [4,13e-23]	53,38 [2,56e-12]
USDPLN \rightarrow SP500	4,324 [0,00201]*	40,65 [8,88e-28]*	56,14 [6,46e-13]	26,255 [0,0000]
SP500 \rightarrow USDPLN	8,556 [1,38e-6]*	0,226 [0,924]	57,43 [3,38e-13]	7,268 [0,0264]
WIG20 \rightarrow USDPLN	2,219 [0,1103]	2,316 [0,100]	84,036 [5,65e-19]	6,004 [0,050]

Źródło: obliczenia własne w pakiecie gretl.

W kolumnie 2 i 3 gwiazdką zaznaczono przypadki, w których hipoteza o braku przyczynowości została odrzucona. W nawiasach kwadratowych podano wartość p dla odpowiedniej statystyki.

TEST DIKSA-PANCHENKI

Do tych samych zmiennych, w próbie obejmującej okres kryzysu oraz okres po kryzysie, zastosowano nieparametryczny test T_n Diksa-Panchenki [2006]. Jest on łatwo dostępny w wersji dla Windows oraz Linuxa na stronie V. Panchenki³: http://research.economics.unsw.edu.au/vpanchenko/software/2006_GC_JEDC_c_and_exe_code.zip.

Idea testu Diksa-Panchenki jest następująca: Zmienna X jest przyczyną w sensie Grangera dla Y jeśli bieżąca i opóźnione wartości $X: x_t, x_{t-1}, \dots, x_{t-k}, \dots$ zawierają dodatkowe informacje o przyszłych wartościach $Y_{t+1}, Y_{t+2}, \dots, Y_{t+h}$, nie zawarte w bieżącej i opóźnionych wartościach tej zmiennej. Autorzy koncentrują się na przypadku, gdy badany jest wpływ X na Y_{t+1} . Testują warunkową niezależność, ale z przyczyn technicznych jedynie skończonej liczby opóźnień:

$$Y_{t+1} | (X_t, X_{t-1}, \dots, X_{t-l_X+1}, Y_t, Y_{t-1}, \dots, Y_{t-l_Y+1}) \sim Y_{t+1} | (X_t, X_{t-1}, \dots, X_{t-l_X+1})$$

gdzie l_X, l_Y oznaczają liczbę uwzględnionych opóźnień danej zmiennej. Test jest poprawioną wersją testu Hiemstry-Jonesa, porównującego różnice pomiędzy rozkładami warunkowymi⁴.

Diks i Panchenko stosują po jednym opóźnieniu dla obu zmiennych, i prognozę na jeden okres. Niech $(X_t, Y_t, Y_{t+1}) = (X, Y, Z)$ i niech X i Y będą zmiennymi ściśle stacjonarnymi. Hipoteza zerowa o braku przyczynowości oznacza, że rozkład warunkowy Z względem X i Y jest taki sam jak rozkład Z względem samego Y . Łączny rozkład i rozkłady brzegowe są związane wzorem: $f_{X,Y,Z}(x, y, z) / f_{X,Y}(x, y) = f_{Y,Z}(y, z) / f_Y(y)$. Diks i Panchenko [2006] wykazują wady przybliżenia różnicy obu stron tego wzoru całkami korelacyjnymi, które zastosowano w pracy Hiemstry-Jonesa, i wyprowadzają własną wersję wzoru: implikacją hipotezy zerowej jest

$$q_g \equiv E \left[\left(\frac{f_{X,Y,Z}(X,Y,Z)}{f_Y(Y)} - \frac{f_{X,Y}(X,Y) f_{Y,Z}(Y,Z)}{f_Y(Y)} \right) g(X, Y, Z) \right] = 0 \quad (5)$$

gdzie $g(X, Y, Z)$ jest dodatnią funkcją wagową; dla $g(X, Y, Z) = f_Y^2(Y)$ wzór upraszcza się do:

$$q_g \equiv E \left[\left(f_{X,Y,Z}(X, Y, Z) f_Y(Y) - f_{X,Y}(X, Y) f_{Y,Z}(Y, Z) \right) \right]$$

Przy założeniu prawdziwości hipotezy zerowej wyrażenie w nawiasach okrągłych zeruje się, więc (5) przyjmuje wartość zero. Hipotezę zerową odrzucamy jeśli obliczona wartość statystyki testu jest wysoka.

³ Dane dla każdej z dwu zmiennych zapisujemy w osobnych plikach tekstowych, *.txt, w pojedynczej kolumnie, bez nagłówka. Liczba obserwacji jest jednakowa, a pliki z danymi umieszczone w podkatalogu, w którym znajduje się plik wykonywalny .exe. Po wywołaniu programu należy podać nazwy plików, liczbę opóźnień dla obu zmiennych, oraz wielkość pasma (bandwith), domyślnie 0.5 (z kropką dziesiętną). Otrzymujemy dwie wartości statystyki testu, dla przyczynowości od X do Y oraz od Y do X .

⁴ Dokładniejszy opis konstrukcji testu można znaleźć w [Osińska 2008], str. 226–229.

Diks i Panchenko [2006] stosują dla (5) estymator oparty na funkcji wskaźnikowej:

$$T_n(\varepsilon) = \frac{(2\varepsilon)^{-d_X-d_Y-d_Z}}{n(n-1)(n-2)} \sum_i \left[\sum_{k,k \neq i} \sum_{j,j \neq 1} (I_{ik}^{X,Y,Z} I_{ij}^Y - I_{ik}^{XY} I_{ij}^{YZ}) \right]$$

gdzie n – liczba obserwacji, d_X – wymiar wektora X itd., $I(\cdot)$ to funkcja wskaźnikowa: $I_{i,j}^W = I(\|W_i - W_j\| \leq \varepsilon)$ równa 1 jeśli $\|W_i - W_j\| \leq \varepsilon$, 0 w przeciwnym przypadku, zaś $\|x\| = \sup_{i=1,2,\dots,d_X} |x_i|$ oznacza normę supremum.

Autorzy testu podkreślają konieczność odpowiedniego doboru parametru ε (*bandwith*), domyślnie równego 0,5, i podają wzory określające ten parametr w zależności od wielkości próby. Wykazują że dla $d_X = d_Y = d_Z = 1$ estymator jest zgodny i ma asymptotyczny rozkład normalny.

Wyniki testu dla badanych zmiennych

W naszych obliczeniach przyjmujemy domyślną wartość 0,5 a statystyki testu T_n wyznaczamy dla opóźnień od 1 do 5. Wyniki obliczeń testu Diksa-Panchenki wskazują na to, że statystyka może mieć różne wartości przy różnej liczbie opóźnień zmiennych X i Y . Jeśli jednak skupimy się na wynikach dla wartości $l_X, l_Y = 1$ i porównamy je z wynikami liniowego testu Grangera, możemy sformułować następujące wnioski. Wyniki obu testów – liniowego testu Grangera oraz testu Diksa-Panchenki – okazały się zbliżone. W okresie kryzysu hipoteza zerowa o braku przyczynowości była w większości przypadków odrzucana. W okresie po kryzysie zależności przyczynowe były słabsze. Różnica między wynikami obu testów wystąpiła dla kursu USDPLN: w okresie po kryzysie test Diksa-Panchenki nie odrzuca hipotezy o braku przyczynowości od USDPLN do indeksu WIG20 oraz do S&P500.

Tabela 5. Wyniki testu Diksa-Panchenki dla par zwrotów z instrumentów finansowych

	Okres kryzysu	Po kryzysie	Okres kryzysu	Po kryzysie
	USDPLN \rightarrow WIG20		WIG20 \rightarrow USDPLN	
1	1,949 [0,0257]*	1,096 [0,1366]	0,881 [0,1890]	0,147 [0,4416]
2	1,988 [0,0234]*	-0,380 [0,6479]	-0,123 [0,5489]	-0,592 [0,7232]
3	1,340 [0,0902]	-0,740 [0,7703]	0,522 [0,3009]	-0,940 [0,8264]
4	-0,168 [0,5666]	-0,057 [0,5226]	1,034 [0,1505]	-0,520 [0,6983]
5	-0,047 [0,5189]	-0,656 [0,7441]	-0,340 [0,6330]	-0,646 [0,7410]
	USDPLN \rightarrow SP500		SP500 \rightarrow USDPLN	
1	2,974 [0,0015]*	3,869 [0,00005]	2,496 [0,0063]*	0,530 [0,2981]
2	1,710 [0,0436]*	1,866 [0,0310]	1,008 [0,1568]	0,626 [0,2657]
3	1,262 [0,1035]	0,842 [0,2000]	0,079 [0,4686]	-0,691 [0,7551]

	Okres kryzysu	Po kryzysie	Okres kryzysu	Po kryzysie
4	0,007 [0,4972]	1,137 [0,1277]	0,395 [0,3465]	0,205 [0,4189]
5	-0,399 [0,6552]	-0,562 [0,7131]	0,346 [0,3648]	0,1277 [0,3435]
	SP500 → WIG20		WIG20 → SP500	
1	2,683 [0,0036]*	-0,059 [0,5236]	3,095 [0,00099]*	3,046 [0,0012]*
2	2,116 [0,0172]*	1,890 [0,0294]*	1,641 [0,0504]*	1,005 [0,1575]
3	1,073 [0,1417]	0,562 [0,2869]	1,273 [0,1015]	0,584 [0,2795]
4	-0,469 [0,6806]	0,330 [0,3707]	-0,286 [0,6124]	1,257 [0,1044]
5	-0,816 [0,7928]	-0,210 [0,5833]	0,478 [0,3163]	0,418 [0,3381]

Źródło: obliczenia własne przy użyciu programu GCTTest.exe V. Panchenki.
Gwiazdką zaznaczono przypadki, gdy hipoteza zerowa o braku przyczynowości zostaje odrzucona.

Tabela 6. Porównanie obu testów

Hipoteza zerowa o braku przyczynowości	Test liniowy Grangera		Test Diksa-Panchenki	
	W kryzysie	Po kryzysie	W kryzysie	Po kryzysie
SP500 → WIG20	Odrzuca	Nie odrzuca	Odrzuca	Nie odrzuca
USDPLN → WIG20	Odrzuca	Odrzuca	Odrzuca	#Nie odrzuca
WIG20 → SP500	Odrzuca	Odrzuca	Odrzuca	Odrzuca
USDPLN → SP500	Odrzuca	Odrzuca	Odrzuca	#Nie odrzuca
SP500 → USDPLN	Odrzuca	Nie odrzuca	Odrzuca	Nie odrzuca
WIG20 → USDPLN	Nie odrzuca	Nie odrzuca	Nie odrzuca	Nie odrzuca

Źródło: opracowanie własne na podstawie Tabeli 4 i 5.

Krzyżykiem (#) zaznaczono przypadki gdy występowała niezgodność między wskazaniami obu testów.

PODSUMOWANIE I WNIOSKI

Przedstawione wyniki testowania przyczynowości nawiązują do wcześniejszych badań nt. wykorzystania indeksów giełdowych jako dodatkowych zmiennych poprawiających jakość modeli i prognoz dziennych notowań kursów walutowych. Tym razem porównujemy ewentualne zależności przyczynowe pomiędzy kursem USDPLN a odpowiadającymi mu indeksami giełdowymi, które reprezentowane są przez S&P500 dla giełdy w Stanach Zjednoczonych i WIG20 dla giełdy warszawskiej. Stosujemy klasyczny już liniowy test Grangera oraz nowszy test Diksa-Panchenki oparty na analizie warunkowych rozkładów, wykrywający zależności niekoniecznie liniowe. Badamy również możliwość zmiany kierunku i siły powiązań przyczynowych w okresie kryzysu i poza nim.

Oba testy dają dość zbliżone wyniki, a powiązania przyczynowe poza kryzysem wydają się nieco słabsze.

Przypomnijmy, że Diks i Panchenko jako jeden z argumentów uzasadniających konieczność modyfikacji testu Hiemstry-Jonesa podali fakt, że test HJ częściej odrzucał hipotezę o braku przyczynowości niż test liniowy Grangera, co ich zdaniem osłabia wyniki wnioskowania na podstawie testu HJ w opublikowanych praktycznych zastosowaniach.

W nawiązaniu do tego porównania należy podkreślić, że w przedstawionej tu analizie test T_n Diksa i Panchenki w większości przypadków dał ten sam wynik co test Grangera. W dwu przypadkach test T_n nie odrzucił hipotezy o braku przyczynowości, podczas gdy test Grangera sugerował jej odrzucenie. Nie zdarzyło się zaś, aby test T_n odrzucił hipotezę zerową w przypadku gdy test liniowy Grangera jej nie odrzucał – w tym sensie modyfikacja dokonana przez Diksa i Panchenkę wydaje się spełniać swoje zadanie.

Wskazania testu Diksa-Panchenki zależą od liczby opóźnień zmiennych. Należałoby m.in. sprawdzić dokładniej wpływ opóźnień przyjętych dla zmiennych w teście Diksa-Panchenki, oraz porównać otrzymane tu rezultaty z innymi metodami wykrywania zależności przyczynowych. Przeprowadzone tu obliczenia stanowią podstawę do tego typu dalszych badań.

BIBLIOGRAFIA

- Barnett L., Barrett A., Seth, A. (2009) Granger causality and transfer entropy are equivalent for Gaussian variables”, *Phys. Rev. Letter* 103, 238701.
- Bauer D., Maynard A. (2012) Persistence-robust surplus-lag Granger causality testing, *Journal of Econometrics* 169, 293-300.
- Bauwens L., Pohlmeier W., Veredas, D. (2008) *High Frequency Financial Econometrics, Recent Developments*, Physica-Verlag A Springer Company, Heidelberg.
- Bauwens L., Rime D., Succarat G. (2008) Exchange Rate Volatility and the Mixture of Distribution Hypothesis, [w:] Bauwens et al., 7–29,
- Bruzda J. (2004) Miary zależności nieliniowej w identyfikacji nieliniowych procesów ekonomicznych, *Acta Universitatis Nicolai Copernici*, 34, 183–203.
- Bruzda J. (2007) *Procesy nieliniowe i zależności długookresowe w ekonomii. Analiza kointegracji nieliniowej*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń 2007.
- Cartwright N. (2000) *Hunting causes and using them: Approaches in Philosophy and Economics*, Cambridge University Press, Cambridge.
- Cartwright N. (2010) Reply to Steel and Pearl, *Economics and Philosophy*, 26, 87–94.
- Cheung Y.W., Ng L.K. (1996) A causality in variance test and its application to financial market prices, *Journal of Econometrics*, 72, 33–48.
- Diks C., Panchenko V. (2005) A note on Hiemstra-Jones test for Granger non-causality, *Studies in Nonlinear Dynamics and Econometrics*, 9 (art. 4).
- Diks C., Panchenko V. (2006) A new statistics and practical guidelines for nonparametric Granger causality testing, *Journal of Economic Dynamics and Control*, 30, 1647-1669.

- Doornik J.A., Hansen H. (2008) An Omnibus test for univariate and multivariate normality, *Oxford Bulletin of Economics and Statistics*, 70, 927-939.
- Fałdziński M., Osińska M., Zdanowicz T. (2012) Detecting risk transfer in financial markets using different risk measures, *Central European Journal of Economic Modelling and Econometrics*, 4, 45-64.
- Granger C.W.J., Maasoumi E., Racine J. (2004) A Dependence Metric for Possibly Nonlinear Processes, *Journal of Time Series Analysis*, 25, 649-669.
- Granger C.W.J., Teräsvirta, T. (1993) *Modelling Nonlinear Economic Relationships*, Oxford University Press, Oksford.
- Hiemstra C., Jones J.D. (1994) Testing for linear and nonlinear Granger causality in the stock price volume relation, *Journal of Finance*, 49(5) 1639-1664.
- Hlaváčková-Schindler K. (2011) Equivalence of Granger causality and transfer entropy: a generalization, *Applied Mathematical Sciences* 5, 3637-3648.
- Hlaváčková-Schindler K., Paluš M., Vejmelka M., Bhattacharya, J. (2007) Causality detection based on information-theoretic approaches in time series analysis, *Physics Reports*, 441, 1-46.
- Hoover K. D. (2001) *Causality in Macroeconomics*, Cambridge University Press, Cambridge.
- Hoover K.D. (2003) Causality: Models, Reasoning, and Inference by Judea Pearl. Review by Hoover K.D., *The Economic Journal*, 113 (488) F411-F413.
- Krawiec M. (2012) Testing the Granger causality for commodity mutual funds in Poland and commodity prices, *Quantitative Methods in Econometrics*, Vol. XIII, No. 2., 84-95.
- Mishkin F.S. (2011) Over the Cliff: From the Subprime to the Global Financial Crisis, *Journal of Economic Perspectives*, 49-70.
- Misiuk A.B., Zajkowska O. (2010) Does simultaneous investing on different stock markets allow to diversify risk? The cointegration analysis with main focus on Warsaw Stock Exchange, *Quantitative Methods of Economics*, XI, 118-127.
- Olbryś J., Majewska E. (2014) Direct identification of crisis period on the CEE stock markets: The Influence of the 2007 U.S. subprime crisis, *Procedia Economics and Finance*, Special Issue: International Conference on Applied Economics, ICOAE2014, Vol. 14, 461-470.
- Orzeszko W. (2012) Nieliniowa identyfikacja rzędu zależności w stopach zmian indeksów giełdowych, *Przegląd Statystyczny*, 59, 369-385.
- Osińska M. (2008) *Ekonometryczna analiza zależności przyczynowych*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń.
- Osińska M. (2011) On the interpretation of causality in Granger sense, *Dynamic Econometric Models*, 129-139.
- Pearl J. (2000) *Causality: Models, Reasoning and Inference*, Cambridge University Press, Cambridge.
- Pearl J. (2010) Nancy Cartwright on Hunting Causes, *Economics and Philosophy*, 26, 69-77.
- Steel D. (2010) Cartwright on Causality: Methods, Metaphysics and Modularity, *Economics and Philosophy*, 26, 77-86.
- Syczewska E.M. (2010a) Increase of exchange rate risk during current crisis, *Roczniki Kolegium Analiz Ekonomicznych*, 21/2010, 99-122.

- Syczewska E.M. (2010b) Changes of exchange rate behaviour during and after crisis, 145-157 w: *Metody Ilościowe w Badaniach Ekonomicznych, (Quantitative Methods in Economics) XI*, 2010.
- Syczewska E.M. (2013) On exchange-rate model with stock indices as additional regressors (during and after crisis), referat na konferencji międzynarodowej „CEST’2013, Current Economic and Social Topic International Colloquium, focused on Gender Disparities and Financial Market Analysis”, 23-24 maja 2013, Warszawa.
- Syczewska E.M. (2014) Przegląd metod analizy współzależności indeksów giełdowych i kursów walutowych oraz badania przyczynowości dla tych szeregów, opracowanie w ramach badania statutowego pt. „Makro- i mikroekonometria – zastosowania ekonomiczne i finansowe”, nr badania KAE/S/07/13-14, kierownik M. Gruszczyński, Szkoła Główna Handlowa, Warszawa.
- Syczewska E.M., Struzik Z.R. (2014) Granger causality and transfer entropy for financial returns, referat na 7 Sympozjum FENS 2014, 14-16 May 2014, Lublin.
- Toda, H.Y., Yamamoto, T. (1995) Statistical inference in vector autoregressions with possibly integrated processes, *Journal of Econometrics* 66, 225-250.
- Software:
- Panchenko, V. Kody do wyznaczania statystyki testu [Diks, Panchenko 2006] dla Linuxa i Windows:
http://research.economics.unsw.edu.au/vpanchenko/software/2006_GC_JEDC_c_and_exe_code.zip [dostęp 17 grudnia 2014 r.]

GRANGER CAUSALITY – SELECTED METHODS

Abstract: Granger causality is defined to allow for construction of statistical tests. Granger himself intended to describe, possibly nonlinear, dependencies between data generating processes and take into account informational content of interdependence between variables. There are tests for nonstationary variables, of nonlinear causality etc., we describe some of the methods. We apply to daily returns of exchange rate and stock indices the linear Granger test and the Diks-Panchenko test (for nonlinear dependencies), while taking into account influence of the recent crisis.

Keywords: Granger causality, linear causality test, Hiemstra-Jones test, Diks.-Panchenko test, nonlinear causality, exchange rate

**ZASTOSOWANIE WYBRANYCH MODELI ADAPTACYJNYCH
W PROGNOZOWANIU BRAKUJĄCYCH DANYCH
W SZEREGACH ZE ZŁOŻONĄ SEZONOWOŚCIĄ
DLA LUK NIESYSTEMATYCZNYCH**

Maria Szmuksta- Zawadzka

Studium Matematyki

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

e-mail: Maria.Szmuksta-Zawadzka@zut.edu.pl

Jan Zawadzki

Katedra Zastosowań Matematyki w Ekonomii

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

e-mail: Jan.Zawadzki@zut.edu.pl

Streszczenie: Artykuł poświęcony jest wykorzystaniu wybranych modeli wyrównywania wykładniczego: Browna, Holta i Holta-Wintersa w prognozowaniu zmiennych ze złożoną sezonowością w warunkach braku pełnej informacji. Prognozy wyjściowe będą budowane na podstawie szeregów oczyszczonych z sezonowości. Prognozy końcowe, uwzględniające wahania sezonowe, będą sumami prognoz wyjściowych i składników sezonowości lub iloczynami prognoz tego rodzaju i wskaźników sezonowości. Rozważania o charakterze teoretycznym zostaną zilustrowane przykładem empirycznym.

Słowa kluczowe: złożona sezonowość, wyrównywanie wykładnicze, prognozowanie, brakujące dane

WSTĘP

W literaturze statystyczno-ekonometrycznej można spotkać wiele przykładów zastosowania modeli adaptacyjnych do modelowania i prognozowania zjawisk, w których występuje jeden rodzaj wahań (np.: miesięczne, dekadowe). Dla danych nieczyszczonych (z sezonowością) najczęściej wykorzystywane były modele Holta-Wintersa (addytywny i multiplikatywny). Natomiast dla danych oczyszczonych z sezonowości metody wyrównywania wykładniczego: Browna

(prosty, liniowy i kwadratowy) oraz liniowy model Holta a także metody numeryczne. Prognozy ostateczne, w zależności od sposobu eliminacji wahań, otrzymuje się po przemnożeniu przez wskaźniki lub przez dodanie składników sezonowości¹.

Prognozy dla danych oczyszczonych mogą być także budowane jako sumy lub iloczyn wartości trendów szacowanych KMNK i odpowiednio składników lub wskaźników sezonowości. Tego rodzaju postępowanie określane jest mianem metody wskaźnikowej lub metody wskaźników sezonowości (por. [Dittmann 2006, s.85], Zeliaś i inni 2003, s.90]).

W pracy [Szmuksta-Zawadzka, Zawadzki; 2014] zaproponowana została procedura wykorzystania modeli adaptacyjnych w prognozowaniu zmiennych ze złożoną sezonowością dla pełnych danych. Niniejsza praca stanowi rozszerzenie rozważań na przypadek występowania niesystematycznych luk w danych.

WPROWADZENIE TEORETYCZNE

W niniejszej pracy podjęta zostanie próba wykorzystania modeli adaptacyjnych do modelowania i prognozowania zmiennych ze złożoną sezonowością dla danych oczyszczonych z jednego lub dwóch rodzajów wahań sezonowych. Zakładamy, że w szeregu czasowym dla danych dziennych występują wahania o cyklu tygodniowym (7 dniowym) i rocznym (12-miesięcznym). Mogą się one nakładać na siebie i na trend w sposób addytywny lub multiplikatywny.

Ogólny zapis modelu addytywnego jest następujący:

$$Y_{t(a)} = P^{(a)}(t) + M^{(a)}(t) + D^{(a)}(t) + V_{t(a)} \quad (1)$$

gdzie:

$P^{(a)}(t)$ – trend,

$M^{(a)}(t)$ – składniki sezonowości o cyklu 12 miesięcznym,

$D^{(a)}(t)$ – składniki sezonowości o cyklu 7 dniowym.

Natomiast postać ogólna modelu multiplikatywnego wyraża się wzorem:

$$Y_{t(m)} = P^{(m)}(t) \cdot M^{(m)}(t) \cdot D^{(m)}(t) \cdot V_{t(m)} \quad (2)$$

gdzie:

$P^{(m)}(t)$ – trend,

$M^{(m)}(t)$ - wskaźniki sezonowości o cyklu 12 miesięcznym,

$D^{(m)}(t)$ - wskaźniki sezonowości o cyklu 7 dniowym.

¹ Przegląd publikacji poświęconych metodom prognozowania, zarówno pełnych jak i brakujących danych oraz ich praktycznym zastosowaniom w odniesieniu do zmiennych z niezłożonymi wahaniami sezonowymi (miesięcznymi, kwartalnymi i dekadowymi), można znaleźć w pracy [Szmuksta-Zawadzka, Zawadzki; 2012].

Bezpośrednie wykorzystanie modeli Holta-Wintersa nie jest możliwe, ponieważ wymagałoby wprowadzenia dodatkowego, czwartego równania opisującego wahania o cyklu rocznym – uwzględniającego różną długość miesięcy.

Jak się wydaje, ze względów praktycznych mogą wchodzić w grę m.in. modele Holta-Wintersa dla danych oczyszczonych, z których wyeliminowano wahania o cyklu rocznym (Y_t^*). Jeżeli eliminacji dokonano odejmując składniki sezonowości to będzie to model w postaci addytywnej. W przypadku podzielenia wartości zmiennej prognozowanej przez wskaźniki sezonowości będziemy mieć do czynienia z postacią multiplikatywną.

Zapis modelu addytywnego Holta-Wintersa (A_HW) jest następujący [Pawłowski 1973]:

$$m_t^* = \alpha(Y_t^* - c_{t-L}) + (1 - \alpha)m_{t-1}^* \quad (3)$$

$$\delta_{1t}^* = \beta(m_t^* - m_{t-1}^*) + (1 - \beta)\delta_{1t-1}^* \quad (4)$$

$$C_t^* = \gamma(Y_t^* - m_t^*) + (1 - \gamma)C_{t-m}^* \quad (5)$$

$$0 \leq \alpha, \beta, \gamma \leq 1 \quad (6)$$

Predyktor oparty na tym modelu przyjmuje postać:

$$\Pi_{A_HW}^* = m_{t_0}^* + \delta_{1t_0}^* h + C_{t_0, 1+h}^* \quad (7)$$

Prognozę ostateczną, uwzględniającą wahania sezonowe, otrzymuje się na podstawie predyktora o postaci:

$$\Pi_{A_HW} = \Pi_{A_HW}^* + M^{(a)}(t) \quad (8)$$

Model multiplikatywny Holta-Wintersa (M_HW) można zapisać następująco:

$$m_t^* = \frac{\alpha Y_t^*}{C_{t-m}^*} + (1 - \alpha)(m_{t-1}^* + \delta_{1t-1}^*) \quad (9)$$

$$\delta_{1t}^* = \beta(m_t^* - m_{t-1}^*) + (1 - \beta)\delta_{1t-1}^* \quad (10)$$

$$C_t^* = \frac{\gamma Y_t^*}{m_t^*} + (1 - \gamma)C_{t-m}^* \quad (11)$$

$$0 \leq \alpha, \beta, \gamma \leq 1 \quad (12)$$

Predyktory; wyjściowy i końcowy, wyrażają się wzorami:

$$\Pi_{M_HW}^* = (m_{t_0}^* + \delta_{1t_0}^* h) C_{t_0-m+h}^* \quad (13)$$

$$\Pi_{M_HW} = \Pi_{M_HW}^* \cdot M^{(m)}(t) \quad (14)$$

Natomiast w przypadku danych, z których wyeliminowano dodatkowo także wahania o cyklu tygodniowym (Y_t^{**}) do budowy prognoz mogą być wykorzystane na przykład proste modele Browna i modele Holta.

Równanie addytywnego prostego modelu Browna (A_BS) jest następujące:

$$m_t^{**} = \alpha Y_t^{**} + (1 - \alpha) m_{t-1}^{**} \quad (15)$$

$$0 \leq \alpha \leq 1 \quad (16)$$

Predyktory wyjściowy i końcowy przyjmują postać:

$$\Pi_{A_BS}^{**} = m_{t_0}^{**} \quad (17)$$

$$\Pi_{A_BS} = \Pi_{A_BS}^{**} + M^{(a)}(t) + D^{(a)}(t) \quad (18)$$

W modelu multiplikatywnym równanie modelu prostego Browna (M_BS) różni się od postaci addytywnej jedynie sposobem wyznaczenia wartości oczyszczonych (Y_t^{**}) - są one ilorazami wartości zmiennej prognozowanej i wskaźników sezonowości o cyklu rocznym i tygodniowym.

Predyktory wyjściowy i końcowy są następujące:

$$\Pi_{M_BS}^{**} = m_{t_0}^{**} \quad (19)$$

$$\Pi_{M_BS} = \Pi_{M_BS}^{**} * M^{(m)}(t) * D^{(m)}(t) \quad (20)$$

Addytywny model liniowy Holta (A_H) można zapisać [Pawłowski 1973]:

$$m_t^{**} = \alpha Y_t^{**} + (1 - \alpha)(m_{t-1}^{**} + \delta_{1t-1}^{**}) \quad (21)$$

$$\delta_{1t}^{**} = \beta (m_t - m_{t-1}) + (1 - \beta) \delta_{1t-1}^{**} \quad (22)$$

$$0 \leq \alpha, \beta \leq 1 \quad (23)$$

Predyktory wyjściowy i końcowy wyrażają się wzorami:

$$\Pi_{A_H}^{**} = m_{t_0}^{**} + \delta_{1t_0}^{**} h \quad (24)$$

$$\Pi_{A_H} = \Pi_{A_H}^{**} + M^{(a)}(t) + D^{(a)}(t) \quad (25)$$

Równania postaci multiplikatywnej modelu Holta (M_H) różnią się, podobnie jak w przypadku modelu Browna, jedynie sposobem eliminacji wahań sezonowych.

Postacie predyktorów są następujące:

$$\Pi_{M_H}^{**} = m_{t_0}^{**} + \delta_{1t_0}^{**} h \quad (26)$$

$$\Pi_{M_H} = \Pi_{M_H}^{**} * M^{(m)}(t) * D^{(m)}(t) \quad (27)$$

Jednym z ważniejszych zagadnień związanych z modelowaniem i prognozowaniem z wykorzystaniem modeli adaptacyjnych w warunkach braku pełnej informacji jest wybór optymalnych wartości stałych wygładzania

Dla pełnych danych za optymalne uznaje się te wartości stałych wygładzania, które minimalizują wartość określonego miernika dokładności. Najczęściej są to błąd średnio-kwadratowy (RMSE) i procentowy błąd absolutny (MSE) albo związane z nimi przeciętne błędy względne i pierwiastek kwadratowy ze współczynnika Theila (I) lub przeciętny względny błąd prognozy (MAPE). Oblicza się je jako różnice absolutne lub względne dla przedziału czasowego próby

oraz realizacji zmiennej prognozowanej. Okazuje się jednak, że otrzymanie minimalnych ocen mierników dla wartości wyrównanych nie jest jednoznaczne z otrzymaniem minimalnych ocen błędów prognoz *ex post*.

Jeżeli celem modelowania jest budowa prognoz *ex ante* to podstawą wyboru optymalnych wartości stałych wygładzania powinny być mierniki dokładności prognoz *ex post*. Proces ten ulega komplikacji w przypadku, gdy w szeregach czasowych występują luki w danych. Wyznacza się wtedy dwa rodzaje prognoz: interpolacyjne i ekstrapolacyjne. Prognozy interpolacyjne odnoszą się do tych okresów należących do przedziału czasowego próby, w których wystąpiły luki. Natomiast prognozy ekstrapolacyjne wybiegają poza ten przedział. Jeżeli zna się wartości realizacji w tym okresie to obliczamy mierniki dokładności *ex post*.

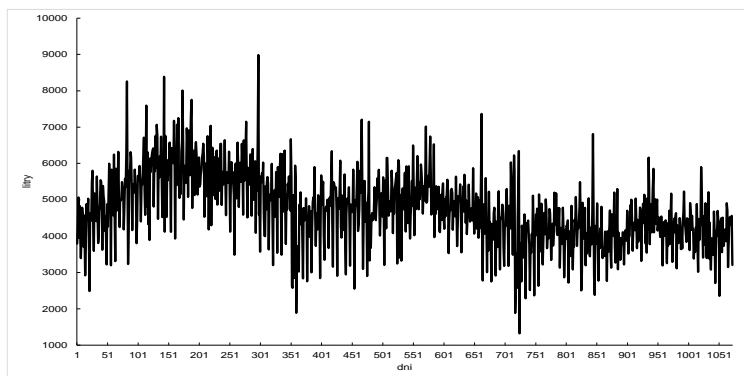
Należy liczyć się z tym, że w omawianej sytuacji otrzyma się najczęściej różne zbiory optymalnych wartości stałych wygładzania odpowiednio dla: wartości wyrównanych (WW), prognoz interpolacyjnych (I) oraz prognoz ekstrapolacyjnych (E)

PRZYKŁAD EMPIRYCZNY

Charakterystyka zmiennej

Modelowaniu i prognozowaniu poddana zostanie dzienna sprzedaż paliw płynnych na stacji benzynowej X (w litrach). Kształtowanie się zmiennej prognozowanej w przedziale czasowym próby przedstawione zostało na rysunku 1.

Rysunek 1. Wielkość sprzedaży paliw płynnych na stacji benzynowej X



Źródło: Bank Danych Katedry Zastosowań Matematyki w Ekonomii ZUT w Szczecinie

W Tabeli 1 zestawione zostały oceny wskaźników i składników sezonowości o cyklach 12 miesięcznym i 7 dniowym..

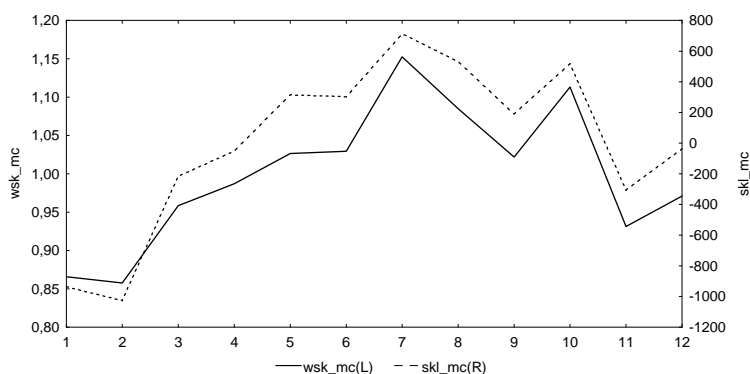
Tabela 1. Oceny wskaźników i składników sezonowości o cyklach 12 miesięcznym i 7 dniowym.

Dzień	Wskaźniki sez.	Składniki sez.	Miesiąc	Wskaźniki sez.	Składniki sez.
Poniedziałek	1,028	131,30	Styczeń	0,866	-936,19
Wtorek	1,015	71,09	Luty	0,858	-1026,38
Środa	1,027	146,10	Marzec	0,958	-216,37
Czwartek	1,108	535,76	Kwiecień	0,987	-50,84
Piątek	1,035	200,12	Maj	1,026	314,46
Sobota	0,829	-832,90	Czerwiec	1,029	303,36
Niedziela	0,957	-251,46	Lipiec	1,153	713,35
			Sierpień	1,085	532,99
			Wrzesień	1,022	191,27
			Październik	1,113	517,72
			Listopad	0,931	-306,09
			Grudzień	0,971	-37,28

Źródło: Szmuksta-Zawadzka, Zawadzki, 2014

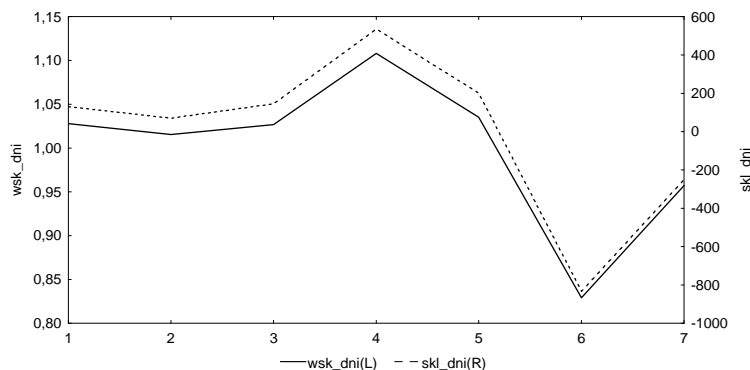
Na podwójnie skalowanych rysunkach 2 i 3 przedstawione zostały w postaci graficznej oceny wskaźników i składników sezonowości wahań o cyklu rocznym i tygodniowym.

Rysunek 2. Oceny wskaźników i składników sezonowości wahań o cyklu rocznym



Źródło: opracowanie własne

Rysunek 3. Oceny wskaźników i składników sezonowości wahań o cyklu tygodniowym.



Źródło: opracowanie własne

Z rysunków wynika, że sprzedaż paliw na badanej stacji charakteryzuje się istotnymi wahaniami zarówno w skali roku jak i tygodnia. Widoczna jest duża zgodność przebiegu wskaźników i składników sezonowości.

Oszacowania wskaźników i składników sezonowości zawarte w Tabeli 1 zostaną wykorzystane najpierw do eliminacji wahań sezonowych a następnie wyznaczania prognoz końcowych

Zakres badań

Rozpatrywany będzie jeden wariant luk niesystematycznych. Luki występować będą w drugim roku przedziału czasowego próby: w poniedziałki, środy i piątki tzn. w 157 spośród 724 dni. Udział luk będzie wynosić zatem 21,69% długości szeregu. Otrzymano je przez „wymazanie” odpowiednich wielkości z pełnego szeregu.

Do budowy prognoz dla danych, z których wyeliminowano wahania o cyklu rocznym (Y_t^*) zostaną wykorzystane modele Holta-Wintersa w postaci addytywnej (A_HW) i multiplikatywnej (M_HW). Natomiast dla danych, z których wyeliminowano zarówno wahania o cyklu rocznym jak i tygodniowym (Y_t^{**}) prognozy będą budowane na podstawie predyktorów opartych na prostych modelach Browna (A_BS i M_BS) oraz modelach Holta (A_H i M_H). Dla każdej kombinacji stałych wygładzania obliczone zostaną oceny błędów względnych: wartości wyrównanych (WW), prognoz interpolacyjnych (I) i prognoz ekstrapolacyjnych (E).

Natomiast analizie poddane zostaną zestawy stałych wygładzania charakteryzujące się minimalnymi ocenami błędów wymienionych wyżej rodzajów wielkości. Dla celów porównawczych zostaną podane także oceny błędów: wartości wyrównanych, prognoz inter- i ekstrapolacyjnych otrzymanych na podstawie predyktorów opartych na modelach szeregu czasowego z liniowym trendem i periodycznym składnikiem sezonowym (A_KL) i trendem

wykładniczym o stałej stopie wzrostu i relatywnie stałych wahaniach sezonowych (M_KL).

Dla modeli charakteryzujących się minimalnymi ocenami błędów prognoz inter- i ekstrapolacyjnych przeprowadzona została dekompozycja błędów prognoz według dni tygodnia i miesięcy.

Pracę kończyć będzie część wnioskowa dotycząca oceny możliwości wykorzystania modeli wyrównania wykładniczego dla danych oczyszczonych z sezonowości w prognozowaniu zmiennych ze złożoną sezonowością w warunkach braku pełnej informacji.

Analiza wyników modelowania i prognozowania

W tabelicy 2 zestawione zostały oceny błędów: wartości wyrównanych (WW), prognoz interpolacyjnych (I) i prognoz ekstrapolacyjnych (E) otrzymane dla parametrów wykładniczych charakteryzujących się minimalnymi ocenami błędów wymienionych wyżej wielkości. Zestawiono w niej także oceny błędów dla modeli klasycznych: A_KL oraz M_KL. W modelach tych wahania cyklu rocznym i tygodniowym były opisane za pomocą osobnych zbiorów zmiennych zero-jedynkowych. Model M_KL różnił się tym od modelu A_KL tym, że po lewej stronie szacowanego równania zamiast zaobserwowanych wartości zmiennej prognozowanej Y_t wystąpiły jej logarytmy naturalne (por. np. [Kufel 2010], [Szmuksta-Zawadzka; Zawadzki 2011]).

Tablica 2. Minimalne oceny średnich błędów względnych: wartości wyrównanych (WW), prognoz interpolacyjnych (I) i prognoz ekstrapolacyjnych (E) oraz optymalne wartości stałych wykładzania

Model	Miernik dotyczy	Stałe wykładzania			MAPE(%)		
		α	β	γ	WW	I	E
MODELE ADDYTYWNE							
A_BS1	WW	0,07			12,00	12,88	16,17
A_BS2	I	0,16			12,26	12,62	15,00
A_BS3	E	0,22			12,48	12,66	14,86
A_H1	WW	0,15	0,15		13,55	12,78	297,72
A_H2	I	0,15	0,03		15,99	12,61	49,02
A_H3	E	0,02	0,06		50,21	17,58	13,31
A_HW1	WW	0,1	0,02	0,01	11,33	13,93	22,97
A_HW2	I	0,1	0,04	0,01	11,36	13,89	45,45
A_HW3	E	0,03	0,08	0,03	12,93	15,04	12,94
A_KL					11,54	15,17	16,59
MODELE MULTIPLIKATYWNE							
M_BS1	WW	0,04			15,68	16,34	16,66
M_BS2	I	0,01			16,30	16,09	18,35
M_BS3	E	0,38			17,22	19,28	12,14

Model	Miernik dotyczy	Stałe wygładzania			MAPE(%)		
		α	β	γ	WW	I	E
M_H1	WW	0,15	0,15		17,22	17,56	258,52
M_H2	I	0,04	0,03		35,13	16,67	17,18
M_H3	E	0,12	0,01		31,48	16,99	11,59
M_HW1	WW	0,1	0,02	0,1	12,74	12,74	17,97
M_HW2	I	0,1	0,04	0,1	12,98	12,69	14,26
M_HW3	E	0,07	0,01	0,03	14,48	13,47	11,20
M_KL					11,51	14,17	15,60

Źródło: opracowanie własne

Występujące przy oznaczeniach modeli kolejne cyfry oznaczają modele charakteryzujące się minimalnymi ocenami błędów: 1 - wartości wyrównanych, 2 - prognoz interpolacyjnych, 3 - prognoz ekstrapolacyjnych.

Analiza błędów wartości wyrównanych oraz błędów obu rodzajów prognoz otrzymanych na podstawie modeli wyrównywania wykładniczego zostanie poprzedzona krótkim omówieniem wyników otrzymanych dla modeli klasycznych. Niższe o 1 p.p. oceny błędów prognoz inter- i ekstrapolacyjnych otrzymano dla predyktora opartego na modelu wykładniczym o stałej stopie wzrostu z relatywnie stałą sezonowością (M_KL). W trakcie omawiania wyników modelowania i prognozowania będziemy odwoływać się do tego modelu klasycznego.

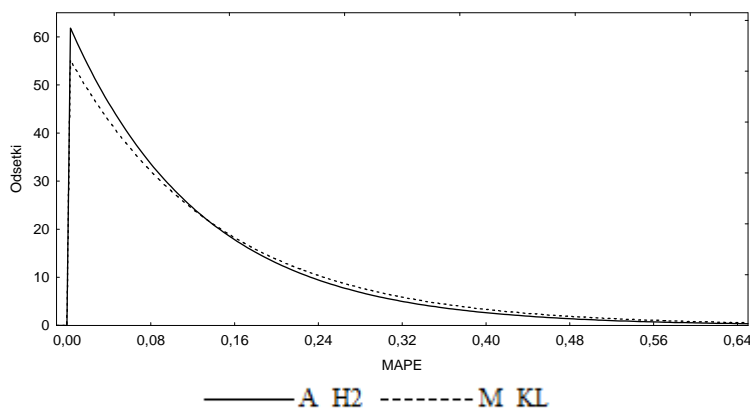
Z informacji zawartych w tabelicy wynika, że minimalną oceną błędu wartości wyrównanych wynoszącą 11,33% otrzymano dla addytywnej postaci modelu Holta-Wintersa (A_HW1) o stałych wygładzania wynoszących: $\alpha=0,1$, $\beta=0,02$ oraz $\gamma=0,1$. Następnym w kolejności był model multiplikatywny Holta-Wintersa (M_HW1) z oceną 12,74%. Spośród modeli z podwójnie eliminowaną sezonowością najniższą ocenę błędu względnego (12,78%) otrzymano dla predyktora opartego na modelu Holta (A_H1) o stałych wygładzania $\alpha=0,15$ i $\beta=0,15$. Oceny przeciętego błędu względnego otrzymanego dla modelu klasycznego (M_KL) jest tylko o 1,68% wyższa od oceny uzyskanej dla modelu A_HW1.

W przypadku prognoz interpolacyjnych z oceną 12,61%, najwyższą efektywnością charakteryzował się model addytywny Holta (A_H2), o takiej samej ocenie parametru α jak w modelu A_HW1, ale ocenie parametru $\beta=0,03$. Model klasyczny (M_KL) charakteryzuje się oceną błędu o 23,7% wyższą. Oceny bardzo zbliżone otrzymano dla dwóch stałych modeli Browna (A_H2 i A_H3) o parametrach wygładzania wynoszących 0,16 oraz 0,22. Oceny tylko nieznacznie wyższe otrzymano dla dwóch modeli Holta w postaci multiplikatywnej (M_H1 i M_H2).

Na rysunku 4 przedstawione zostały aproksymanty teoretyczne empirycznych rozkładów błędów prognoz interpolacyjnych otrzymane na

podstawie najlepszego predyktora adaptacyjnego (A_H2) oraz predyktora klasycznego (M_KL). Z rysunku wynika, że błędy prognoz adaptacyjnych przyjmują częściej niższe wartości.

Rysunek 4. Rozkłady błędów prognoz interpolacyjnych otrzymanych na podstawie modeli A_H2 oraz M_KL



Źródło: opracowanie własne

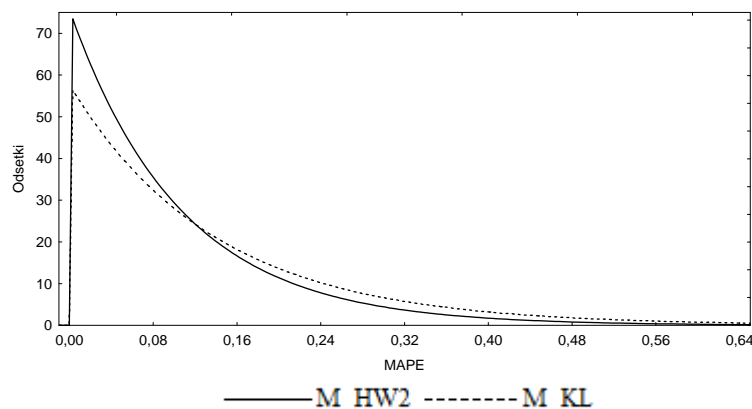
Zdecydowanie najniższą ocenę błędu prognoz ekstrapolacyjnych otrzymano dla multiplikatywnej postaci modelu Holta-Wintersa (M_HW3) o stałych wygładzania wynoszących odpowiednio: $\alpha=0,07$, $\beta=0,01$, $\gamma=0,08$. Wynosiła ona 11,20% i była o 28,21% niższa od oceny dla predyktora klasycznego. Dwie najniższe w kolejności oceny błędu otrzymano także dla postaci multiplikatywnych modeli Holta (M_H3) i Browna (M_BS3) o stałych wygładzania wynoszących: 0,12;0,01 oraz 0,38. Przyjęły one wartości 11,58% oraz 12,14% i były niższe o: 25,77% i 22,18 od oceny otrzymanej dla klasycznego predyktora z trendem wykładniczym i relatywnie stałą sezonowością.

Widzimy zatem, że przy niemal takiej samej ocenie błędu wartości wyrównanych efektywność prognoz interpolacyjnych a zwłaszcza ekstrapolacyjnych okazała się zdecydowanie wyższa od prognoz otrzymanych na podstawie predyktora klasycznego.

Spośród modelu addytywnych najniższą oceną charakteryzuje się model Holta-Wintersa (A_HW3) o stałych wygładzania wynoszących odpowiednio: $\alpha=0,03$, $\beta=0,08$ oraz $\gamma=0,03$. Otrzymana ocena błędu 12,94% jest o 17,06% niższa od uzyskanej dla predyktora klasycznego.

Kształtowanie się aproksymant teoretycznych empirycznych rozkładów błędów prognoz ekstrapolacyjnych otrzymanych na podstawie najlepszego predyktora adaptacyjnego (M_HW2) oraz predyktora klasycznego (M_KL) zostało przedstawione graficznie na rysunku 5.

Rysunek 5 Aproksymanty teoretyczne rozkładów błędów prognoz ekstrapolacyjnych otrzymanych na podstawie modeli M_HW2 oraz M_KL



Źródło: opracowanie własne

Z rysunku wynika, że błędy prognoz adaptacyjnych, podobnie jak w przypadku prognoz interpolacyjnych, przyjmują częściej niższe wartości niż dla modelu klasycznego. Różnice w rozkładach są bardziej widoczne niż dla prognoz interpolacyjnych.

W tabelicy 3 zestawione zostały zdezagregowane oceny błędów prognoz interpolacyjnych i ekstrapolacyjnych otrzymane na podstawie równań o minimalnych ocenach błędów. W celach porównawczych zestawiono w niej także zdezagregowane błędy prognoz dla modelu klasycznego. Z uwagi na to, że prognozy interpolacyjne odnoszą się do poniedziałków, śród i piątków i dotyczą one okresów, w których wystąpiły luki, zostanie podana ich liczba. To samo dotyczy miesięcy.

Tablica 3. Zdezagregowane oceny błędów prognoz inter- i ekstrapolacyjnych według dni tygodnia i miesięcy.

	Błędy prognoz interpolacyjnych (%)			Błędy prognoz ekstrapolacyjnych (%)	
	Liczba prognoz	A_H2	M_KL	M_HW3	KL
Ogółem	157	12,61	14,18	11,20	15,65
Poniedziałek	52	8,04	8,72	10,71	21,09
Wtorek	0			8,45	12,49
Środa	53	18,65	18,24	9,91	14,05
Czwartek	0			10,28	12,93
Piątek	52	11,02	15,49	8,95	18,53
Sobota	0			18,99	18,89
Niedziela	0			11,23	16,35
Styczeń	14	6,70	9,53	11,13	12,30

	Błędy prognoz interpolacyjnych (%)			Błędy prognoz ekstrapolacyjnych (%)	
	Liczba prognoz	A_H2	M_KL	M_HW3	KL
Luty	12	9,10	10,40	12,94	9,62
Marzec	13	12,94	11,25	7,05	9,71
Kwiecień	13	13,28	13,81	10,93	12,34
Maj	13	11,99	13,84	11,48	23,92
Czerwiec	13	12,51	14,23	9,26	21,87
Lipiec	13	5,76	9,94	6,96	14,38
Sierpień	13	9,30	11,54	7,91	12,78
Wrzesień	13	5,06	6,04	11,98	17,08
Październik	14	11,29	16,98	14,12	20,75
Listopad	12	8,59	10,30	16,93	17,61
Grudzień	14	42,49	39,84	17,72	14,65

Źródło: opracowanie własne

Z informacji zawartych w tabeli 3 wynika, że dla poniedziałków i śródtę dni zbliżone są oceny błędów prognoz interpolacyjnych otrzymanych na podstawie najlepszego z modeli adaptacyjnych i otrzymanych na podstawie predyktora klasycznego. Natomiast około 4,5 p.p. niższe błędy otrzymano na podstawie modelu Holta.

W przypadku dezagregacji na miesiące najniższą ocenę dla modelu Holta wynoszącą 5,06% otrzymano dla września. Błędami nieznacznie wyższymi charakteryzują się lipiec i styczeń. Oceny błędów w granicach 9% otrzymano także dla lutego, sierpnia i listopada. Najwyższą ocenę przekraczającą 40% otrzymano dla grudnia. Dla modelu klasycznego najniższą ocenę otrzymano także dla września – jest ona ok. 1 p.p. wyższa niż dla predyktora adaptacyjnego. Oceny w granicach 10% otrzymano dla stycznia, lipca, lutego oraz listopada. Niższe oceny błędów ok. 1,5 p.p. i 2,7 p.p. otrzymano na podstawie predyktora klasycznego odpowiednio dla marca i listopada. Największą różnicę wynoszącą ponad 5,5 p.p. odnotowano dla października. Najwyższą wartość błędu dla predyktora klasycznego przyjął także w grudniu. Wynikać to mogło, z nagłego załamania się pogody w porównaniu z rokiem poprzednim.

W przypadku prognoz ekstrapolacyjnych przeciętna ocena błędu dla predyktora adaptacyjnego była niższa o 4,45 p.p. od błędu dla predyktora klasycznego. Oceny błędów prognoz ekstrapolacyjnych dla dni tygodnia dla modelu Holta-Wintersa wahają się do 8,45% dla wtorku do 18,95% i 18,99% dla piątku i soboty. Dla pozostałych dni kształtują się na poziomie ok. 10%-19%.

Dla modelu klasycznego oceny błędów wahają się od 12,45% i 12,93% dla wtorku i czwartku do 21,09% dla poniedziałku. Najwyższe różnice dla dni tygodnia, wynoszące 10,38p.p. i 9,58 otrzymano dla poniedziałku i piątku a najniższą – minus 0,0010 dla soboty. Najniższą ocenę błędów w przypadku

miesiący dla predyktora opartego na modelu multiplikatywnym Holta-Wintersa otrzymano dla lipca i marca – wynoszą one odpowiednio: 6,96% oraz 7,05%. Oceny w granicach 8%-9% otrzymano dla sierpnia i czerwca. Natomiast ocenami najwyższymi charakteryzują się grudzień (17,72%) i listopad (16,93%).

Oceny błędów otrzymane dla predyktora klasycznego zawarte są w przedziale od 9,62% - 9,71% dla lutego i marca do 23,92% dla maja. Maksymalne różnice dokładności prognoz wynoszące 12,62 p.p. i 12,44 p.p. otrzymano odpowiednio dla czerwca i maja. Jedynie w grudniu błąd prognozy dla predyktora klasycznego był niższy (o 3,07 p.p.) od błędu dla predyktora adaptacyjnego.

PODSUMOWANIE

Z przeprowadzonych w pracy badań wyprowadzić można następujące wnioski:

1. Kryterium wyboru modelu dla celów prognozowania nie mogą być przeciętne błędy względne wartości wyrównanych, lecz błędy względne prognoz interpolacyjnych lub prognoz ekstrapolacyjnych.
2. W przypadku modelowania i prognozowania zmiennych o niezbyt silnej dynamice, w których występują luki niesystematyczne, optymalne wartości stałych wygładzania w modelach Holta i Holta-Wintersa przybierają wartości bliskie zeru.
3. Minimalne oceny błędów: wartości wyrównanych, prognoz interpolacyjnych oraz prognoz ekstrapolacyjnych otrzymano na podstawie różnych modeli adaptacyjnych.
4. Dokładność prognoz interpolacyjnych otrzymana na podstawie najlepszego predyktora adaptacyjnego była ok. 23,7% wyższa od dokładności prognoz dla predyktora klasycznego.
5. Dokładność prognoz ekstrapolacyjnych otrzymanych na podstawie najlepszych predyktorów adaptacyjnych był ok. 22,2% - 28,2 % wyższa od dokładności dla predyktora klasycznego.
6. W toku badań wykazano, modele wyrównywania wykładniczego dla danych oczyszczonych z sezonowości mogą być użytecznym narzędziem prognozowania zmiennych ekonomicznych ze złożoną sezonowością.

BIBLIOGRAFIA

- Dittmann P. (2006) Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie, Wolters Kluwer Polska, Kraków.
- Kufel T. (2010) Ekonometryczna analiza cykliczności procesów gospodarczych o wysokiej częstotliwości obserwowania, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń.
- Pawłowski Z. (1973) Prognozowanie ekonometryczne, PWN, Warszawa.

- Szmuksta-Zawadzka M., Zawadzki J. (2011) Zastosowanie modelowania ekonometrycznego w prognozowaniu brakujących danych w szeregach o wysokiej częstotliwości, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu. *Ekonometria* 34, Wrocław, str. 303-314.
- Szmuksta-Zawadzka M., Zawadzki J. (2012) Z badań nad metodami prognozowania na podstawie niekompletnych szeregów czasowych z wahaniami okresowymi (sezonowymi), *Przegląd Statystyczny*, Tom 59, s.140-154, Warszawa.
- Szmuksta-Zawadzka M., Zawadzki J. (2014) Modele wyrównywania wykładniczego w prognozowaniu zmiennych ekonomicznych ze złożoną sezonowością (w druku).
- Zeliaś A., Pawełek B., Wanat S. (2003) Prognozowanie ekonomiczne. Teoria, przykłady, zadania. PWN, Warszawa.

THE APPLICATION OF SELECTED ADAPTATION MODELS IN FORECASTING THE MISSING DATA IN THE TIME SERIES WITH COMPLEX SEASONALITY FOR UNSYSTEMATIC GAPS

Abstract: The paper is devoted to the application of selected exponential smoothing models: Brown, Holt and Holt-Winters in prediction of variables with complex seasonality in the condition of lack of full information. Output forecasts will be built on the basis of time series cleansed from seasonality. Final forecasts, taking into account seasonal fluctuations, will be a sum of output forecasts and seasonal components or multiply of forecasts and the seasonal indicators. Theoretical considerations will be illustrated by an empirical example.

Keywords: complex seasonality, exponential smoothing, forecasting, gaps in data

WYKORZYSTANIE DATA ENVELOPMENT ANALYSIS DO PORÓWNIANIA POLITYKI WIEKOWEJ W KRAJACH UNII EUROPEJSKIEJ

Andrzej Szuwarzyński

Wydział Zarządzania i Ekonomii, Politechnika Gdańska

e-mail: Andrzej.Szuwarzynski@zie.pg.gda.pl

Streszczenie: Celem artykułu jest zbudowanie modelu, bazującego na Data Envelopment Analysis (DEA), pozwalającego na porównanie działań związanych z polityką wiekową w krajach Unii Europejskiej. Zastosowano model z ograniczeniami na wagi, pozwalający na stworzenie rankingu i określenie kierunków kształtowania polityki wiekowej. Wykorzystano dane pochodzące z bazy danych Global AgeWatch Index 2013.

Słowa kluczowe: DEA, efektywność, polityka wiekowa

WSTĘP

Starzenie się społeczeństw i związane z tym skutki społeczno-ekonomiczne są jednymi z głównych problemów polityków na całym świecie. Kształtowanie polityk na poziomie rządów wymaga zarówno wiarygodnych danych jak również metod ich analizy. Takie inicjatywy jak Active Ageing Index 2012 [Zaidi i in. 2013], czy Global AgeWatch Index 2013 [HelpAge 2013], mają na celu wspieranie rozwoju programów poprawiających sytuację osób starszych. Wyniki tych analiz pozwalają na określenie słabych i mocnych stron poszczególnych krajów oraz na stworzenie rankingu, wykorzystując złożone wskaźniki (Composite Indicators – CIs). Stwierdzono jednak, że same wskaźniki nie zapewniają wystarczających informacji, jakie konkretne polityki i programy są konieczne [HelpAge 2013].

W 1950 było na świecie 205 mln osób w wieku 60 lat i powyżej, a do 2012 roku ich liczba wzrosła do 810 mln (11,5% całkowitej światowej populacji). Przewiduje się, że w roku 2050 ich udział się podwoi, osiągając 2 miliardy. Liczba osób, które każdego roku na całym świecie przekraczają 60 lat, wynosi prawie 58 mln, co odpowiada blisko dwóm osobom w każdej sekundzie. W roku 2050 po

raz pierwszy będzie więcej ludzi starszych niż dzieci poniżej 15 lat [UNFPA 2012].

Proces starzenia się jest wielowymiarowy i obejmuje zmiany fizyczne, psychologiczne i społeczne. Zwiększanie się długości życia rodzi pytanie, czy będzie to więcej lat życia w dobrym zdrowiu, czy też wiąże się z większą liczbą lat spędzonych w długotrwałej niesprawności i zależności. Średnia długość życia w krajach rozwiniętych wzrosła z 45-50 lat w roku 1900 do około 80 obecnie [UNFPA 2012]. Starzenie się społeczeństwa powoduje szybki wzrost wydatków na zdrowie, gdyż osoby starsze wymagają większej opieki zdrowotnej [UN 2013].

Celem tego artykułu jest dokonanie oceny efektywności polityki wiekowej w krajach Unii Europejskiej (UE), poprzez utworzenie rankingu oraz wskazanie kierunków zmian, jakie powinny być wprowadzone w krajach, gdzie efektywność polityki wiekowej nie jest zadowalająca. Obliczenia wykonano na podstawie danych Global AgeWatch Index 2013, przy zastosowaniu metody DEA.

PRZYKŁADY OCENY EFEKTYWNOŚCI POLITYKI WIEKOWEJ

Ostatnim pełnym badaniem polityki wiekowej jest inicjatywa AgeWatch Index 2013 [HelpAge 2013], dotycząca 91 krajów z całego świata. Nadrzędnym celem jest wspieranie rozwoju polityk i programów mających poprawić jakość życia i dobrobyt obecnych a przede wszystkim przyszłych pokoleń starszych ludzi. Badania uwzględniają cztery obszary: bezpieczeństwo dochodów; stan zdrowia; aktywność na rynku pracy i poziom wykształcenia oraz ograniczenia funkcjonowania w środowisku, które opisane są trzynastoma wskaźnikami cząstkowymi. Innym przykładem jest Active Ageing Index 2012 [Zaidi i in. 2013], który ma podobny cel, czyli dostarczenie mierników, które mogą być wykorzystane do kształtowania polityki wiekowej. Badaniami objęto 27 krajów UE. Zdefiniowano 22 wskaźniki podstawowe, w czterech obszarach: zatrudnienie; uczestnictwo w społeczeństwie; niezależność, zdrowie i bezpieczeństwo życia oraz sprzyjające możliwości środowiska dla aktywnego starzenia się.

W obu tych badaniach zastosowano metodykę złożonych wskaźników, powszechnie stosowaną w analizach polityk makroekonomicznych [np. Cherchye i in. 2009; Despotis 2005; Ramanathan 2006]. Na podobnych zasadach budowane są również znane, światowe rankingi uczelni, takie jak Shanghai Academic Ranking of World Universities, czy też brytyjski Times Higher Education Supplement. Dyskutowana jest jednakże wiarygodność rankingów stosujących metodykę złożonych wskaźników, głównie z powodu zbyt dużego subiektywizmu w doborze struktury stosowanych modeli, w tym głównie wag, przypisywanych do poszczególnych wskaźników cząstkowych, które ustalane są zwykle na podstawie ocen eksperckich [Marginson 2014; Paruolo i in. 2013; Saisana i in. 2011]. W wielu przypadkach przyjmuje się jednakowe wagi dla wszystkich wskaźników wchodzących w skład wskaźnika złożonego [np. Hoffman i in. 2008].

Zastosowana metodyka składa się z czterech etapów [HelpAge 2013; Zaidi i in. 2013]. Pierwszy, to analiza surowych danych mająca na celu sprawdzenie, czy wszystkie wskaźniki cząstkowe są stymulantami (jeżeli nie, to konieczne jest ich przekształcenie). Każdy wskaźnik cząstkowy musi spełniać warunek, że jego wyższa wartość oceniana jest pozytywnie. Drugi etap to normalizacja danych, z reguły wykorzystująca metodę min-max, co pozwala sprowadzić wszystkie wskaźniki do przedziału [0-100]. Następnie agreguje się wskaźniki z każdego obszaru, wykorzystując arytmetyczną lub geometryczną średnią ważoną. W ostatnim etapie dokonuje się na podobnej zasadzie agregacji wskaźników obszarowych w jeden złożony wskaźnik, będący podstawą do stworzenia rankingu.

W obu przedstawionych przypadkach badań polityki wiekowej podkreśla się, że budowanie wskaźnika złożonego opiera się jedynie na rezultatach i nie uwzględnia procesów, które prowadzą do ich osiągnięcia ani nakładów, jakie są do tego potrzebne [HelpAge 2013; Zaidi i in. 2012].

PROPONOWANY MODEL DEA DO OCENY EFEKTYWNOŚCI POLITYKI WIEKOWEJ

Metoda DEA wykorzystuje programowanie liniowe, nieuwzględniające wpływu czynnika losowego i błędów pomiaru, a także niewymagające określenia zależności funkcyjnej między nakładami a rezultatami. Optymalne wagi są obliczane na podstawie danych, a nie ustalane subiektywnie [Cooper i in. 2007]. DEA bada efektywność względną obiektów, zwanych jednostkami decyzyjnymi (DMU – Decision Making Units). Powszechnie stosowany jest model CCR (od nazwisk twórców Charnes, Cooper i Rhodes), ze stałymi efektami skali, pozwalający na obliczenie efektywności technicznej [Charnes i in. 1978]. Ocenia on radialną efektywność, nie uwzględniając jednakże luzów. Zgodnie z definicją efektywności DEA, funkcjonowanie DMU jest w pełni (100%) efektywne, wtedy i tylko wtedy, gdy zarówno wynik efektywności jest równy jedności oraz nadwyżki nakładów i niedobory rezultatów są zerowe.

Pomiar efektywności polega na wyznaczeniu obiektów wzorcowych oraz przyrównywaniu do nich pozostałych obiektów. Wynik efektywności θ_o w modelu CCR zorientowanym na rezultaty, dla grupy odniesienia DMU_j ($j=1, \dots, n$), jest obliczany dla rezultatów (y_{rj} , $r=1, \dots, s$) i nakładów (x_{ij} , $i=1, \dots, m$), co można zapisać w następujący sposób [Zhu i in. 2007]:

$$\min \theta_o = \sum_{i=1}^m v_i x_{io} \quad (1)$$

z warunkami:

$$\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} \geq 0 \quad (2)$$

$$\sum_{r=1}^s u_r y_{ro} = 1 \quad (3)$$

$$u_r, v_i > 0 \quad (4)$$

gdzie: u_r, v_i są wagami zmiennych, które są określane przez rozwiązanie powyższego problemu, na podstawie danych ze wszystkich DMU.

Stosowanie klasycznych modeli radialnych pozwala każdej DMU uzyskać najbardziej korzystny wynik efektywności, co często wiąże się z zerowymi wartościami wag, które nie są akceptowalne w rzeczywistych zastosowaniach [Murias i in. 2008; Roll, Golany 1993]. Takie modele nie uwzględniają więc wszystkich źródeł nieefektywności [Ramón i in. 2010]. Całkowita elastyczność wag DEA jest zaletą, ale również poważną wadą. Prowadzi to często do nieracjonalnych wyników, gdyż ocenia się efektywność DMU, poprzez przypisanie niezerowych wag tylko do niektórych nakładów i rezultatów [Cooper i in. 2011; Murias i in. 2008]. Można tego uniknąć przez nałożenie ograniczeń na wagi, które poprawiają siłę dyskryminacji pomiędzy badanymi DMU [Angulo-Meza, Lins 2002]. Nie istnieją żadne formalne reguły nakładania ograniczeń na wagi [Roll, Golany 1993]. Można je ustalić, odwołując się do opinii ekspertów [Cherchye i in. 2009] lub określić je na podstawie wartości zmiennych ocenianych DMU, stosując poniższą technikę [Ramón i in. 2010, Roll, Golany 1993]:

1. Uruchamia się nieograniczony model radialny, uzyskując „macierz wag”, na podstawie której określa się średnie wagi u_r i v_i dla rezultatów i nakładów.
2. Ustala się wielkość dopuszczalnych zmienności w obrębie wag dla tego samego czynnika, jako stosunek d , wartości najwyższej do najniższej.
3. Rozszerza się podstawowy model radialny o dodatkowe ograniczenia typu:

$$L_r = \frac{2 \times u_r}{1+d} \leq u_{rj} \leq \frac{2 \times d \times u_r}{1+d} = U_r \quad (5)$$

4. Uruchamia się model „z ograniczeniami”.

Średnie wagi oblicza się ze zredukowanego wektora wag, przez pominięcie ekstremalnych wartości [Roll, Golany 1993] lub wykorzystując jedynie w pełni efektywne DMU [Angulo-Meza, Lins 2002]. W podobny sposób można wprowadzić ograniczenia na wirtualne wyjścia. Jest to udział wirtualnego wyjścia DMU_j powiązany z rezultatem r , czyli „znaczenie” przypisane do tego rezultatu dla DMU_j, ograniczone do zakresu $[L_r, U_r]$ w formie [Angulo-Meza, Lins 2002; Cherchye i in. 2007; Ramón i in. 2010]:

$$L_r \leq \frac{u_r y_{rj}}{\sum_{r=1}^s u_r y_{rj}} \leq U_r \quad (6)$$

Dodanie tych ograniczeń tworzy model Assurance Region Global (ARG), który został wykorzystany w tym artykule. Obliczenia przeprowadzono w programie DEA Solver-LV(3) firmy Saitech.

Problemem, w zastosowaniach DEA do budowy złożonych wskaźników, przeznaczonych do oceny polityk makroekonomicznych, jest określenie tego, co ma być nakładami (zgodnie z metodyką obliczania AgeWatch Index rozważane są tylko rezultaty). Wykorzystuje się tu koncepcję helmsman'a (władz podejmujących decyzje), którą wprowadził Koopmans, badając zagadnienia efektywności w systemach zdecentralizowanych decyzji [Koopmans 1951]. Zakłada się, że każdy kraj ma do dyspozycji narzędzia umożliwiające prowadzenie własnej polityki makroekonomicznej (helmsman) i tak steruje wszystkimi wskaźnikami cząstkowymi, aby osiągnąć ich najwyższe dopuszczalne poziomy [Lovell 1995; Murias i in. 2008]. Tak więc w modelu DEA wektor wejść ogranicza się do jednej zmiennej dummy, o wartości równej jedności dla każdej DMU [Despotis 2005; Lovell i in. 1995]. Taki model zakłada, że wszystkie kraje mają, w równym stopniu, możliwości do osiągnięcia pełnej efektywności [Cherchye i in. 2009], a ocena działania badanej jednostki uwzględnia tylko rezultaty a nie sposób, w jaki są one osiągnęte. Ponieważ nie ma transformacji nakładów na rezultaty, określa się to raczej, jako pomiar skuteczności a nie efektywności [Cooper i in. 2009].

STRUKTURA DANYCH MODELU

Wykorzystano bazę danych AgeWatch Index 2013, która zawiera dane pochodzące z baz: Banku Światowego, OECD, Eurostatu, Międzynarodowej Organizacji Pracy, Światowej Organizacji Zdrowia i Instytutu Gallupa. Trzydzieści zmiennych podzielono na cztery obszary [HelpAge 2013].

Obszar bezpieczeństwa dochodowego zawiera cztery zmienne:

1. (R1) Wskaźnik pokrycia dochodów ze świadczeń emerytalnych – udział beneficjentów programów emerytalnych do liczby osób w wieku 65 plus.
2. (R2) Wskaźnik ubóstwa w starszym wieku – udział osób w wieku 60 plus, których dochód jest poniżej progu ubóstwa.
3. (R3) Względny dobrobyt starszych ludzi – średni dochód osób w wieku 60 plus w stosunku do reszty populacji danego kraju.
4. (R4) Produkt Krajowy Brutto per capita – zmienna proxy określająca ogólny poziom życia w danym kraju.

Obszar stanu zdrowia zawiera trzy zmienne:

5. (R5) Oczekiwana długość życia w wieku 60 lat – średnia liczba lat życia, jaka pozostała osobom, które osiągnęły wiek 60 lat.
6. (R6) Oczekiwana długość życia w zdrowiu w wieku 60 lat – średnia liczba lat życia w dobrym zdrowiu fizycznym, jaka pozostała osobom w wieku 60 lat.
7. (R7) Względne dobre samopoczucie psychiczne – stosunek udziału osób w wieku 50 plus, które odpowiedziały „tak” na pytanie: „Czy uważasz, że życie ma ważny cel lub sens?”, do udziału osób w wieku 35-49 lat, które odpowiedziały „tak” na to same pytanie.

Obszar zatrudnienia i edukacji zawiera dwie zmienne:

8. (R8) Zaangażowanie starszych ludzi na rynku pracy – procent osób w wieku 55-64 lata, które są zatrudnione – miara dostępu osób starszych do rynku pracy.
9. (R9) Poziom wykształcenia starszych osób – procent populacji w wieku 60 plus mających wykształcenie co najmniej średnie.

Obszar ograniczeń funkcjonowania w środowisku zawiera cztery zmienne:

10. (R10) Więzy społeczne – udział osób w wieku 50 plus, które odpowiedziały twierdząco na pytanie, „Jeżeli masz kłopoty, czy masz krewnych lub przyjaciół, od których możesz liczyć na pomoc?”.
11. (R11) Bezpieczeństwo fizyczne – udział osób w wieku 50 plus, które odpowiedziały twierdząco na pytanie, „Czy czujesz się bezpiecznie spacerując samotnie w nocy w mieście, gdzie mieszkasz?”.
12. (R12) Wolność obywatelska – udział osób w wieku 50 plus, które udzieliły odpowiedzi twierdzącej na pytanie, „Czy jesteś zadowolony z wolności wyboru, co zrobić ze swoim życiem?”.
13. (R13) Dostęp do transportu publicznego – udział osób w wieku 50 plus, które udzieliły odpowiedzi twierdzącej na pytanie, „Czy tam gdzie mieszkasz jesteś zadowolony z systemu transportu publicznego?”.

Stosując model DEA konieczne jest zapewnienie, aby zbiór DMU był jednorodny, a wzrost wartości rezultatów, z punktu widzenia celu badania, powinien być oceniany pozytywnie [Guzik, 2009]. Przeprowadzono analizę na obecność obserwacji odstających korzystając z wykresu ramka-wąsy (program Statistica). Jedynie dla jednej zmiennej R3, dwa kraje: Cypr i Luksemburg są obserwacjami odstającymi. Ponieważ zastosowany model ma ograniczenia na wagi, zdecydowano o włączeniu tej zmiennej do modelu. Zmienną R2, która jest destymulantą, przekształcono odejmując wartość wskaźnika od 100.

Siła dyskryminacji DEA jest zależna od relacji między liczbą ocenianych DMU i łączną liczbą nakładów i rezultatów. Obowiązuje praktyczna zasada [Cook i in. 2014], że liczba DMU, uwzględnianych w analizie, powinna być trzy razy większa niż łączna liczba nakładów i rezultatów. Przy 28 porównywanych krajach (DMU), konieczne było zredukowanie zestawu zmiennych. Dokonując doboru zmiennych przyjęto zasadę, że każdy obszar musi być opisany przynajmniej dwoma zmiennymi. Wyeliminowano zmienną R1 dotyczącą świadczeń emerytalnych, gdyż dla niektórych krajów przyjmuje ona wartości większe od 100%, co wynika z tego, że wiek emerytalny w niektórych krajach jest niższy od granicy 65 lat, do której ten wskaźnik jest odnoszony. W obliczaniu AgeWatch Index, dla krajów, gdzie wskaźnik ten przekraczał 100%, przyjęto wartość 100%, aby uniknąć nieuzasadnionej przewagi tych krajów [HelpAge 2013]. Dla pozostałych zmiennych obliczono w obrębie każdego obszaru współczynniki korelacji. W obszarze bezpieczeństwa dochodowego, pozostałe trzy zmienne mają współczynniki korelacji w przedziale [0,32-0,42]. Zmienne R5 i R6 mają współczynnik korelacji równy 0,97 i przenoszą praktycznie te same informacje. Na tej podstawie wyeliminowano zmienną R5. W obszarze zatrudnienia i edukacji

pozostawiono obie zmienne (korelacja na poziomie 0,05). Natomiast w obszarze warunków otoczenia usunięto zmienne R11 i R12, ponieważ współczynniki korelacji z pozostałymi zmiennymi były na poziomie 0,6-0,8. Dwie zmienne R10 i R13, jakie pozostają w zestawie, mają współczynnik korelacji 0,40.

WYNIKI OBLICZEŃ I DYSKUSJA

W przypadku zastosowania modelu CCR bez ograniczeń występuje problem zerowych wag. Żadna DMU, przy obliczaniu wyników nie wykorzystuje wszystkich 9 rezultatów. Największa liczba, pięciu zmiennych, była wykorzystana do obliczenia wyniku efektywności dla jednej DMU. Dla 11 DMU wykorzystano dwa rezultaty a dla 2 DMU jeden. Połowa badanych krajów osiągnęła pełną efektywność, co potwierdza zasadność wprowadzenia ograniczeń na wagi.

W modelu z ograniczeniami, cztery kraje osiągnęły pełną efektywność. Na liście referencyjnej dla nieefektywnych DMU, Holandia znalazła się 19 razy, Niemcy 7 razy, Luksemburg 5 razy a Francja 4 razy. Wszystkie wagi jak również luzy są niezerowe dla wszystkich rezultatów. Wyniki przedstawione są w Tabeli 1.

Tabela 1. Porównanie rankingów DEA i AgeWatch Index 2013

Kraj	Ranking		θ	DEA-AWI	Kraj	Ranking		θ	DEA-AWI
	AWI	DEA				AWI	DEA		
Holandia	2	1	1	-1	Węgry	18	15	0,926	-3
Niemcy	3	2	1	-1	Estonia	16	16	0,917	0
Luksemburg	7	3	1	-4	Łotwa	21	17	0,907	-4
Francja	10	4	1	-6	Dania	9	18	0,904	9
Szwecja	1	5	0,994	4	Słowacja	19	19	0,904	0
Austria	5	6	0,987	1	Malta	20	20	0,898	0
Irlandia	4	7	0,967	3	Grecja	22	21	0,886	-1
Słowenia	11	8	0,963	-3	Polska	23	22	0,875	-1
Wielka Brytania	6	9	0,950	3	Litwa	26	23	0,867	-3
Włochy	14	10	0,950	-4	Rumunia	28	24	0,851	-4
Belgia	13	11	0,940	-2	Cypr	24	25	0,840	1
Finlandia	8	12	0,933	4	Bułgaria	25	26	0,834	1
Czechy	15	13	0,930	-2	Chorwacja	27	27	0,829	0
Hiszpania	12	14	0,930	2	Portugalia	17	28	0,743	11

Źródło: obliczenia własne

Ranking opracowany na podstawie modelu DEA (DEA) został porównany z rankingiem AgeWatch Index 2013 (AWI). Kraje uszeregowane są według wyniku efektywności (θ), czyli pozycji w rankingu DEA. Porównano też rankingi wyliczone według proponowanego modelu oraz według metodyki AgeWatch Index (DEA-AWI). Cztery kraje nie zmieniły swojej pozycji, siedem krajów zmieniło swoją pozycję o jedno miejsce, trzy kraje zmieniły swoją pozycję o dwa miejsca, pięć krajów zmieniło swoją pozycję o trzy miejsca a dziewięć krajów o więcej niż trzy miejsca, z czego Dania o 9 a Portugalia o 11 miejsc. Różnice te wynikają z zastosowania innego systemu wag. W modelu AWI, wagi zostały ustalone przez

ekspertów, jednakowe dla wszystkich krajów. Każdy z czterech obszarów ma wagę 0,25, co oznacza równy udział każdego z nich w zagregowanym wskaźniku. Poszczególne zmienne mają wagi zróżnicowane, np. w obszarze „zatrudnienie i edukacja”, zmienne R8 i R9 mają równe wagi, a więc ich udział w całym wskaźniku wynosi po 0,125. Natomiast w modelu DEA średni udział zmiennej R8 wynosi 0,006 a zmiennej R9 wynosi 0,1. Ważną zaletą DEA, w porównaniu do metody złożonych wskaźników, jest możliwość ilościowego określenia zmian wartości rezultatów, jakie są konieczne, dla osiągnięcia pełnej efektywności. Tabela 2 przedstawia projekcję dla kilku krajów, gdzie podana jest procentowa zmiana wartości zmiennych, która zapewni efektywność porównywalną z liderami.

Tabela 2. Projekcja dla wybranych krajów [%]

Wskaźnik	Bułgaria	Cypr	Polska	Portugalia	Rumunia	Słowenia	Dania
R2	23,37	26,90	8,55	0,00	20,65	7,53	8,69
R3	0,00	40,31	11,54	2,70	0,00	0,00	19,30
R4	221,38	42,23	117,20	54,95	239,98	34,04	9,32
R6	23,21	2,30	11,54	1,16	18,87	2,96	7,23
R7	8,97	0,00	38,84	10,52	31,47	18,97	12,74
R8	18,57	0,00	36,12	16,90	33,11	61,58	-12,44
R9	62,05	161,49	15,19	432,15	18,87	0,00	45,02
R10	15,45	24,32	3,75	16,48	35,50	4,61	0,91
R13	23,21	21,92	11,54	2,70	18,87	0,00	9,32

Źródło: obliczenia własne

Dla Bułgarii, Polski i Rumunii jedną z podstawowych przyczyn nieefektywności jest zbyt niska wartość PKB per capita (R4), co oczywiście trudne jest do zmiany w krótkim horyzoncie czasowym, lecz wskazuje, że prowadzenie polityki wiekowej wymaga dużych nakładów finansowych. Za wyjątkiem Polski, Rumunii i Słowenii konieczne jest zintensyfikowanie działań edukacyjnych. W przypadku Danii wartość dla zmiennej R8 jest ujemna – zaangażowania osób starszych na rynku pracy, co jest to wynikiem aktywnej polityki rynku pracy, bazującej na koncepcji flexicurity. Oczywiście należy to interpretować, jako pewną nadwyżkę tego wskaźnika i na pewno nie należy go redukować, a wykorzystać duńskie doświadczenia w pozostałych krajach. Wyniki wskazują kierunki zmian w polityce, z czego część musi być kierowana do młodszych pokoleń, gdyż efekty w podnoszeniu poziomu wykształcenia, zaangażowaniu na rynku pracy czy zwiększeniu oczekiwanej długości życia w zdrowiu, to działania długofalowe, związane z kształtowaniem polityki edukacyjnej, rynku pracy i opieki zdrowotnej.

PODSUMOWANIE

Uzyskane wyniki pozwoliły na zbudowanie rankingu oraz wskazanie kierunków, w jakich poszczególne kraje powinny prowadzić swoje polityki, aby osiągnąć wyniki porównywalne z liderami. Zastosowanie ograniczeń na wirtualne wyjścia zwiększyło siłę dyskryminacji modelu i uwzględnienie w obliczeniach

efektywności wszystkich wskaźników, poprzez eliminację zerowych wag. Dwa rankingi wykazują w kilku przypadkach duże różnice, co spowodowane jest innym systemem przypisywania wag. W modelu DEA, wagi dla poszczególnych wskaźników cząstkowych są ustalane dla każdej DMU indywidualnie, uwzględniając łączny wpływ wszystkich czynników a w przypadku złożonych wskaźników są subiektywnie ustalane dla wszystkich krajów na tym samym poziomie dla poszczególnych wskaźników cząstkowych.

BIBLIOGRAFIA

- Angulo-Meza L., Lins M.P.E. (2002) Review of Methods for Increasing Discrimination in Data Envelopment Analysis, *Annals of Operations Research*, 116, s. 225-242.
- Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, s. 429-444.
- Cherchye L., Moesen W., Rogge N., Van Puyenbroeck T. (2007) An introduction to 'benefit of the doubt' composite indicators, *Social Indicators Research*, 82, s. 111-145.
- Cherchye L., Moesen W., Rogge N., Van Puyenbroeck T. (2009) Constructing a Knowledge Economy Composite Indicator with Imprecise Data, *Hub Research Paper – Economics & Management*, 2009/16, Katholieke Universiteit Leuven.
- Cook W.D., Tone K., Zhu J. (2014) Data envelopment analysis: Prior to choosing a model. *Omega-International Journal of Management Science*, 44, s. 1-4.
- Cooper W.W., Seiford L.M., Tone K. (2007) *Data Envelopment Analysis, A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Springer, New York.
- Cooper W.W., Ruiz J.L., Sirvent I. (2009) Selecting non-zero weights to evaluate effectiveness of basketball players with DEA, *European Journal of Operational Research*, 195, s 563-574.
- Cooper W.W., Seiford L.M., Zhu, J. (2011) *Handbook on Data Envelopment Analysis*, Springer, New York.
- Despotis D.K. (2005) Measuring human development via Data Envelopment Analysis: the case of Asia and the Pacific, *Omega-International Journal of Management Science*, 33, s. 385-390.
- Guzik B. (2009) *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.
- HelpAge (2013) *Global AgeWatch Index 2013. Purpose, methodology and results*, HelpAge International, London, www.globalagewatch.org
- Hoffman A., Giovanni E., Nardo M., Saisana M., Saltelli A., Tarantola S. (2008) *Handbook on Constructing Composite Indicators. Methodology and User Guide*, OECD Publishing, Paris.
- Koopmans T.C. (1951) Analysis of production as an efficient combination of activities, in: T.C. Koopmans (Ed.), *Activity Analysis of Production and Allocation*, Wiley, New York.
- Lovell K.C.A. (1995) Measuring the macroeconomic performance of the Taiwanese economy, *Int. J. Production Economics*, 39, s. 165-178.

- Lovell K.C.A., Pastor J.T., Turner J.A. (1995) Measuring macroeconomic performance in the OECD: A comparison of European and non-European countries, *European Journal of Operational Research*, 87 s. 507-518.
- Marginson S. (2014) University Rankings and Social Science. *European Journal of Education*, 49 (1), s. 45-59.
- Murias, P., de Miguel, J.C. i Rodriguez D. (2008) A Composite Indicator for University Quality Assesment: The Case of Spanish Higher Education System, *Social Indicators Research*, 89, s. 129-146.
- Paruolo P, Saisana M., Saltelli, A. (2013). Ratings and rankings: voodoo or science?, *Journal of the Royal Statistical Society. Series A, Statistics in Society*, 176 (3), 609-634.
- Ramanathan R. (2006) Evaluating the comparative performance of countries of the Middle East and North Africa: A DEA application, *Socio-Economic Planning Sciences*, 40, s. 156-167.
- Ramón N., Ruiz J.L., Sirvent I. (2010) A multiplier bound approach to assess relative efficiency in DEA without slacks, *European Journal of Operational Research*, 203, s. 261-269.
- Roll Y., Golany B. (1993) Alternate Methods of Treating Factor Weights in DEA, *Omega-International Journal of Management Science*, vol. 21, no. 1, s. 99-109.
- Saisana M., d'Hombres B., Saltelli A. (2011) Rickety numbers: Volatility of university rankings and policy implications. *Research Policy*, 40, s. 165-177.
- UN (2013) *World Population Ageing 2013*, Department of Economic and Social Affairs, Population Division, United Nations, New York.
- UNFPA (2012) *Ageing in the Twenty-First Century: A Celebration and A Challenge*, United Nations Population Fund, New York and HelpAge International, London.
- Zaidi A., Gasior K., Hofmarcher M.M., Lelkes O., Marin B., Rodrigues R., Schmidt A., Vanhuysse P., Zolyomi E. (2013) *Active Ageing Index 2012. Concept, Methodology and Final Results*, Project: 'Active Ageing Index (AAI)' UNECE Grant No: ECE/GC/2012/003, United Nations Economic Commission for Europe (UNECE), Geneva.
- Zhu J., Cook W.D. (2007) *Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis*, Springer, New York.

USE OF DATA ENVELOPMENT ANALYSIS TO COMPARE AGE POLICIES IN THE EUROPEAN UNION COUNTRIES

Abstract: The aim of this paper is to build a model, based on Data Envelopment Analysis (DEA), allowing the comparison of activities related to the age policy in the European Union countries. A model with bounded weights was adopted to create a ranking and determine the direction of age policy-making. Data from AgeWatch Global Index 2013 database were used.

Keywords: DEA, efficiency, ageing policy

WYKORZYSTANIE WSKAŹNIKÓW ŚRODKÓW PIENIĘŻNYCH W ANALIZIE FINANSOWEJ PRZEDSIĘBIORSTWA

Tomasz Śpiewak

Katedra Finansów, Uniwersytet Rzeszowski
e-mail: tspiewak@op.pl

Streszczenie: Wskaźniki analizy ekonomicznej są powszechnie wykorzystywane do oceny poziomu kondycji finansowej podmiotów gospodarczych. Do podstawowych zadań analizy wskaźnikowej należy identyfikacja tych procesów gospodarczych zachodzących w przedsiębiorstwach, które w istotny sposób wpływają na poziom jego kondycji finansowej. W związku z powyższym analiza finansowa opierająca się jedynie na danych o charakterze memoriałowym nie jest w stanie dostarczyć wymaganych informacji. Głównym celem artykułu jest przedstawienie wybranych gotówkowych wskaźników finansowych oraz wskazanie ich znaczenia dla diagnozy rzeczywistego poziomu kondycji finansowej przedsiębiorstwa na przykładzie spółki Hydrobudowa S.A.

Słowa kluczowe: analiza finansowa, przepływy pieniężne, gotówkowe wskaźniki finansowe

WSTĘP

Wskaźniki analizy ekonomicznej są powszechnie wykorzystywane do oceny poziomu kondycji finansowej podmiotów gospodarczych. Badania nad możliwością stosowania narzędzi analizy finansowej w zarządzaniu przedsiębiorstwem mają w krajach o rozwiniętej gospodarce rynkowej długą historię, sięgającą XIX wieku. W Polsce dynamiczny rozwój analizy ekonomicznej nastąpił po 1989 roku i był następstwem transformacji systemu gospodarczego oraz zmian zachodzących w otoczeniu funkcjonowania przedsiębiorstw. Głównym celem analizy ekonomicznej jest identyfikacja czynników, które mają istotny wpływ na poziom kondycji finansowej przedsiębiorstwa. Analiza ekonomiczna dzieli się na analizę finansową i analizę ekonomiczno-techniczną. Kondycja finansowa przedsiębiorstwa określana jest

w literaturze jako ocena stanu jego finansów i efektywności finansowej jego gospodarki.¹ W literaturze najczęściej prezentowane są wskaźniki poziomu kondycji finansowej, których konstrukcja oparta jest na bilansie oraz rachunku zysków i strat - elementach sprawozdania finansowego bazujących na zasadzie memoriałowej rachunkowości. Jednak analiza opierająca się jedynie na danych o charakterze memoriałowym nie jest w stanie dostarczyć informacji niezbędnej dla obiektywnej oceny poziomu kondycji finansowej jednostki. Analiza rachunku przepływów pieniężnych stanowi ważne uzupełnienie analizy bilansu oraz rachunku zysków i strat poprzez dostarczenie dodatkowych informacji charakteryzujących zdolność firmy do maksymalizacji wartości kapitału własnego. Głównym celem artykułu jest przedstawienie wybranych gotówkowych wskaźników finansowych oraz wskazanie ich znaczenia dla diagnozy rzeczywistego poziomu kondycji finansowej przedsiębiorstwa na przykładzie spółki Hydrobudowa S.A. Celem przeprowadzonej analizy jest również porównanie objętości informacyjnej wskaźników memoriałowych oraz wskaźników opartych na rachunku przepływów pieniężnych. W opracowaniu przedstawione zostały wyniki analizy wskaźnikowej przeprowadzonej na podstawie danych finansowych przedsiębiorstwa, obejmujących lata 2003-2010.

PRZEPIŁYWY PIENIĘŻNE W ANALIZIE KONDYCJI FINANSOWEJ PRZEDSIĘBIORSTWA- ICH ROLA ORAZ ZASADY KONSTRUKCJI

Głównym zadaniem analizy ekonomicznej jest określenie rzeczywistego poziomu kondycji finansowej przedsiębiorstwa. Kluczowym etapem dla poprawności sformułowania ostatecznej diagnozy jest dobór właściwych narzędzi wykorzystywanych w procesie badawczym. Przeprowadzający analizę winien dostosować wykorzystywane wskaźniki do specyfiki podmiotu oraz branży, w jakiej działa badane przedsiębiorstwo. Poprawny dobór zróżnicowanych wskaźników wpływa na wszechstronną ocenę poziomu kondycji finansowej badanej jednostki. W literaturze przedmiotu bardzo często prezentowane wskaźniki ograniczają się do mierników opartych na memoriałowych składnikach sprawozdania finansowego - bilansu i rachunku wyników. Pominięcie wskaźników rachunku przepływów pieniężnych może doprowadzić do błędnej diagnozy w odniesieniu do badanych obszarów - np. w odniesieniu do płynności lub wypłacalności jednostki.

Determinanty znaczenia rachunku przepływów pieniężnych w ocenie kondycji finansowej:

¹ Dudycz T., Wrzosek S. (2000), *Analiza finansowa. Problemy metodyczne w ujęciu praktycznym*, Wydawnictwo AE, Wrocław, s. 28.

1. Wpływ zasad rachunkowości wykorzystywanych przy tworzeniu sprawozdań finansowych na wielkość szacowanych mierników kondycji finansowej.² Zasady rachunkowości, zwłaszcza memoriałowa, pomimo standaryzacji, których źródłem są Międzynarodowe Standardy Sprawozdawczości Finansowej (MSSF), umożliwiają wpływanie na ostateczny poziom wielkości kluczowych z punktu widzenia poziomu kondycji finansowej, w efekcie prowadzi to do odmiennego obrazu tej sytuacji w różnych krajach, co utrudnia obiektywną ocenę przedsiębiorstw w skali międzynarodowej - występuje problem porównywalności osiągniętych np. wyników finansowych.
2. Dla inwestorów istotne znaczenie mają wskaźniki oparte na realnych przepływach pieniężnych, a nie memoriałowych kategoriach finansowych - w rezultacie nastąpił rozwój wskaźników opartych na konstrukcji miernika ekonomicznej wartości dodanej (EVA) oraz zarządzania nastawionego na zwiększanie wartości przedsiębiorstwa.
3. Według opinii inwestorów i analityków finansowych przepływy pieniężne oraz wskaźniki na nich oparte są mniej podatne na manipulacje księgowe, a zatem bardziej wiarygodne w ocenie poziomu kondycji finansowej podmiotu.³
4. Zarządzanie przepływami pieniężnymi stanowi podstawę zarządzania płynnością finansową w przedsiębiorstwie, której poziom decyduje o jego przetrwaniu.

Kategoria przepływów pieniężnych wykorzystywana w konstrukcji wskaźników analizy finansowej obliczana jest przy wykorzystaniu formuł różnicowanych z punktu widzenia poziomu ich dokładności, np.:

wynik netto + amortyzacja (1)

wynik netto + amortyzacja + zmiana stanu kapitału obrotowego (2)

przepływy pieniężne obliczone
w oparciu o rachunek przepływów pieniężnych (3)

Bardzo istotny z punktu widzenia doboru wskaźników w analizie finansowej jest stopień korelacji pomiędzy wynikiem finansowym netto i przepływami pieniężnymi jednostki

² Black A, Wright Ph., Bachman J. (2000), W poszukiwaniu wartości dla akcjonariuszy, Dom Wydawniczy ABC, Warszawa, s. 51.

³ Sharma D., Iselin E. (2003), The decision Usefulness of Reported Cash Flow and Accrual Information in a Behavioral Field Experiment, Accounting and Business Research, nr 2/2003.

W swoich badaniach L. Austin i M. Bradbury zweryfikowali następujące hipotezy:⁴

1. operacyjne przepływy pieniężne mają istotnie różną pojemność informacyjną od wyniku finansowego netto,
2. przepływy pieniężne obliczone według metody uproszczonej (2) znacząco różnią się od przepływów pieniężnych netto ze sprawozdania.

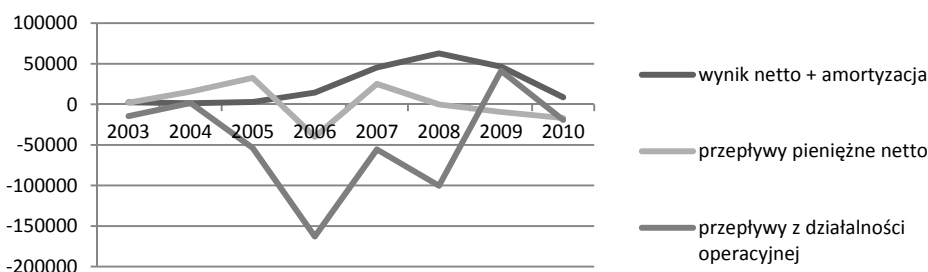
W rezultacie przeprowadzonych badań, potwierdzili oni, że przepływy pieniężne netto są istotnie różne od wyniku finansowego, a także, że formuły (2) i (3) różnią się od siebie.

Stopień korelacji wyniku finansowego i przepływów pieniężnych w warunkach polskiej gospodarki badał D. Wędzki, który dokonał:⁵

1. weryfikacji hipotezy, że przepływy pieniężne netto różnią się istotnie od wyniku finansowego netto, co uzasadnia potrzebę sporządzania rachunku przepływów pieniężnych,
2. wyboru formuły uproszczonej (1) lub (2) do oszacowania operacyjnych przepływów pieniężnych netto.

Badanie potwierdziło istotność przepływów pieniężnych netto jako kategorii finansowej różnej od wyniku finansowego netto. Wykazało także, że formuła (2) jest umiarkowanie skorelowana z przepływami operacyjnymi, a zatem prawdopodobnie również z przepływami pieniężnymi netto.

Rysunek 1. Przepływy pieniężne badanego przedsiębiorstwa wyznaczone na różnym poziomie dokładności (w tys. PLN)



Źródło: opracowanie własne

Na Rysunku 1 przedstawiony został poziom przepływów pieniężnych badanego przedsiębiorstwa, obliczonych na podstawie wzorów wykorzystywanych w analizie finansowej. W obliczeniach zostały wykorzystane dane sprawozdania finansowego badanego przedsiębiorstwa, okres objęty analizą obejmuje lata 2003-

⁴ Austin L., Bradbury M. (1995), The Accuracy of Cash Flow Estimation Procedures, Accounting and Finance, nr 1.

⁵ Wędzki D. (2003), Statystyczna weryfikacja istotności przepływów pieniężnych, Zeszyty Teoretyczne Rachunkowości, nr 15 (71), s. 92.

2010. Poziom poszczególnych kategorii finansowych jest bardzo zróżnicowany, co niewątpliwie wpływa na poziom wskaźników finansowych w zależności od wykorzystanego wariantu definicji przepływów.

Prowadzone w przeszłości badania naukowe potwierdzają obserwacje procesów zachodzących w praktyce w zakresie znaczenia przepływów pieniężnych:

- przepływy pieniężne netto w istotnym stopniu różnią się od wyniku finansowego netto,
- uproszczone formuły obliczania przepływów pieniężnych są słabiej skorelowane z przepływami pieniężnymi netto, obliczonymi na podstawie sprawozdania finansowego.

Przedstawione wnioski z przeprowadzonych badań stanowią podstawę do potwierdzenia słuszności wykorzystania w ocenie poziomu kondycji finansowej przedsiębiorstwa rachunku przepływów pieniężnych, ponieważ mierniki oparte na tych przepływach mogą dostarczyć innych informacji niż mierniki oparte na danych sprawozdania finansowego opartego na zasadzie memoriałowej rachunkowości.

CHARAKTERYSTYKA I KLASYFIKACJA WSKAŹNIKÓW ANALIZY FINANSOWEJ OPARTYCH NA RACHUNKU PRZEPŁYWÓW PIENIĘŻNYCH

Rachunek przepływów pieniężnych weryfikuje informację zawartą w bilansie i rachunku zysków i strat oraz pozwala na ocenę gotówkowych zmian zachodzących w aktywach i pasywach jednostki. Jest bardzo istotnym uzupełnieniem pozostałych składników sprawozdania finansowego z punktu widzenia analizy finansowej. Znajomość struktury oraz wielkości przepływów pieniężnych pozwala na realną ocenę osiągniętych wyników memoriałowych wykazywanych w rachunku zysków i strat.⁶ Analiza rachunku przepływów pieniężnych obejmuje dwa poziomy oceny: analizę wstępną oraz analizę wskaźnikową.

W ramach analizy wstępnej przeprowadzana jest ogólna ocena struktury, wielkości oraz kierunków zmian poszczególnych pozycji rachunku przepływów pieniężnych. Na tym etapie oceny dokonuje się analizy przepływów według rodzajów działalności i typów operacji. W literaturze przedmiotu bardzo często analiza wstępna przedstawiana jest na przykładzie analizy ośmiu przypadków znaków przepływów pieniężnych z działalności operacyjnej, inwestycyjnej oraz finansowej.

⁶ E.Śniezek (2005), Przepływy pieniężne jako szczególny obszar integracji rachunkowości finansowej i rachunkowości zarządczej, Zeszyty Teoretyczne Rachunkowości, nr 30, s.110.

Analiza wskaźnikowa jest rozwinięciem wstępnej analizy sprawozdania finansowego. Jest wariantem analizy ilościowej, która za pomocą określonych relacji wielkości ekonomicznych przedstawia obraz kondycji finansowej ocenianego przedsiębiorstwa. W analizie wskaźnikowej rachunku przepływów pieniężnych wykorzystywane są wskaźniki charakteryzujące wybrane obszary przepływów pieniężnych oraz ich powiązania z pozostałymi składnikami sprawozdania finansowego.

Według T. Waśniewskiego i innych można wyróżnić trzy grupy wskaźników opartych na rachunku przepływów pieniężnych, a mianowicie:⁷

- struktury przepływów pieniężnych,
- wystarczalności środków pieniężnych,
- wydajności gotówkowej.

Wskaźniki struktury charakteryzują relacje zachodzące pomiędzy podstawowymi pozycjami rachunku przepływów pieniężnych. Ich konstrukcja może być dostosowana jest do indywidualnego przypadku badanego przedsiębiorstwa. Jeśli w sprawozdaniu występują np. dodatnie przepływy operacyjne oraz ujemne inwestycyjne oraz finansowe, licznik wskaźnika obejmuje przepływy dodatnie, zaś mianownik wartość bezwzględną sumy przepływów ujemnych. W efekcie otrzymujemy informację o stopniu pokrycia ujemnych przepływów inwestycyjnych i finansowych przez dodatnie przepływy operacyjne.

Wskaźniki wystarczalności środków pieniężnych stanowią podstawę oceny zdolności przedsiębiorstwa do pokrycia wybranych wydatków inwestycyjnych oraz finansowych przez przepływy operacyjne. Wskaźniki zaliczane do tej grupy stanowią istotne uzupełnienie klasycznych wskaźników płynności oraz wypłacalności wykorzystywanych w analizie finansowej.

Wskaźniki wydajności środków pieniężnych charakteryzują poziom efektywności gospodarowania gotówką w przedsiębiorstwie. Mierniki zaliczane do tej grupy wskazują na zdolność generowania operacyjnych przepływów pieniężnych w relacji do wybranych pozycji rachunku zysków i strat oraz bilansu.⁸

WYKORZYSTANIE WYBRANYCH WSKAŹNIKÓW PIENIĘŻNYCH W ANALIZIE FINANSOWEJ PRZEDSIĘBIORSTWA - STUDIUM PRZYPADKU

W celu ukazania roli rachunku przepływów pieniężnych w analizie finansowej przedsiębiorstwa, przedstawione zostało porównanie następujących

⁷ Waśniewski T., Skoczylas W. (2004), Teoria i praktyka analizy finansowej w przedsiębiorstwie, FRR w Polsce, Warszawa, s.458-463.

⁸ G. Susek-Borek (2001), Ocena płynności finansowej przedsiębiorstwa, Rachunkowość, nr 7, s.432.

grup wskaźników: opartych na memoriałowych składnikach sprawozdania finansowego, ich branżowych odpowiedników (średnia dla branży) oraz wskaźników, których konstrukcja opiera się na rachunku przepływów pieniężnych. Zarząd badanego przedsiębiorstwa złożył w 2011 r. wniosek o jego upadłość, analiza obejmuje dane sprawozdania finansowego za okres 2003-2010.

Wskaźnik bieżącej płynności oraz wskaźnik pokrycia pieniężnego zobowiązań bieżących

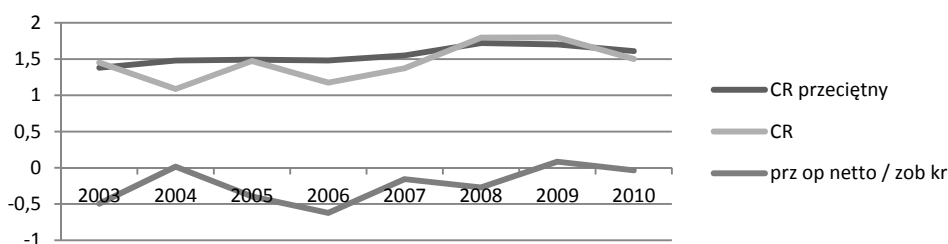
Wskaźnik bieżącej płynności CR (4) jest podstawowym miernikiem potencjalnej zdolności przedsiębiorstwa do spłacenia jego zobowiązań bieżących poprzez upłynnienie posiadanych zasobów majątku obrotowego. Wskaźnik ten informuje o kwocie aktywów obrotowych przypadających na jednostkę zobowiązań krótkoterminowych przedsiębiorstwa. Wskazuje, ile razy można spłacić zobowiązania krótkoterminowe aktywami obrotowymi⁹.

Wskaźnik pokrycia pieniężnego zobowiązań bieżących WPPZB (5) charakteryzuje stopień pokrycia zobowiązań bieżących przedsiębiorstwa przez środki pieniężne, których źródłem jest działalność operacyjna. Wskaźnik ten przedstawia realną zdolność jednostki do spłaty zobowiązań krótkoterminowych przez wypracowane środki pieniężne w ramach działalności podstawowej.

$$CR = \frac{\text{aktywa obrotowe}}{\text{zobowiązania bieżące}} \quad (4)$$

$$WPPZB = \frac{\text{przepływy operacyjne netto}}{\text{zobowiązania bieżące}} \quad (5)$$

Rysunek 2. Poziom wybranych wskaźników badanego przedsiębiorstwa w latach 2003-2010 - wskaźnik bieżącej płynności oraz wskaźnik pokrycia pieniężnego zobowiązań bieżących.



Źródło: obliczenia własne

⁹ M. R. Tyran (2004), Wskaźniki finansowe, Oficyna Ekonomiczna, Kraków, s. 106.

Na Rysunku 2 przedstawiony został poziom omawianych wskaźników. W okresie od 2008 do 2009 roku wskaźnik bieżącej płynności badanego przedsiębiorstwa był wyższy od średniego poziomu dla całej branży. Badana spółka złożyła wniosek o upadłość w 2011 r. Poziom wskaźnika WPPZB był niższy od zera na skutek ujemnej wartości operacyjnych przepływów pieniężnych (wyjątkiem był okres 2009 r.).

Marża zysku ze sprzedaży oraz wskaźnik pieniężnej wydajności sprzedaży

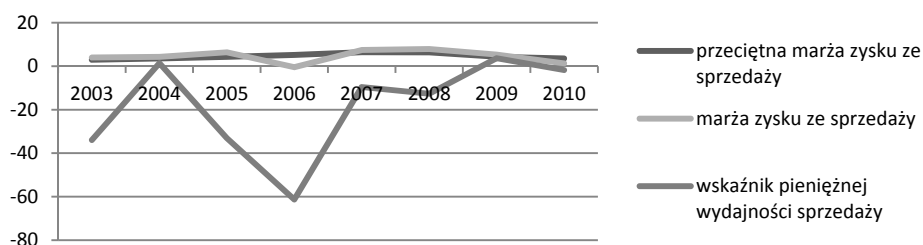
Wskaźniki rentowności sprzedaży określają, jaki jest poziom różnego rodzaju korzyści w relacji do współmiernym tym korzyściom przychodom ze sprzedaży. Wskaźniki marży informują o kwocie określonego wyniku osiągniętego z jednostki sprzedaży.

Marża zysku ze sprzedaży MZS (6) określa ilość zysku ze sprzedaży przypadającą na jednostkę sprzedaży. Pozytywny jest wzrost omawianego wskaźnika. Wysoki poziom omawianego wskaźnika jest szczególnie ważny w przedsiębiorstwach, w których największy wpływ na kondycję finansową wywiera obszar działalności podstawowej. Występowanie straty na poziomie działalności operacyjnej jednostki prowadzi w dłuższym okresie do utraty płynności oraz upadłości. Wskaźnik pieniężnej wydajności sprzedaży WPWS (7) przedstawia wartość środków pieniężnych wypracowywanych w ramach działalności operacyjnej przedsiębiorstwa przypadającą na jednostkę przychodów ze sprzedaży produktów, towarów i materiałów.

$$\text{MZS} = \frac{\text{wynik ze sprzedaży}}{\text{przychody ze sprzedaży produktów, towarów i materiałów}} \quad (6)$$

$$\text{WPWS} = \frac{\text{przepływy operacyjne netto}}{\text{przychody ze sprzedaży produktów, towarów i materiałów}} \quad (7)$$

Rysunek 3. Poziom wybranych wskaźników badanego przedsiębiorstwa w latach 2003-2010 - marża zysku ze sprzedaży oraz wskaźnik pieniężnej wydajności sprzedaży



Źródło: opracowanie własne

Wartości poziomu wskaźnika marży zysku ze sprzedaży badanego przedsiębiorstwa są zbliżone do średniej branżowej tego wskaźnika. Analiza wskaźnika pieniężnej wydajności sprzedaży wskazuje na potencjalne problemy finansowe przedsiębiorstwa - jego wartość jest pochodną ujemnych przepływów operacyjnych jednostki. Podstawowa działalność powinna stanowić źródło środków pieniężnych w przedsiębiorstwie, które mogą być wykorzystane w ramach działalności inwestycyjnej lub finansowej.

Rentowność ekonomiczna sprzedaży oraz gotówkowa wydajność ekonomiczna sprzedaży

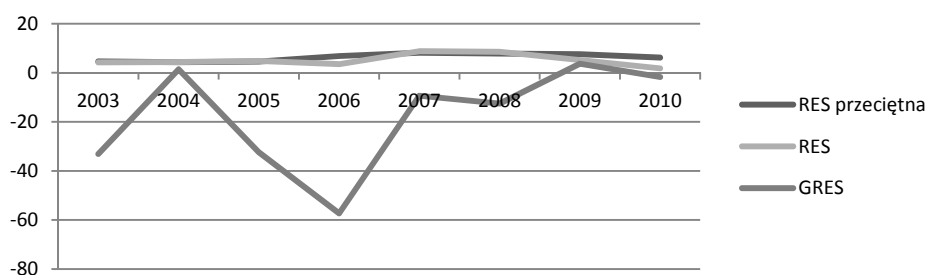
Rentowność ekonomiczna sprzedaży RES (8) wykorzystuje w swojej konstrukcji potencjalny strumień środków pieniężnych wypracowanych w ramach działalności operacyjnej (wynik operacyjny + amortyzacja). Wskaźnik ten powinien wykazywać tendencję rosnącą, co może oznaczać wzrost efektywności działalności operacyjnej w relacji do osiągniętych przychodów.

Wskaźnik gotówkowej wydajności ekonomicznej GRES (9) określa wartość przepływów operacyjnych netto generowanych przez przychody ze sprzedaży oraz pozostałe przychody operacyjne. Określa rzeczywistą efektywność pieniężną przychodów operacyjnych jednostki. Wzrost tego wskaźnika jest oceniany pozytywnie, spadek może świadczyć o obniżającym się poziomie efektywności wykorzystania aktywów operacyjnych przedsiębiorstwa.

$$RES = \frac{\text{wynik operacyjny} + \text{amortyzacja}}{\text{przychody ze sprzedaży} + \text{pozostałe przychody operacyjne}} \quad (8)$$

$$GRES = \frac{\text{przepływy operacyjne netto}}{\text{przychody ze sprzedaży} + \text{pozostałe przychody operacyjne}} \quad (9)$$

Rysunek 4. Poziom wybranych wskaźników badanego przedsiębiorstwa w latach 2003-2010 - wskaźnik rentowności ekonomicznej sprzedaży oraz wskaźnik gotówkowej wydajności ekonomicznej sprzedaży



Źródło: opracowanie własne

W okresie objętym analizą wskaźnik rentowności ekonomicznej sprzedaży nie wykazuje znacznych odchyłeń od wartości średniej dla branży, w jakiej działa przedsiębiorstwo. Natomiast wskaźnik pieniężnej wydajności sprzedaży charakteryzuje się większą zmiennością. Dwa lata przed ogłoszeniem upadłości, poziom obu wskaźników obniża się. Obserwacja poziomu badanych wskaźników potwierdza stawianą tezę o słabej korelacji przepływów pieniężnych obliczanych na podstawie uproszczonej definicji przepływów operacyjnych oraz przepływów rzeczywistych.

Wskaźnik rentowności operacyjnej aktywów ROA oraz wskaźnik pieniężnej wydajności majątku ROA CF

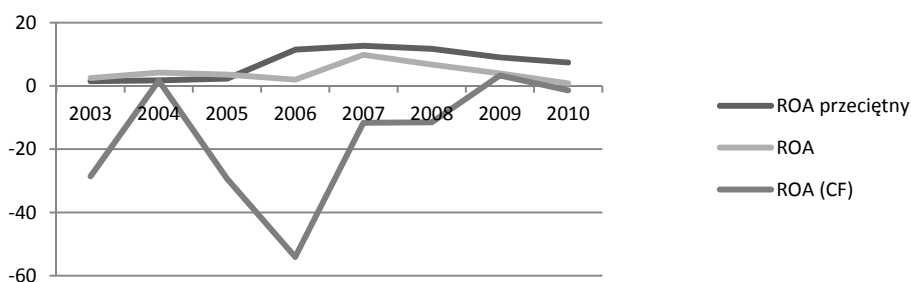
Stopa zwrotu z aktywów ROA (10) informuje, jaka ilość wypracowanego w przedsiębiorstwie zysku przed spłatą odsetek i opodatkowaniem przypada na jednostkę posiadanego majątku. Omawiany wskaźnik określa rentowność operacyjną majątku i stanowi podstawę oceny opłacalności zaangażowania kapitału obcego w przedsiębiorstwie z punktu widzenia poziomu dźwigni finansowej.

W formule wskaźnika pieniężnej wydajności majątku ROA CF (11) kategorię wyniku operacyjnego (charakter memoriałowy) zastąpiła kasowa pozycja przepływów operacyjnych netto. Wskaźnik informuje o ilości przepływów operacyjnych netto przypadających na jednostkę średnich aktywów jednostki. Wskaźnik ten wskazuje rzeczywistą efektywność operacyjną majątku.

$$ROA = \frac{\text{wynik operacyjny}}{\text{aktywa średnioroczne}} \quad (10)$$

$$ROA (CF) = \frac{\text{przepływy operacyjne netto}}{\text{aktywa średnioroczne}} \quad (11)$$

Rysunek 5. Poziom wybranych wskaźników badanego przedsiębiorstwa w latach 2003-2010 - Wskaźnik rentowności operacyjnej aktywów ROA oraz wskaźnik pieniężnej wydajności majątku ROA CF



Źródło: opracowanie własne

Również w tym przypadku poziom wskaźnika kasowego (ROA CF) wyraźnie wskazuje na potencjalne problemy finansowe jednostki - w okresie objętym analizą wykazuje wartości ujemne (wyjątkiem jest rok 2009). Wskaźnik rentowności operacyjnej aktywów przedsiębiorstwa tylko w 2010 roku przyjmuje wartości poniżej zera - w pozostałym okresie nie odbiega znacząco od wartości średniej wskaźnika dla branży, w jakiej działa jednostka.

Reasumując, we wszystkich przedstawionych przypadkach analiza trendów kształtowania się poszczególnych wskaźników o charakterze memoriałowym w nieznacznym stopniu wskazuje na wzrost ryzyka upadłości badanego przedsiębiorstwa. Wskaźniki wykazują nieznaczne odchylenia od wartości średnich branżowych, co może wpłynąć na pozytywną ocenę poziomu kondycji finansowej jednostki przedstawioną przez podmiot badający. Wskaźniki wykorzystujące dane rachunku przepływów pieniężnych wykazują większe wahania, zaś ich ujemny poziom wyraźnie wskazuje na wzrost ryzyka upadłości przedsiębiorstwa. Wynika to z konstrukcji rachunku przepływów pieniężnych - jeśli w przedsiębiorstwie występuje ujemny poziom gotówkowych przepływów z działalności operacyjnych, może to oznaczać problemy płatnicze jednostki.

ZAKOŃCZENIE

Alternatywą dla memoriałowej metody ustalania wyniku finansowego, wykorzystywanej w ocenie poziomu kondycji finansowej przedsiębiorstwa, jest oparty na zasadzie kasowej rachunkowości rachunek przepływów pieniężnych. Staje się on podstawą dla coraz większej liczby przeprowadzanych analiz finansowych jednostek.¹⁰ Wyniki przeprowadzonego porównania wskaźników memoriałowych oraz wskaźników pieniężnych potwierdziły znaczenie gotówkowych wskaźników w analizie finansowej przedsiębiorstwa - ich poziom wskazywał na potencjalne problemy finansowe jednostki z wyprzedzeniem w porównaniu z ich memoriałowymi odpowiednikami. Sytuacja, w której przeprowadzający analizę finansową coraz częściej zwracają uwagę na słabą wiarygodność diagnostyczną wyniku finansowego sporządzanego zgodnie z zasadą memoriałową, może spowodować wzrost zainteresowania rachunkiem przepływów pieniężnych jako podstawą do oceny poziomu kondycji finansowej przedsiębiorstwa. Z nielicznych badań wynika, że model mieszany, tj. oparty zarówno na wskaźnikach memoriałowych, jak i kasowych posiada wyższą

¹⁰ Śnieżek E., (2013) Wykorzystanie relacji między elementami pieniężnymi i memoriałowymi w procesach decyzyjnych - przegląd badań empirycznych, *Acta Universitatis Lodzensis, Folia Oeconomica* 263, s.124.

zdolność prognostyczną od modelu bazującego tylko na wskaźnikach memoriałowych.¹¹

Należy podkreślić, że z punktu widzenia jakości przeprowadzanej analizy finansowej, istotny jest wybór odpowiedniej definicji przepływów pieniężnych wykorzystanej w konstrukcji wskaźników. Aktualne badania empiryczne sugerują, iż wykorzystanie rachunku przepływów pieniężnych w modelowaniu poziomu kondycji finansowej przedsiębiorstwa dostarcza lepszych prognoz zarówno od modeli preferujących uproszczone sposoby kalkulacji przepływów jak też modeli opartych tylko na wskaźnikach memoriałowych.

BIBLIOGRAFIA

- Austin L., Bradbury M. (1995), *The Accuracy of Cash Flow Estimation Procedures*, Accounting and Finance, nr 1.
- Black A., Wright Ph., Bachman J. (2000), *W poszukiwaniu wartości dla akcjonariuszy*, Dom Wydawniczy ABC, Warszawa.
- Dudycz T., Wrzosek S. (2000), *Analiza finansowa. Problemy metodyczne w ujęciu praktycznym*, Wydawnictwo AE, Wrocław.
- Sharma D., Iselin E. (2003), *The decision Usefulness of Reported Cash Flow and Accrual Information in a Behavioral Field Experiment*, Accounting and Business Research, nr 2/2003.
- Susek-Borek G. (2001), *Ocena płynności finansowej przedsiębiorstwa*, Rachunkowość, nr 7.
- Śniezek E. (2005), *Przepływy pieniężne jako szczególny obszar integracji rachunkowości finansowej i rachunkowości zarządczej*, Zeszyty Teoretyczne Rachunkowości, nr 30.
- Śniezek E., (2013) *Wykorzystanie relacji między elementami pieniężnymi i memoriałowymi w procesach decyzyjnych - przegląd badań empirycznych*, Acta Universitatis Lodziensis, Folia Oeconomica 263.
- Tyran M. R. (2004), *Wskaźniki finansowe*, Oficyna Ekonomiczna, Kraków.
- Waśniewski T., Skoczylas W. (2004), *Teoria i praktyka analizy finansowej w przedsiębiorstwie, FRR w Polsce*, Warszawa.
- Wędzki D. (2003), *Statystyczna weryfikacja istotności przepływów pieniężnych*, Zeszyty Teoretyczne Rachunkowości, , nr 15 (71).

¹¹ Wędzki D. (2003), *Statystyczna weryfikacja istotności przepływów pieniężnych*, Zeszyty Teoretyczne Rachunkowości, nr 15 (71), s. 102.

**USE OF THE CASH INDICATORS
IN THE FINANCIAL ANALYSIS OF THE ENTERPRISE**

Abstract: The most common indicators presented in the literature usually focus only on balance sheets and income statements - elements of the financial statements based on the accrual basis of accounting. The statement of cash flows complements the balance sheet and income statement by providing additional information concerning an organization's ability to maximize the value of its equity. The purpose of this article is to provide an overview of selected cash flow ratios and an indication of their importance for proper diagnosis level of financial condition.

Keywords: financial analysis, cash flow ratios, statement of cash flows

ZASTOSOWANIE MODELU POTENCJAŁU W ANALIZIE ZRÓŻNICOWANIA REGIONALNEGO ROLNICTWA W KRAJACH UNII EUROPEJSKIEJ

Agnieszka Tłuczak

Zakład Ekonometrii i Metod Ilościowych, Wydział Ekonomiczny
Uniwersytet Opolski
e-mail: atluczak@uni.opole.pl

Streszczenie: Modele potencjału stosowane są w analizach poziomu rozwoju regionalnego, w przypadku których prowadzi się analizę korelacyjną, pozwalającą na zbadanie zależności między potencjałem a innymi zjawiskami społeczno-gospodarczymi. Potencjał określa intensywność oddziaływania między regionami nie tylko jako zmienną zależną od wielkości regionów, ale również od ich lokalizacji. Korzystne usytuowanie kraju w regionalnym systemie oddziaływań może wpłynąć na wzrost jego małego potencjału. W pracy zostaną zaprezentowane potencjał dochodu i potencjał ludności. Dutton w 1970 roku wprowadzając do badań iloraz potencjałów, złożył że potencjał dochodu jest proporcjonalny do popytu nominalnego, a potencjał ludności jest wskaźnikiem popytu rzeczywistego. Jednocześnie określił iloraz tych potencjałów jako miarę możliwości zaspokojenia popytu. Iloraz potencjału uwzględnia wpływ relacji międzyregionalnych na kształtowanie się poziomu badanego zjawiska, jest miarą systemową oraz zmienną o ciągłym rozkładzie przestrzennym. Rozkład przestrzenny ilorazu potencjałów ujmowanego jako miara poziomu rozwoju, jest podstawą wyróżnienia w strukturze regionalnej regionów rdzeniowych i obszarów peryferyjnych.

Celem artykułu jest zbadanie zróżnicowania regionalnego rolnictwa Unii Europejskiej oraz wyodrębnienie na podstawie zastosowanego modelu potencjału regionów rdzeniowych oraz peryferyjnych. Badania zostały przeprowadzone na poziomie krajów członkowskich UE. Uznano, że w ramach poszczególnych krajów nie zachodzi znaczące zróżnicowanie poziomu rozwoju rolnictwa i przeprowadzanie analiz na poziomie jednostek niższego rzędu nie będzie przedmiotem rozważań niniejszego opracowania.

Słowa kluczowe: model potencjału, rolnictwo, Unia Europejska

WSTĘP

Rolnictwo oraz obszary wiejskie Unii Europejskiej charakteryzują się dość dużym zróżnicowaniem regionalnym. Zróżnicowanie to w dużej części uzależnione jest od warunków przyrodniczych, ale coraz częściej również pozaprzyrodniczych. Występujące różnice w przestrzennym poziomie rozwoju rolnictwa oddziałują na potencjał wytwórczy, efektywność oraz możliwość generowania dochodów. Produkcja rolna odgrywa ważną rolę w gospodarce europejskiej, stanowi ona 10% całego PKB Unii Europejskiej. Warto zauważyć dodatkowo, że obszary wiejskie to ponad 91% terytorium Unii Europejskiej, a obszary te zamieszkuje ponad 56% obywateli. Rolnictwo pozostaje ważną gałęzią gospodarki każdego kraju, zapewnia zdrową żywność dla ludności oraz jest kluczem do europejskiej suwerenności żywnościowej. Duże znaczenie rolnictwa w gospodarkach krajów unijnych powoduje, że każdorazowo w budżecie unijnym na realizację działań i instrumentów WPR zarezerwowane ok. 10% wszystkich środków. Wydaje się, że do wyrównania występujących dysproporcji konieczne jest zintegrowanie polityki rozwoju wsi realizowanej w poszczególnych krajach UE ze Wspólną Polityką Rolną [Stanny 2009]. Zasadne zatem staje się rozpoznanie stopnia i struktury występującego zróżnicowania [Krasowicz 2009].

W literaturze przedmiotu coraz częściej do analiz poziomu rozwoju regionalnego stosuje się modele potencjału. W przypadku tych modeli prowadzi się analizę korelacyjną, która pozwala na zbadanie zależności pomiędzy potencjałem a innymi zjawiskami społeczno-gospodarczymi. W analizach regionalnych potencjał to miara oddziaływania regionów wchodzących w skład danego systemu. Potencjał określa intensywność oddziaływania między regionami nie tylko jako zmienną zależną od wielkości regionów, ale również od ich lokalizacji [Czyż 2002(a); Czyż 2002(b)]. Korzystne usytuowanie regionu w regionalnym systemie oddziaływać może wpłynąć na wzrost jego małego potencjału.

Praca prezentuje wyniki badań nad regionalnym zróżnicowaniem rolnictwa i obszarów wiejskich Unii Europejskiej przy użyciu potencjału dochodu i potencjału ludności. Potencjał dochodu regionalnego uznawany jest za miarę dostępności działalności gospodarczej [Keeble i in. 1982], zaś potencjał ludności określa różnice w dostępności regionów w skali Europy [Vickerman i in. 1999]. W roku 1970 roku Dutton wprowadził do badań iloraz potencjałów. Zakłada on, że potencjał dochodu jest proporcjonalny do popytu nominalnego, a potencjał ludności jest wskaźnikiem popytu rzeczywistego i określa iloraz tych potencjałów jako miarę możliwości zaspokojenia popytu oraz jako wskaźnik dochodu na osobę [Czyż 2012]. Rozkład przestrzenny ilorazu potencjałów ujmowanego jako miara poziomu rozwoju, jest podstawą wyróżnienia w strukturze regionalnej Unii Europejskiej regionów rdzeniowych i obszarów peryferyjnych [Czyż 2002(b)].

MODELE POTENCJAŁU W BADANIACH REGIONALNYCH

Twórcą koncepcji modelu potencjału jest amerykański astrofizyk J. Q. Stewart, którego badania koncentrowały się na badaniu rozmieszczenia ludności (potencjał ludności), wykorzystując do tego celu szereg praw zapożyczonych z fizyki [Pawlik 2014; Chojnicki 1964; Rosik 2012, s. 25]. Istotą modelu potencjału jest założenie, że wzajemne oddziaływanie dwóch obszarów jest wprost proporcjonalne do ich masy mierzonej np. wielkością produkcji przemysłowej, potencjałem gospodarki, poziomem innowacyjności, liczbą ludności, itp., a odwrotnie proporcjonalne do dzielącej ich odległości. Pojęcie potencjału przestrzennego, opiera się na regule potencjału cząstkowego E_{ij} [Pawlik 2014; Lewandowska-Gwarda, Antczak 2010, s. 230-235], który definiowany jest jako wielkość energii między dwoma obiektami i oraz j przypadającą na jednostkę masy P_j [Lewandowska-Gwarda, Antczak 2010, s. 230-235]:

$$E_{ij} = \alpha_0 \frac{P_i P_j}{d_{ij}} \quad (1)$$

gdzie: α_0 – parametr strukturalny modelu, P_i – masa i -tej jednostki, d_{ij} – odległość między jednostką i oraz j .

W analizie systemu regionalnego potencjał jest interpretowany jako miara oddziaływania regionów wchodzących w skład rozpatrywanego systemu. Region może mieć mały potencjał własny, ale dzięki korzystnemu usytuowaniu w regionalnym systemie oddziaływań jego potencjał ulega zwiększeniu [Czyż 2002(a); Czyż 2002(b); Werner 2003, s. 17-23; Kowalski Ł., Łobodzińska A. 2013].

Modele potencjału stosowane są szeroko w badaniach poziomu rozwoju społeczno-gospodarczego. Prekursorami analiz przestrzennych w Europie byli Clark [1969], Dicken i Lloyd [1977] oraz Keeble [1982]. Współcześnie tematyką tą zajmują się Roos [2001], Overman [2003], Mion [2004], Redding i Venables [2004], Hanson [2005], oraz Brakman [2009]. Zastosowali oni modele potencjału do analiz identyfikacji obszarów centralnych i peryferyjnych Europejskiej Wspólnoty Gospodarczej oraz zbadania zmian jego wartości w wyniku postępującej integracji ekonomicznej powodowanej redukcją barier handlowych. Z analiz tych wynikało, że głównymi beneficjentami integracji europejskiej były wówczas najbardziej rozwinięte i najgęściej zaludnione regiony [Cieślak, Rokicki 2013].

Ogólny zapis modelu potencjału przedstawia się następująco:

$$V_i = \alpha_0 \sum_{\substack{j=1 \\ j \neq i}}^n \frac{M_j^{\alpha_{1,j}} P_j^{\alpha_{2,j}}}{d_{ij}^\beta}, \quad i=1, 2, \dots, n. \quad (2)$$

gdzie V_i – potencjał i -tej jednostki, P_i – masa i -tej jednostki, M_i – waga masy i -tej jednostki, d_{ij} – odległość między jednostką i oraz j , α_0 , $\alpha_{1,j}$, $\alpha_{2,j}$ – parametry strukturalne modelu, β – parametr „oporu” odległości.

W analizach regionalnego zróżnicowania z zastosowaniem modeli potencjału ważnym zagadnieniem jest właściwe określenie odległości. W badaniach przestrzenno-ekonomicznych przyjmuje się, że odległość rozumiana jest jako odległość fizyczna (euklidesowa), fizyczna rzeczywista (np. odległość drogowa), czasowa (czas podróży/przewozu) lub ekonomiczna (koszt podróży/przewozu) między dwoma regionami [Rosik 2012, s. 25].

W literaturze przedmiotu wyróżnia się trzy podstawowe przestrzenne modele potencjału [Lewandowska-Gwarda, Antczak 2010, s. 230-235; Czyż 2002(a), s. 13-24; Czyż 2002(b), s. 5-7]:

- model potencjału ludności – określa dostępność danego, jest on miarą dostępności do tej jednostki z innych lokalizacji danego systemu,
- model potencjału dochodu – jest funkcją dochodu wytworzonego w danym regionie oraz dochodów w innych regionach i dzielących je odległości.
- model potencjału lokalizacji.

Dutton w 1970 roku do badań geograficznych wprowadził pojęcie ilorazu potencjałów, zakładał on, że potencjał dochodu jest proporcjonalny do popytu, a potencjał ludności jest wskaźnikiem popytu rzeczywistego [Lewandowska-Gwarda, Antczak 2010, s. 230-235]. Zastosowanie ilorazu potencjałów w badaniach empirycznych przedstawił Coffey w 1978 roku, określając iloraz potencjałów jako miarę możliwości zaspokojenia popytu [Coffey 1978]. Stosowanie ilorazu potencjałów w badaniu regionalnego zróżnicowania obejmuje kilka etapów. W pierwszy etapie oblicza się wartości potencjałów ludności V_i oraz dochodu U_i według formuł:

$$U_i = \frac{z_i}{d_{ii}} + \sum_{j=1}^n \frac{z_j}{d_{ij}}, \quad i=1, 2, \dots, n; \quad (3)$$

$$V_i = \frac{l_i}{d_{ii}} + \sum_{j=1}^n \frac{l_j}{d_{ij}}, \quad i=1, 2, \dots, n; \quad (4)$$

gdzie: z_i – dochody w i -tym regionie, d_{ij} – odległość między regionem i oraz j , l_i – ludność w i -tym regionie.

Porównanie ze sobą dwóch powierzchni potencjału pozwala na określenie obszarów o nadwyżce lub niedoborze dostępności (suma potencjałów badanego układu wynosi 100%). Daje to możliwość wyrażenia potencjału w punkcie i jako procentu sumy potencjału całego układu. Wartości potencjału od zera do jeden informują o przewadze dostępności w jednym układzie, a powyżej jedności w drugim [Guzik 2003, s. 33-40]. W kolejności oblicza się dla każdego regionu iloraz potencjałów P_i :

$$P_i = \frac{U_i}{V_i}. \quad (5)$$

W kolejnym etapie określa się relację pomiędzy ilorazem potencjałów P_i a wskaźnikiem dochodu regionalnego na mieszkańca g_i w postaci równania [Czyż 2002(a); Czyż 2002(b)]:

$$P_i = \left(1 + \frac{\sum_{j=1}^n z_j / d_j}{z_i} \right) / \left(1 + \frac{\sum_{j=1}^n l_j / d_j}{l_i} \right) g_i \quad (6)$$

gdzie: g_i – wskaźnik dochodu regionalnego na mieszkańca.

Wartość P_i/g_i w i -tym regionie zależy od:

- proporcji pomiędzy potencjałem generowanym przez otoczenie i -tego regionu a potencjałem własnym regionu i -tego w zakresie dochodu, a także pod względem liczby mieszkańców,
- relacji równości lub nierówności między tymi proporcjami.

W sytuacji, gdy w regionie występuje większy stopień koncentracji przestrzennej potencjału dochodu niż potencjału ludności zachodzi nierówność $P_i < g_i$. Oznacza to, że większa jest rola otoczenia regionalnego danego regionu i w generowaniu potencjału ludności tego regionu niż w generowaniu jego potencjału dochodu [Czyż 2002(a); Czyż 2002(b)].

WYNIKI

Postępowanie badacze w analizie regionalnego zróżnicowania rolnictwa w Unii Europejskiej rozpoczęto od obliczenia potencjałów ludności i potencjału dochodów dla każdego kraju. Obliczeń dokonano na podstawie danych Eurostatu. Do obliczenia potencjału ludności wykorzystano liczbę ludności aktywnej zawodowo w rolnictwie (w tysiącach osób), zaś do obliczenia potencjału dochodu wykorzystano przeciętny dochód w euro na 1 gospodarstwo rolne¹. Analiza obejmuje lata 2007 – 2013. Porównując potencjał ludności (Tabela 1) dla analizowanych jednostek przestrzennych – krajów UE, obserwujemy względną stabilną sytuację. Wartości potencjału ludności w kolejnych latach dla poszczególnych państw nie różnią się znacząco od siebie. Największym potencjałem ludności aktywnej zawodowo w rolnictwie charakteryzuje się Polska, dla której wartości potencjału ludności są prawie dwukrotnie większe niż na przykład w Hiszpanii. Do krajów z najwyższym potencjałem ludności zaliczyć należy również Rumunię i Włochy. Najmniejsze wartości potencjału ludności

¹ <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/> [dostęp 25.07.2014].

obserwujemy dla takich krajów jak Cypr i Luksemburg (ok. 0,5), należy jednak pamiętać, że są to kraje z rolnictwem o mało znaczącym wkładzie w rolnictwo UE.

Tabela 1. Potencjał ludności w krajach UE w latach 2007-2013

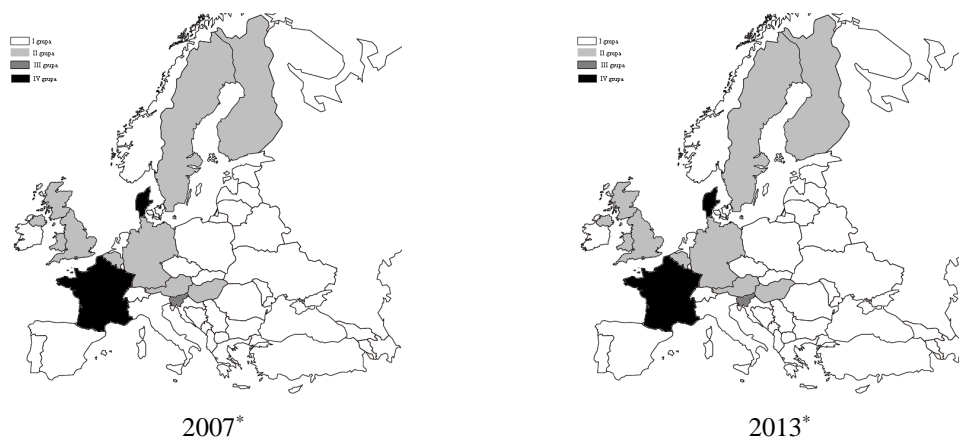
Kraj \ Rok	2007	2008	2009	2010	2011	2012	2013
Austria	2,05	2,06	2,05	2,05	2,04	2,04	2,04
Belgia	1,13	1,14	1,14	1,14	1,14	1,14	1,14
Bułgaria	1,66	1,63	1,58	1,55	1,46	1,46	1,41
Cypr	0,50	0,50	0,55	0,55	0,58	0,58	0,60
Czechy	3,28	3,30	3,31	3,33	3,36	3,36	3,39
Dania	1,36	1,37	1,37	1,37	1,37	1,37	1,37
Estonia	0,96	0,97	0,97	0,98	1,00	1,00	1,01
Finlandia	1,25	1,25	1,25	1,24	1,24	1,24	1,24
Francja	5,30	5,24	5,15	5,09	4,94	4,94	4,85
Grecja	5,35	5,37	5,43	5,45	5,51	5,51	5,55
Hiszpania	8,44	8,43	8,49	8,46	8,43	8,43	8,43
Holandia	2,24	2,25	2,25	2,26	2,29	2,29	2,30
Irlandia	1,50	1,53	1,52	1,53	1,56	1,56	1,58
Litwa	1,74	1,74	1,73	1,73	1,72	1,72	1,71
Luksemburg	0,62	0,62	0,62	0,62	0,62	0,62	0,62
Łotwa	1,43	1,44	1,46	1,48	1,51	1,51	1,54
Niemcy	6,22	6,17	6,11	6,06	5,94	5,94	5,87
Polska	13,63	13,82	13,98	14,19	14,61	14,61	14,85
Portugalia	4,51	4,53	4,49	4,49	4,54	4,54	4,55
Rumunia	7,90	7,68	7,53	7,36	6,98	6,98	6,76
Słowacja	2,38	2,40	2,43	2,45	2,49	2,49	2,52
Słowenia	0,67	0,67	0,66	0,66	0,65	0,65	0,64
Szwecja	1,25	1,26	1,27	1,28	1,30	1,30	1,31
Węgry	3,29	3,28	3,27	3,26	3,23	3,23	3,22
Wielka Brytania	4,04	4,12	4,18	4,27	4,45	4,45	4,55
Włochy	7,30	7,25	7,20	7,15	7,03	7,03	6,97

Źródło: opracowanie własne

W przypadku potencjału dochodu zaobserwować można zmiany wartości pomiędzy 2007 rokiem a 2013 rokiem. W grupie I (kraje o najniższych wartościach potencjału dochodu) odnotowano spadek wartości potencjału przeciętnie o 0,14,

zmienił się też skład tej grupy². Najkorzystniej zmieniła się sytuacja w Bułgarii i na Litwie, w 2007 oba kraje należały do pierwszej grupy pod względem wartości potencjału dochodu, zaś w roku 2013 do drugiej (Rysunek 1). Model przestrzennego potencjału dochodu wskazuje na silną koncentrację potencjału dochodu we Francji i w Danii. Najniższym natężeniem potencjału dochodu charakteryzują się kraje Europy Środkowej (m.in. Polska, Czechy, Słowacja, Węgry) oraz Europy Południowo-Zachodniej (m.in. Hiszpania, Portugalia i Włochy). Taki rozkład potencjału dochodu pokrywa się praktycznie z poziomem rozwoju rolnictwa, w krajach o których można mówić, że rolnictwo jest na wysokim poziomie rozwoju widoczne są wyższe wartości potencjału dochodu aniżeli w pozostałych przypadkach.

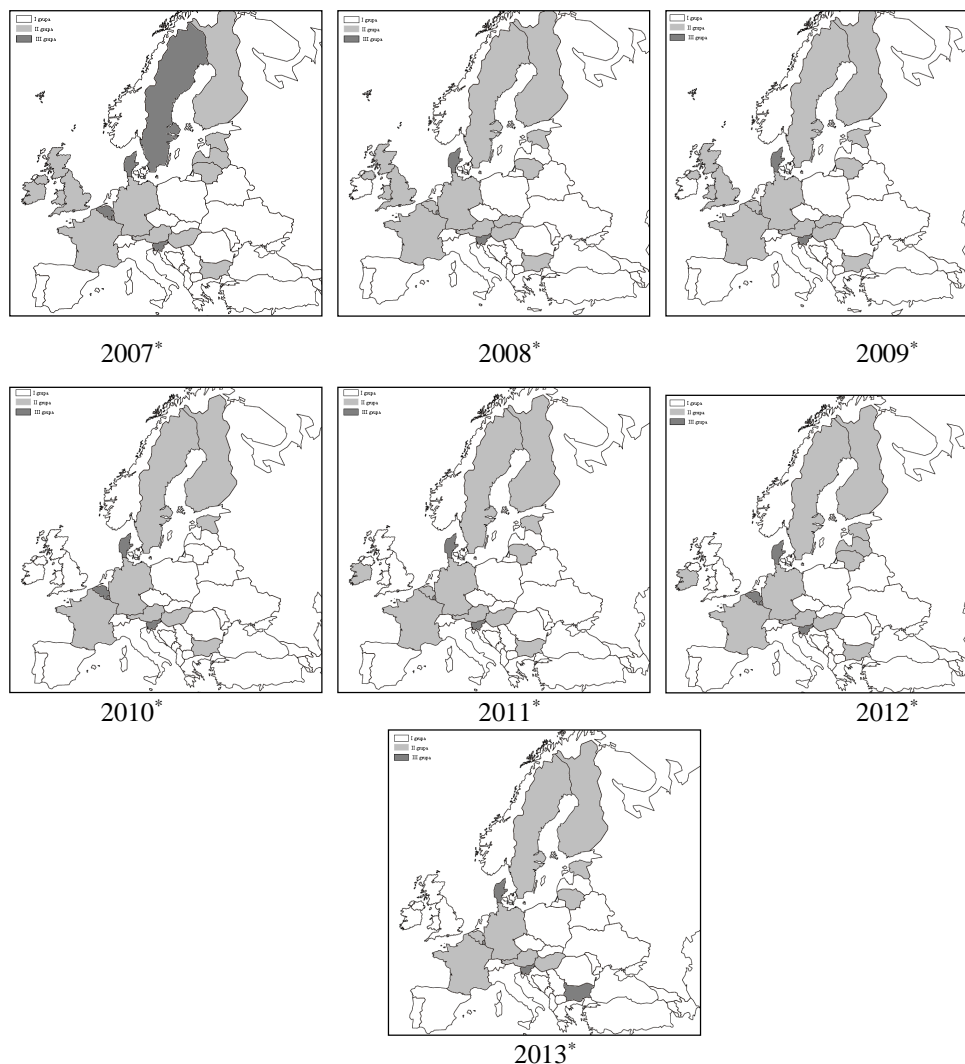
Rysunek 1. Potencjał dochodu w roku 2007 i 2013



Źródło: opracowanie własne

² Podział na grupy odbył się wedle zasady: grupa I - obejmuje kraje znajdujące się pomiędzy najmniejszą wartością różnic a połową średniej arytmetycznej obliczoną z najmniejszej i największej wartości różnic odległości; grupa II - kraje znajdujące się w przedziale od połowy średniej do średniej z przedziału; grupa III - kraje znajdujące się w przedziale od średniej do trzech drugich średniej; grupa IV - obejmuje kraje, których wartości znajdują się powyżej trzech drugich średniej. Szerzej patrz: Kauf S., Tłuczak A., Logistyka miasta i regionu. Metody ilościowe w badaniach przestrzennych, Difin Warszawa 2014.

Rysunek 2. Ilorazy potencjałów dochodu i ludności w latach 2007-2013



*grupa I: $P_I < 1$, grupa II: $1 < P_I < 4$, grupa III: $P_I > 4$.

Źródło: opracowanie własne

Rozkład przestrzenny ilorazu potencjałów ujmowanego jako miara poziomu rozwoju, jest podstawą wyróżnienia w strukturze regionalnej Europy, na gruncie koncepcji rdzeń-peryferie, regionów rdzeniowych i obszarów peryferyjnych. Na mapie rozkładu ilorazu potencjałów (Rysunek 2) układy ciągłe subregionów z wysokimi wartościami tego ilorazu odpowiadają układom oddziaływania w postaci regionów rdzeniowych. Rozkład wartości ilorazu potencjałów na mapie Europy ma charakter wielobiegunowy. Wyróżnia się dwa bieguny: Dania,

Słowenia, występuje tu najwyższa wartość ilorazu potencjałów w skali całej Unii Europejskiej. Kraje te wyróżniają się również najwyższymi wartościami potencjału dochodu (Rysunek 1). Strefa oddziaływania obu krajów obejmuje swym zasięgiem regiony sąsiednie: Niemcy oraz Szwecję. Obszary o niższych wartościach ilorazu potencjałów wyraźnie kontrastują z regionem obu krajów. Kraje odstające to przede wszystkim Polska, Portugalia i Rumunia, każdorazowo zajmują one najniższe pozycje pod względem wartości ilorazu potencjałów (poniżej 0,5).

W systemie Unii Europejskiej najwięcej państw dla których zachodzi związek $P_i > g_i$ odnotowano w latach 2008, 2009, 2011, 2012 i 2013, są to Dania, Szwecja, Bułgaria, Słowenia, Finlandia, Luksemburg, Francja, Litwa. Oznacza to, że kraje te wykazują się niższym stopniem koncentracji przestrzennej potencjału dochodu niż potencjału ludności. Dodatkowo oznacza to, że mniejsza jest rola otoczenia wymienionych krajów w generowaniu potencjału ludności niż w generowaniu potencjału dochodu krajów z tej grupy. Najwięcej krajów dla których zachodzi relacja $P_i > g_i$ odnotowano w roku 2009.

PODSUMOWANIE

Modele potencjału są stosowane w przestrzennych analizach społeczno-gospodarczych. Wykorzystywane są w badaniach dotyczących np. lokalizacji przemysłu, rozwoju ekonomicznego, handlu międzynarodowego, ochrony środowiska. Zaletą tych modeli jest możliwość zbadania wzajemnych powiązań pomiędzy poszczególnymi jednostkami terytorialnymi. Modele te, poza potencjałem własnym, uwzględniają wpływ oddziaływań sąsiadujących obiektów oraz położenie w przestrzeni.

Celem artykułu było zbadanie zróżnicowania rolnictwa Unii Europejskiej, wyodrębnienie regionów rdzeniowych oraz peryferyjnych na podstawie zastosowanego modelu potencjału. Przeprowadzone badania potwierdzają hipotezę o zróżnicowaniu rolnictwa na poziomie krajów członkowskich Unii Europejskiej w latach 2007-2013. Należy pamiętać, że badania nad regionalnym zróżnicowaniem rozwoju rolnictwa są badaniami wielowątkowymi i powinny uwzględniać wiele czynników, zarówno ekonomicznych jak i uwarunkowań przyrodniczych. Analizując jednak potwierdza się teza o niższym poziomie rozwoju rolnictwa w krajach tzw. byłego bloku wschodniego (Polska, Czechy, Słowacja, Węgry). Na tle tychże krajów i pozostałych krajów członkowskich można zaobserwować duże różnice w poziomie rozwoju rolnictwa. W krajach tych przez długie lata rolnictwo było niedofinansowane, co powodowało ograniczone możliwości korzystania ze środków chemicznych wspomagających uprawy. Rolnictwa wspomnianych krajów charakteryzowały się wysokimi nakładami pracy. Członkostwo krajów w strukturach Unii Europejskiej oraz objęcie wszystkich krajów jednakowymi zasadami Wspólnej Polityki Rolnej może w niedalekiej przyszłości doprowadzić do zmniejszenia, a może nawet zniwelowania, istniejących różnic.

BIBLIOGRAFIA

- Brakman S., Garretsen H., Marrewijk C. V. (2009), Economic Geography within and between European Nations: The Role of Market Potential and Density across Space and Time, "Journal of Regional Science", vol. 49, p. 777–800.
- Cieślik A., Rokicki B. (2013), Rola sieci transportowej w rozwoju polskich regionów: zastosowanie modelu potencjału ekonomicznego, "AUC Ekonomia", vol. XLIV nr 1 (13), s. 113-126.
- Clark C., Wilson F., Bradley J. (1969), Industrial Location and Economic Potential in Western Europe, "Regional Studies", No. 3, p. 197–212.
- Coffey W. (1978), Income Relationships in Boston and Toronto: A Tale for Two Countries?, "Canadian Geographer", No. 2(22), p. 112-129.
- Czyż T. (2002b), Zastosowanie modelu potencjału w analizie zróżnicowania regionalnego Polski, „Studia Regionalne i Lokalne”, Nr 2–3/2002, s. 5-7.
- Czyż T. (2002a), Application of the potential model to the analysis of regional differences in Poland, "Geographia Polonica", Nr 75(1), s. 13-24.
- Czyż T. (2012), Poziom rozwoju społeczno-gospodarczego Polski w ujęciu subregionalnym, „Przegląd Geograficzny”, Nr 84, s. 219-236.
- Dicken P., Lloyd P. (1977), Location in Space, Harper and Row, New York.
- Dutton G. (1970), Macroscopic Aspects of Metropolitan Evolution, "Harvard Papers in Theoretical Geography", Geography of Income Series", No 1.
- Guzik R. (2003), Przestrzenna dostępność szkolnictwa ponadpodstawowego, Instytut Geografii i Gospodarki Przestrzennej UJ Kraków.
- Hanson G. (2005), Market Potential, Increasing Returns and Geographic Concentration, „Journal of International Economics”, Vol. 67, p. 1–24.
- Keeble D., Owens P.L., Thompson Ch. (1982), Regional Accessibility and Economic Potential in The European Community, "Regional Studies", No. 6(16), p. 419-432.
- Kowalski Ł., Łobodzińska A. (2013), Modelowanie lokalizacji miejsc w przedszkolach w Małopolsce na podstawie dostępności przestrzennej, „AUL Folia Geographica Socio-Oeconomica”, No. 14, s. 93-106.
- Krasowicz S. (2009), Regionalne zróżnicowanie polskiego rolnictwa a możliwości wdrażania polityk wspólnotowych, „Zeszyty Naukowe SGGW, Polityki Europejskie, Finanse i Marketing”, Nr 1(50), s. 21-31.
- Mion G. (2004), Spatial Externalities and Empirical Analysis: The Case of Italy, Journal of Urban Economics, Vol. 56, p. 98 – 117.
- Overman H. G., Redding S., Venables A. (2003), Economic Geography: A Survey of Empirics, Handbook of International Trade, ed. Choi E. K., Harrigan J., Malden, Blackwell.
- Pawlik A., Badanie innowacyjności w skali regionalnej, <http://www.pte.pl/kongres/referaty/Pawlik%20Andrzej/Pawlik%20Andrzej%20-%20BADANIE%20INNOWACY%20JNO%20C5%9ACI%20W%20SKALI%20REGIONALNEJ.pdf>, dostęp 28.06.2014.
- Redding S., Venables A. J. (2004), Economic Geography and International Inequality, "Journal of International Economics", No. 62, p. 53–82.

- Rosik P.(2012), Dostępność przestrzeni Polski w wymiarze europejskim, PAN IG i PZ Warszawa
- Roos M. (2001), Wages and Market Potential in Germany, Jahrbuch für Regionalwissenschaft, No. 21, p. 171–195.
- Stanny M. (2009), Zróżnicowanie poziomu rozwoju obszarów wiejskich w Polsce a problem realizacji polityki spójności, Zeszyty Naukowe SGGW, Polityki Europejskie, Finanse i Marketing, nr 1(50), s. 47-56.
- Lewandowska-Gwarda, Antczak 2010.(2010), Modele potencjału, w Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych, red. nauk. B. Suchecki, C.H. Beck Warszawa.
- Vickerman, R., Spiekermann, K., Wegener, M. (1999), Accessibility and economic development in Europe, Regional Studies, No. 33(1), p. 1-15.
- Werner P. (2003), Geograficzne uwarunkowania rozwoju infrastruktury społeczeństwa informacyjnego w Polsce, WGiSR UW Warszawa.

THE APPLICATION OF POTENTIAL MODEL IN ANALYSIS OF REGIONAL DIVERSIFICATION OF AGRICULTURE IN EUROPEAN UNION

Abstract: In the analysis of the level of regional development the potential models are used, where the correlation analysis is conducted, and it allows to examine the relationship between the potential and the other socio-economic phenomena. The potential it is the intensity of the interaction between the regions not only as a variable dependent on the size of the regions, but also on their location. The preferred location of the country in the regional system of interactions can affect the growth of its small capacity. In this paper the model of income potential and the model of population potential. In 1970 Dutton introduced to the research the potential quotient, it assumed that the income potential is proportional to the nominal demand and the potential of the population is an indicator of real demand and determines the ratio of these potentials as far as possible to meet the demand. The quotient takes into account potential impact of interregional relations on the level of the studied phenomenon. It is a measure of the system and a variable with a continuous spatial distribution. The spatial distribution of the quotient of the recognized potential as a measure of the level of development, is the basis of distinctions in the regional structure of the core regions and remote areas.

The aim of this study is to examine the regional diversity of agriculture in the European Union and to extract the core and outlying regions on the basis of the potential model. The research was conducted at the level of EU member states. It was assumed that within each country there is no significant differences in the level of agricultural development and analysis at the level of the downstream units will not be considered by this study.

Keywords: potential model, agriculture, European Union

OCENA ZRÓŻNICOWANIA PRODUKCJI MLECZNEJ W UKRAINIE Z ZASTOSOWANIEM WIELOWYMIAROWYCH METOD STATYSTYCZNYCH

Maria Parlińska
Łukasz Pietrych
Iryna Petrovska

Katedra Ekonomiki Rolnictwa i Międzynarodowych Stosunków Gospodarczych
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: maria_parlinska@sggw.pl

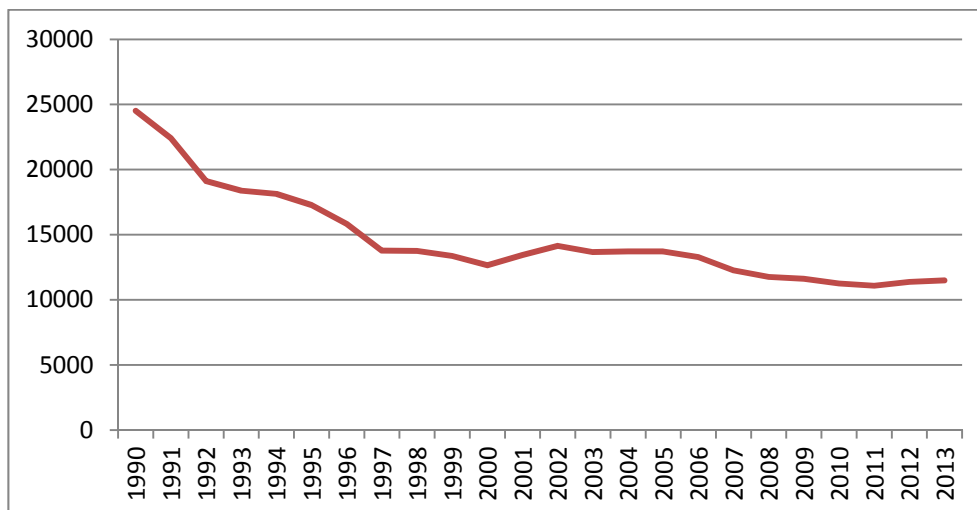
Streszczenie: W artykule przedstawiono przykład zastosowania wielowymiarowych metod statystycznych do oceny możliwości produkcji mleka w poszczególnych regionach Ukrainy. Dokonując analizy zmienności oraz macierzy współczynników korelacji do ostatecznych obliczeń wybrano cztery zmienne charakteryzujące omawiany problem. W pracy zastosowano trzy metody umożliwiające wielowymiarową analizę statystyczną, tj.: metodę Hellwiga, metodę średniej arytmetycznej oraz diagram Czekanowskiego. Należy stwierdzić, że Ukraina jako kraj, który w przyszłości ma zamiar wstąpić do Unii Europejskiej może mieć istotny wpływ na rynek mleka, ze względu na swój potencjał produkcyjny.

Słowa kluczowe: produkcja mleka, wielowymiarowe metody statystyczne, Ukraina

DYNAMIKA PRODUKCJI MLEKA W UKRAINIE

Celem pracy jest zastosowanie wielowymiarowych metod statystycznych do budowy rankingu regionów Ukrainy pod względem potencjału produkcji mleka. Podnoszona tematyka jest istotna, ponieważ w ciągu ostatnich trzynastu lat można zauważyć systematyczny spadek produkcji tego surowca rolnego. W latach 90-tych zmniejszenie ilości produkcji mleka było spowodowane rozpadem Związku Radzieckiego. Do roku 2011 można zaobserwować wyraźny spadek produkcji i tylko w roku 2012 oraz 2013 można zauważyć niewielki przyrost.

Rysunek 1. Dynamika produkcji mleka na Ukrainie w tys. ton.



Źródło: dane do wykresu zaczerpnięte z Rocznika Statystycznego Regionów Ukrainy

Oczywiście poruszony problem wymaga zdecydowanie szerszej analizy. Należy zwrócić uwagę na specyficzne cechy rozpatrywanego rynku. Pierwszą z nich jest to, iż na Ukrainie produkcja mleka opiera się przede wszystkim na gospodarstwach chłopskich, w których to jest ona przeznaczana na wewnętrzne potrzeby. W gospodarstwach towarowych natomiast da się zauważyć zmniejszanie produkcji. Taki stan rzeczy jest ściśle związany z brakiem majątku niezbędnego do rozszerzania skali produkcji oraz trudnościami w podejmowaniu działań inwestycyjnych. Istotne znaczenie ma również niski poziom popytu wewnętrznego oraz to, że Ukraina jest eksporterem netto produktów mlecznych, gdzie główny produkt eksportowy to sery [Sulyma, Parzonko 2009].

METODOLOGIA

Najistotniejszym elementem w procedurze wielowymiarowej analizy statystycznej jest wybór odpowiedniego zestawu zmiennych. W literaturze wymienia się następujące etapy: ustalenie pierwotnego zestawu zmiennych na podstawie kryteriów merytorycznych, analiza formalno – statystyczna oraz wybór optymalnego zestawu [Kisielińska, Stańko 2009]. Wyboru zmiennych opisujących postawiony cel badawczy dokonano metodą burzy mózgów, mając na uwadze takie czynniki, jak dostępność danych oraz celowość w stosunku do tematu pracy.

Wybrano następujące zmienne:

- X1 – produkcja mleka (w tys. litrów),
- X2 – powierzchnia użytków rolnych,

- X3 –średnia cena mleka,
- X4 –liczba firm z branży przetwórstwa rolnego.

Wszystkie dane dotyczyły roku 2013, a ich źródłem był Rocznik Statystyczny Regionów Ukrainy.

W pierwszym kroku obliczono współczynnik zmienności ze wzoru: $V_i = S_i / X_i$ oraz przyjmując progową wartość współczynnika równą 10% stwierdzono, że żadna ze zmiennych nie ma charakteru quasi - stałego [Kola – Bezka, 2012].

Wielowymiarowe metody porządkowania obiektów można podzielić na dwie zasadnicze grupy: metody porządkowania liniowego oraz nieliniowego. Metody porządkowania liniowego pozwalają na uporządkowanie analizowanych obiektów w taki sposób, że możliwe jest ich przeniesienie na linię prostą z uwzględnieniem hierarchii, czyli od obiektów stojących najwyżej, do tych zajmujących najniższe pozycje. Można tutaj wymienić takie metody jak: metody diagramowe (Czekanowskiego), metody oparte na zmiennej syntetycznej (wzorcowe i bez wzorcowe) oraz metody iteracyjne. Na potrzeby pracy wykorzystano miarę Hellwiga, metodę średniej arytmetycznej oraz diagram Czekanowskiego.

Druga grupa metod, to metody porządkowania nieliniowego. W wyniku ich zastosowania możliwe jest rzutowanie analizowanych obiektów na płaszczyznę, w celu porównywania podobieństwa obiektów, bez dokonywania ich hierarchizacji. Można wymienić tutaj takie metody jak: metody dendrytowe (taksonomia, metoda Prima), metoda aglomeracyjna, czy też metoda najbliższego sąsiedztwa [Petrovska, Pietrych 2013].

WYNIKI BADAŃ

Dane po standaryzacji poddano dalszym analizom, których celem było wyznaczenie tzw. miary rozwoju Hellwiga oraz miary syntetycznej (w przypadku metody średniej arytmetycznej).

Tabela 1. Ranking regionów: metoda średniej arytmetycznej oraz miara Hellwiga

Region	Miara Hellwiga	Klasa	Region	Metoda średniej arytmetycznej	Klasa
Vinnytska oblast	0,440634813	1	Vinnytska oblast	1,031202497	1
Dnipropetrovska	0,372284813	1	Odeska	0,970614963	1
Odeska	0,355810699	1	Dnipropetrovska	0,698245302	1
Charkivska	0,343063871	1	Poltavska	0,643519412	1
Poltavska	0,325721423	1	Charkivska	0,562196443	1
Kyivska	0,29095697	2	Mykolaivska	0,350761317	2
Mykolaivska	0,288073354	2	Zakarpatska	0,232008205	2
Republika Krym	0,262998827	2	Zaporizka	0,186505057	2
Chersonska	0,253202038	2	Kyivska	0,127790633	2
Zakarpatska	0,248562195	2	Chersonska	0,090008006	2
Cherkaska	0,247338943	2	Republika Krym	0,075003781	2
Donetska	0,239306184	2	Donetska	0,027222985	2
Zaporizka	0,237175033	2	Chmelnytska	-0,0175586	2
Lwiwska	0,218270083	2	Cherkaska	-0,041907405	2
Chmelnytska	0,217500158	2	Luganska	-0,106836075	2
Luganska	0,202922157	2	Chernigiwska	-0,107901784	2
Zhytomyrska	0,185247216	2	Lwiwska	-0,108750479	2
Ternopilska	0,168930653	2	Zhytomyrska	-0,163764483	2
Chernigiwska	0,161218148	2	Kirovogradska	-0,171085624	2
Kirovogradska	0,154251936	2	Ternopilska	-0,371558902	2
Sumska	0,08182921	3	Sumska	-0,540177388	3
Rivnenska	0,071473125	3	Rivnenska	-0,721738647	3
Wolynska	0,055540383	3	Wolynska	-0,75444516	3
Ivano-Frankiwska	0,031689014	3	Ivano-Frankiwska	-0,886159073	3
Chernivetska	0,0187187	3	Chernivetska	-1,00319498	3

Źródło: obliczenia własne

W celu porównania zgodności otrzymanych wyników obliczono współczynnik rang Spearmana: $r_s=0,96$. Zbudowano również następujące przedziały klasowe:

- Klasa 1 (wysoki potencjał): $d_i \geq \bar{d}_i + S_{di}$;
- Klasa 2 (średni potencjał): $\bar{d}_i - S_{di} \leq d_i < \bar{d}_i + S_{di}$;
- Klasa 3 (niski potencjał): $d_i < \bar{d}_i - S_{di}$;

gdzie:

d_i – wartość mierników rozwoju i syntetycznego d_i ,

\bar{d}_i – średnia arytmetyczna cechy d_i ,

S_{di} - odchylenie standardowe cechy d_i .

Należy stwierdzić, że obie metody dały zbliżone wyniki, zwłaszcza pozycje ostatnie rankingu – dotyczące regionów: Sumska, Rivnenska, Wołyńska, Ivano-Frankiwska, Chernivetska – są identyczne. W przypadku wyższych miejsc można zaobserwować pewne rozbieżności, natomiast pierwsze miejsce zajmuje region Vinnytska oblast. Jeszcze większą zgodność da się zauważyć dzięki zbudowanym przedziałom klasowym, gdyż wszystkie składają się z tych samych regionów.

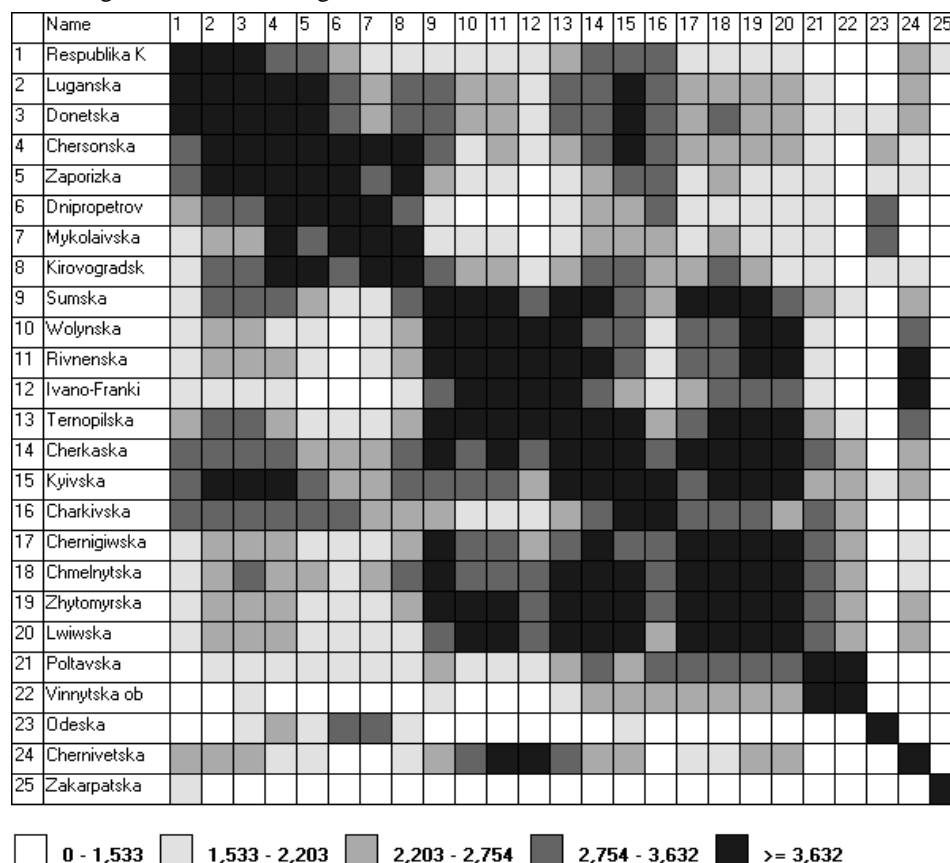
W kolejnym kroku zastosowano jedną z metod diagramowych porządkowania obiektów – diagram Czekanowskiego. Obliczeń dokonano z użyciem programu MaCzek. Procedura postępowania przedstawia się następująco:

- standaryzacja danych (metodą uwzględniającą odchylenia standardowe),
- wyznaczenie uporządkowanego diagramu.

Poszczególne wiersze i kolumny odpowiadają kolejnym regionom Ukrainy. Ciemniejsze pola na przecięciu wiersza i kolumny, odpowiadają regionom bardziej podobnym do siebie, pod względem badanych cech. Na zamieszczonym Rysunku 1 można dostrzec, że regiony tworzą grupy, które powstają z obiektów najbardziej podobnych.

Najbardziej podobne obiekty znajdują się najbliżej głównej przekątnej, im dalej od głównej przekątnej tym obiekty są mniej podobne do siebie pod względem badanych cech. Nie jest to jednak podział jednoznaczny, ponieważ występują obiekty, które można przypisać do kilku grup.

Rysunek 2. Diagram Czekanowskiego



Źródło: obliczenia własne

W tym przypadku można wyróżnić cztery grupy najbardziej podobnych regionów po względem potencjału produkcji mleka:

- Poltavska, Vinnytska oblast,
- Chernigiwska, Chmelnytska, Zhytomyrska, Lwiwska,
- Republika Krym, Luganska, Donetska, Chersonska, Zaporizka, Dnipropetrovska,
- Wolynska, Rivnenska, Ivano-Frankiwska, Ternopilska.

PODSUMOWANIE

Metody porządkowania liniowego pozwoliły na hierarchizację regionów Ukrainy ze względu na potencjał produkcji mleka. Dokonano tego pod kątem czterech zmiennych: produkcja mleka (w tys. litrów), powierzchnia użytków

rolnych, średnia cena mleka, liczba firm z branży przetwórstwa rolnego. Zastosowano metodę bezwzorcową (średniej arytmetycznej) oraz wzorcową (miarę Hellwiga), które dały zbliżone rezultaty. Diagram Czekanowskiego pozwolił na wyznaczenie liniowych skupisk regionów podobnych do siebie, jednakże należy pamiętać, że ze względu na subiektywizm w określaniu skali podobieństwa nie jest to metoda bardzo dobra.

Ukraina posiada zróżnicowany potencjał produkcji mleka. Sektor rolnictwa na Ukrainie wymaga, także przeprowadzenia szeregu zmian, przede wszystkim w kierunku koncentracji produkcji. Można stwierdzić, że przedstawione metody mogą służyć do wykazywania dysproporcji pomiędzy regionami, pod względem rozwoju w określonej dziedzinie, a to z kolei może służyć pobudzeniu aktywności władz w celu pomniejszania różnic.

BIBLIOGRAFIA

- Kisielińska J., Stańko S. (2009) Wielowymiarowa analiza danych w ekonomice rolnictwa. T. 96, z. 2, s. 63-76, Rocznik Nauk Rolniczych, SERIA G chyba najpierw powinien być tytuł czasopisma a potem nr tomu i strony
- Kola – Bezka M. (2012) Wielowymiarowa analiza porównawcza jako narzędzie zarządzania regionem na przykładzie województwa kujawsko – pomorskiego. Rok 16, Nr 2/2012, s. 51-64, Studia i Materiały, Miscellanea Oeconomicae
- Petrovska I., Pietrych Ł. (2013) Ocena zróżnicowania potencjału produkcji żywca wołowego w Polsce z zastosowaniem wielowymiarowych metod statystycznych. T. 15, z. 2, s. 289-292, Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu
- Rocznik statystyczny regionów Ukrainy: <http://ukrstat.org/uk>
- Sulyma N., Parzonko A. (2009) Stan i kierunki zmian na rynku mleka Ukrainy. T. 96, z. 1, s. 130-140, Rocznik Nauk Rolniczych, SERIA G.

EVALUATION OF MILK PRODUCTION DIVERSIFICATION IN UKRAINE WITH USING MULTIDIMENSIONAL STATISTICAL METHODS

Abstract: In the article it is presented example of using multivariate statistical methods in purpose to assess capacity and possibilities of milk production in Ukrainian regions. With use of variability analysis and matrix of correlation coefficients there were chosen explanatory variables. In the article there were presented the following multivariate statistical methods as Hellwig method, method of middle average and Czekanowski diagram. In the conclusions it was defined the regions with the biggest and the smallest production capabilities.

Keywords: milk production, multivariate statistical methods, Ukraine

MODEL DYNAMIKI RYNKU PRACY Z HETEROGENICZNĄ SIŁĄ ROBOCZĄ

Małgorzata Wrzosek

Instytut Ekonometrii, Szkoła Główna Handlowa
e-mail: malgorzata.wrzosek@sgh.waw.pl

Streszczenie: W pracy przedstawiono model dynamiki na rynku pracy w sytuacji, gdy pracownicy są heterogeniczni – różnią się produktywnością, która zależy od kwalifikacji pracownika. Prezentowany model pozwala badać przepływy na rynku pracy i ich zależność od takich czynników, jak występowanie i wysokość płacy minimalnej oraz struktura kwalifikacji pracowników. Zastosowanie układów dynamicznych do opisu zagadnienia umożliwia analizę istnienia, charakteru i stabilności równowagi na rynku pracy. Budowa modelu teoretycznego została uzupełniona symulacjami komputerowymi pozwalającymi na stawianie i weryfikowanie hipotez dotyczących dynamiki zatrudnienia i bezrobocia.

Słowa kluczowe: rynek pracy, modelowanie przepływów, układy dynamiczne

WSTĘP

Wysoki poziom bezrobocia w wielu krajach, szczególnie wśród młodych ludzi wchodzących na rynek pracy powoduje, że ekonomiści wciąż zajmują się analizą procesów zachodzących na tym rynku oraz identyfikowaniem czynników kształtujących popyt na siłę roboczą. W ostatnich latach bardzo często, jako ważny czynnik wpływający w sposób szczególny na rynek pracy, wskazywany jest znaczący postęp technologiczny i wynikające z niego zapotrzebowanie na wysoko wykwalifikowanych pracowników¹. Siła robocza postrzegana jest jako heterogeniczna i jest to cecha istotnie wpływająca na procesy na rynku pracy.

¹ Zjawisko to, szczególnie obserwowane i analizowane w USA oraz UK, występuje w opracowaniach pod nazwą SBTC, która jest skrótem od Skill-Biased Technological Change.

Postęp technologiczny, szczególnie w drugiej połowie XX i na początku XXI wieku jest istotnym motorem wzrostu gospodarczego. Jednak już od około 30 lat obserwowany jest także złożony wpływ postępu na rynek pracy. Z jednej strony – rosnące zapotrzebowanie na odpowiednie kwalifikacje pracowników spowodowało wzrost inwestycji w umiejętności i wykształcenie. Z drugiej strony – wysokie kwalifikacje były coraz wyżej cenione, co doprowadziło do znacznego rozwarstwienia płac. Właśnie postęp technologiczny jako przyczynę stale rosnących nierówności płacowych² wskazują [Card, DiNardo 2002] oraz [Acemoglu, Autor 2011], zaś [Acemoglu, Autor 2012] proponują inwestycje w kapitał ludzki jako sposób na redukcję tych nierówności.

Z punktu widzenia polityki rynku pracy kluczowe są dwie kwestie: jak wpływa wzrost kwalifikacji na poziom bezrobocia w grupie nisko i wysoko wykwalifikowanych pracowników oraz jak osiągnąć pożądany wzrost kwalifikacji. Analizy zależności wielkości bezrobocia od poziomu kwalifikacji siły roboczej wskazują na ogół, że wzrost poziomu kwalifikacji wpływa na ogólny spadek bezrobocia, w obydwu grupach pracowników – zarówno nisko, jak i wysoko wykwalifikowanych (por. [Moore, Ranjan 2005]). Natomiast [Weiss, Garloff 2011] proponują model, z którego wynika, że relatywny wzrost produktywności pracowników wysoko wykwalifikowanych w stosunku do nisko wykwalifikowanych powoduje wzrost bezrobocia.

Problemem dla ekonomistów pozostaje sposób oceny wysokości kwalifikacji. Jako naturalne przybliżenie miernika kwalifikacji [Katz, Margo 2008] przyjmują poziom edukacji społeczeństwa. Jednak często badania empiryczne wskazują, że edukacja i kwalifikacje nie są ekwiwalentami. Model ekonometryczny, który opisują [Altonji i in. 2012] wskazuje, że o rozbieżności płac w mniejszym stopniu decyduje poziom edukacji niż wybrany profil kształcenia.

Jako przyczyna wzrostu i utrzymywania się bezrobocia w grupie najmłodszych pracowników często wskazywane są występowanie i wysokość płacy minimalnej (por. [Meer, West 2013]). Artykuł ma na celu zaprezentowanie opracowanego przez autorkę modelu dynamiki rynku pracy uwzględniającego oddziaływanie na tę dynamikę obserwowanego poziomu kwalifikacji oraz płacy minimalnej. Wykorzystanie do tego celu układów dynamicznych wpisuje się w szeroki nurt prac analizujących rynek pracy, których autorzy zakładają, że najlepszą podstawą do zrozumienia zjawisk zachodzących na tym rynku jest wskazanie przyczyn i wielkości przepływów między grupami uczestników rynku (por. [Elsby i in. 2009],[Shimer 2012]).

² Wśród innych czynników mogących mieć wpływ na nierówności płacowe rozpatruje się także globalizację rynków [Moore, Ranjan 2005], jednak zjawisko to jest obserwowane znacznie krócej niż SBTC.

MODEL DYNAMIKI RYNKU PRACY

Opis modelu

Prezentowany model rynku pracy opisuje przepływy między trzema grupami osób: zatrudnieniem, bezrobociem i grupą nieaktywnych na rynku pracy. Stan rynku w okresie t charakteryzowany jest za pomocą punktu

$$r_t = (e_t, u_t, n_t),$$

w którym przyjęto następujące oznaczenia:

e_t - wielkość zatrudnienia w okresie t ,

u_t - wielkość bezrobocia w okresie t ,

n_t - wielkość grupy nieaktywnych na rynku pracy w okresie t ,

gdzie $t \in [t_0, T] \subset \mathbb{R}^+$ oraz $e_t, u_t, n_t \in \mathbb{R}^+$.

Zakłada się, że w każdym okresie występują przepływy w obu kierunkach między bezrobociem i zatrudnieniem oraz między bezrobociem i grupą nieaktywnych. Natomiast między zatrudnieniem i grupą nieaktywnych występują przepływy tylko w jednym kierunku – przyjmuje się, że część zatrudnionych rezygnuje z aktywności na rynku pracy i przechodzi do grupy nieaktywnych. Jednak nieaktywni, którzy chcą wrócić do aktywnego uczestnictwa w rynku pracy trafiają najpierw do bezrobocia i dopiero z tego stanu mogą przejść do zatrudnienia.

Siła robocza jest heterogeniczna, ponieważ pracownicy różnią się posiadanymi kwalifikacjami. Wśród pracowników zatrudnionych obserwuje się dwa poziomy kwalifikacji: wysoki i niski. Pracodawcy nie znają jednak kwalifikacji konkretnego pracownika przed zatrudnieniem go. Znają tylko strukturę kwalifikacji w grupie bezrobotnych. Struktura ta jest charakteryzowana parametrem $s \in (0,1)$, który oznacza udział osób wysoko wykwalifikowanych w bezrobociu. Przyjmuje się, że wielkość przepływu bezrobotnych w okresie t do zatrudnienia w okresie $t+1$ jest wprost proporcjonalna do struktury kwalifikacji s . Natomiast wielkość przepływu zatrudnionych w okresie t do bezrobocia w okresie $t+1$ jest proporcjonalna do frakcji nisko wykwalifikowanych pracowników, którzy zostali zatrudnieni w okresie poprzednim.

W modelu zakłada się, że przepływy między grupą nieaktywnych na rynku i pozostałymi grupami zależą od wielkości obowiązującej płacy minimalnej. Przepływ z grupy nieaktywnych do bezrobocia jest wprost proporcjonalny do płacy minimalnej, a wielkości przepływów z bezrobocia i zatrudnienia do grupy nieaktywnych są odwrotnie proporcjonalne do płacy minimalnej. Dla uproszczenia modelu przyjęto, że ostatnie dwa przepływy charakteryzowane są tym samym parametrem.

Matematyczny zapis modelu dynamiki rynku pracy

Opisane powyżej zależności na rynku pracy charakteryzuje układ równań różnicowych

$$\begin{cases} e_{t+1} = \left(1 - \gamma \frac{1}{w_m}\right) e_t + \alpha s u_t - \alpha s (1-s) u_{t-1}, \\ u_{t+1} = \left(1 - \alpha s - \gamma \frac{1}{w_m}\right) u_t + \alpha s (1-s) u_{t-1} + \beta w_m n_t, \\ n_{t+1} = \gamma \frac{1}{w_m} e_t + \gamma \frac{1}{w_m} u_t + (1 + \beta w_m) n_t. \end{cases} \quad (1)$$

W modelu (1) przyjęto następujące oznaczenia:

w_m – płaca minimalna,

α – parametr opisujący przepływ z bezrobocia do zatrudnienia, wielkość αs oznacza odsetek osób bezrobotnych w okresie t , które znajdują zatrudnienie w okresie $t+1$, α przyjmuje wartości z przedziału $(0,1)$,

β – parametr charakteryzujący wielkość przepływu z grupy nieaktywnych na rynku do bezrobocia, wskaźnik βw_m oznacza część nieaktywnych w okresie t , którzy wracają na rynek pracy w okresie $t+1$, β przyjmuje wartości z przedziału $(0, \frac{1}{w_m})$,

γ – parametr charakteryzujący przepływy z zatrudnienia oraz bezrobocia do grupy nieaktywnych na rynku, wielkość $\gamma \frac{1}{w_m}$ oznacza frakcję zatrudnionych oraz bezrobotnych w okresie t , którzy rezygnują z aktywności na rynku pracy w okresie $t+1$, γ przyjmuje wartości z przedziału $(0, w_m)$.

Zakłada się, że w każdym okresie $t \in [t_0, T]$ zachodzi równość

$$e_t + u_t + n_t = N = const.$$

Przy powyższym założeniu układ równań (1) jest równoważny następującemu układowi

$$\begin{cases} e_{t+1} = (1 - \varphi) e_t + \alpha s u_t - \alpha s (1-s) x_t, \\ u_{t+1} = -\varepsilon e_t + (1 - \alpha s - \varphi - \varepsilon) u_t + \alpha s (1-s) x_t + \varepsilon N, \\ x_{t+1} = u_t, \end{cases} \quad (2)$$

przy czym dla uproszczenia dalszej analizy przyjęto oznaczenia: $\varphi = \gamma \frac{1}{w_m}$, $\varepsilon = \beta w_m$. Parametry φ , ε spełniają warunek $\varphi, \varepsilon \in (0,1)$.

Dla przyjętych zakresów zmian wartości parametrów modelu (2) zawsze istnieje jeden punkt stały układu. Zatem rynek pracy, którego dynamikę opisuje układ równań (2) w długim okresie posiada równowagę, w której wielkości zatrudnienia i bezrobocia wynoszą odpowiednio

$$\begin{cases} \bar{e} = \frac{\varepsilon \alpha s^2 N}{\varphi^2 + \alpha s^2 (\varphi + \varepsilon) + \varphi \varepsilon}, \\ \bar{u} = \frac{\varepsilon \varphi N}{\varphi^2 + \alpha s^2 (\varphi + \varepsilon) + \varphi \varepsilon}. \end{cases} \quad (3)$$

Postać analityczna punktu równowagi wskazuje, że zarówno poziom zatrudnienia, jak i bezrobocia w równowadze zależą od struktury kwalifikacji i od płacy minimalnej.

WYNIKI SYMULACJI DYNAMIKI RYNKU PRACY

Dla zbudowanego modelu dynamiki rynku pracy zostały przeprowadzone symulacje komputerowe. Graficzna ilustracja zmian wielkości zatrudnienia i bezrobocia nie tylko pozwala obserwować wpływ poszczególnych parametrów modelu na charakter dynamiki rynku, ale umożliwia także stawianie i weryfikowanie hipotez dotyczących badanych zależności. Szczególnie istotna jest ocena wpływu struktury bezrobocia oraz płacy minimalnej na dynamikę rynku i stan równowagi. W celu oceny wpływu tych czynników na dynamikę zatrudnienia i bezrobocia wykonano serię symulacji dla różnych wartości badanego czynnika i ustalonych wartości pozostałych parametrów i punktów początkowych.

Rysunki poniżej przedstawiają różne warianty dynamiki zatrudnienia i bezrobocia w zależności od obserwowanej na rynku struktury kwalifikacji bezrobotnych, charakteryzowanej parametrem s . Wykresy pokazują 8 wariantów przebiegów wartości zmiennych dla parametru s przyjmującego kolejno wartości od 0,5 do 0,64 co 0,02. Punkt początkowy symulacji oraz wartości pozostałych parametrów są jednakowe dla wszystkich wykresów, a ich wartości prezentuje Tabela 1. Wartości te zostały dobrane tak, aby możliwe było odniesienie uzyskanych wyników do polskiego rynku pracy. Wielkości bezrobocia i zatrudnienia wyrażone są w tysiącach osób.

Tabela 1. Wartości początkowe zatrudnienia i bezrobocia oraz przyjęte parametry charakteryzujące rynek pracy, dla którego obserwowany jest wpływ struktury kwalifikacji bezrobotnych

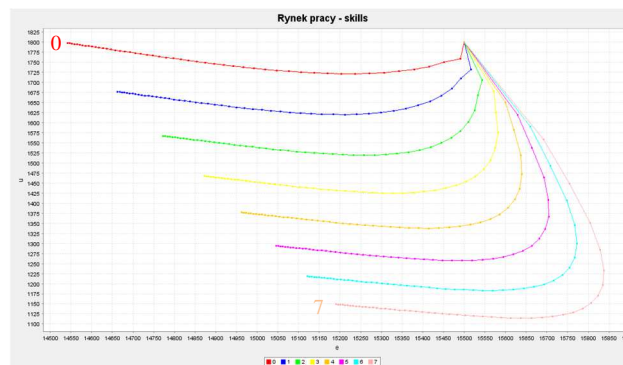
e_0	u_0	N	φ	ε	α
15500	1800	30000	0,022	0,026	0,7

Źródło: obliczenia własne

Rysunek 1 prezentuje trajektorie wielkości zatrudnienia i bezrobocia dla różnych poziomów struktury kwalifikacji bezrobotnych. Na osi poziomej zaznaczone zostały wielkości zatrudnienia, a na osi pionowej – bezrobocia, wyrażone w tysiącach osób. Kolejne trajektorie oznaczone liczbami od 0 do 7 odpowiadają coraz większym wartościom parametru s charakteryzującego strukturę kwalifikacji. Każdy z przedstawionych tu rynków stabilizuje się na innym

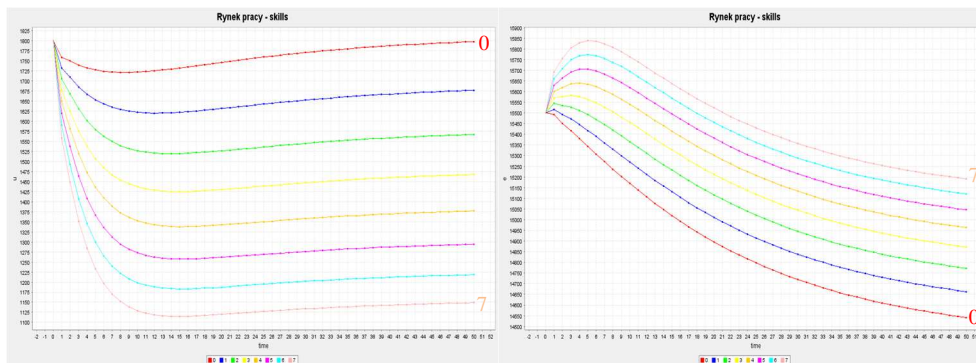
poziomie wielkości zatrudnienia i bezrobocia, przy czym im wyższy jest poziom kwalifikacji, tym wyższe jest zatrudnienie i niższe bezrobocie, do których zbiega rynek pracy. Zależność tę dokładniej ilustrują wykresy wielkości bezrobocia i zatrudnienia w kolejnych okresach, przedstawione na rysunku 2.

Rysunek 1. Trajektorie rynku pracy wyznaczone dla poziomów kwalifikacji bezrobotnych od 0,5 do 0,64



Źródło: obliczenia własne

Rysunek 2. Wielkości zatrudnienia i bezrobocia w kolejnych okresach wyznaczone dla poziomów kwalifikacji bezrobotnych od 0,5 do 0,64



Źródło: obliczenia własne

Zmiana wysokości płacy minimalnej wpływa na zmianę dwóch współczynników w modelu: φ oraz ε . Pierwszy maleje wraz ze wzrostem płacy minimalnej, drugi współczynnik jest wprost proporcjonalny do płacy minimalnej. Rysunki 3 i 4 ilustrują zmiany dynamiki rynku pracy w zależności od wysokości płacy minimalnej. Kolejne krzywe oznaczone numerami od 0 do 7 ilustrują zmiany dynamiki na rynku pracy przy wzroście parametru ε i spadku parametru φ , co odpowiada wzrostowi płacy minimalnej. Zmiany wartości parametrów φ oraz ε dla krzywych 0-7 można interpretować jako wzrost wysokości płacy minimalnej od

310 do 370 EUR³. Tabela 2 przedstawia wspólne dla wszystkich trajektorii wartości dla punktu początkowego oraz pozostałych parametrów modelu.

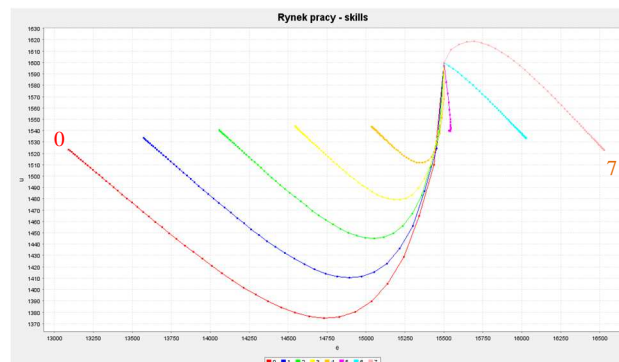
Tabela 2. Wartości początkowe zatrudnienia i bezrobocia oraz przyjęte parametry charakteryzujące rynek pracy, dla którego obserwowany jest wpływ wysokości płacy minimalnej

e_0	u_0	N	A	s
15500	1600	30000	0,7	0,6

Źródło: obliczenia własne

Uzyskane w wyniku symulacji trajektorie zatrudnienia i bezrobocia wskazują na wyraźny wzrost wielkości zatrudnienia wraz ze wzrostem płacy minimalnej. Natomiast wpływ płacy minimalnej na wielkość bezrobocia zmienia się w czasie – w krótkim okresie powoduje wzrost bezrobocia, ale dłuższym horyzoncie nie ma dużego wpływu na poziom bezrobocia. Rysunek 4 wskazuje, że krzywe ilustrujące dynamikę bezrobocia dla różnych poziomów płacy minimalnej nie tylko zbliżają się do siebie, ale mogą się także przecinać, co oznacza brak jednoznacznej zależności między wysokością płacy minimalnej i wielkością bezrobocia.

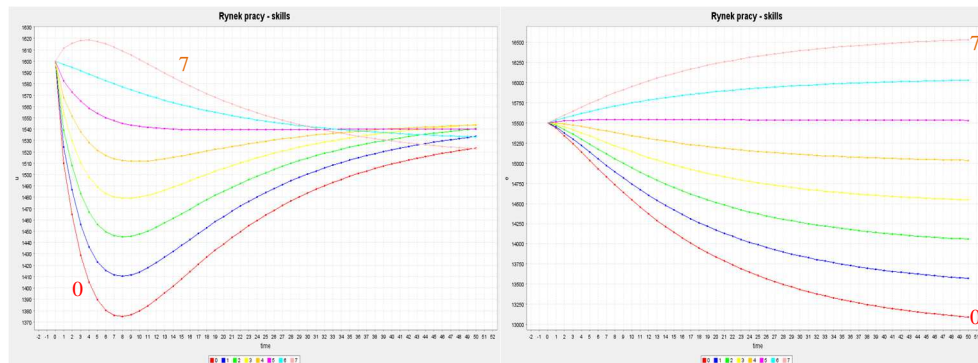
Rysunek 3. Trajektorie rynku pracy wyznaczone przy założeniu wzrostu płacy minimalnej w zakresie 310-370 EUR



Źródło: obliczenia własne

³ Parametry φ oraz ε przyjmują wartości, odpowiednio, od 0,03 do 0,025 oraz od 0,028 do 0,033. Odpowiada to wzrostowi płacy minimalnej od poziomu około 310 EUR do około 370 EUR, przy założeniu, że wartości parametrów β oraz γ są stałe i wynoszą odpowiednio 0,00009 oraz 9,2.

Rysunek 4. Wielkości zatrudnienia i bezrobocia w kolejnych okresach wyznaczone przy założeniu wzrostu płacy minimalnej w zakresie 310-370 EUR



Źródło: obliczenia własne

PODSUMOWANIE

Zaprezentowany w artykule model dynamiki rynku pracy pozwala na analizę wpływu poziomu kwalifikacji oraz płacy minimalnej na zmiany wielkości zatrudnienia i bezrobocia. Uwzględniona w modelu struktura kwalifikacji jest odzwierciedleniem heterogeniczności pracowników, która w sposób istotny determinuje popyt na siłę roboczą. Postać punktu równowagi oraz przeprowadzone symulacje komputerowe wskazują, że w przedstawionym modelu struktura kwalifikacji pracowników wpływa na dynamikę rynku zgodnie z oczekiwaniami – jej wzrost redukuje poziom bezrobocia i zwiększa zatrudnienie. Problemem pozostaje sposób mierzenia kwalifikacji, gdy możliwość wykorzystania do tego celu poziomu edukacji jest kwestionowana. Wydaje się, że lepszym miernikiem kwalifikacji pracowników może być wskaźnik edukacji osób dorosłych w społeczeństwie. Badanie empiryczne [Boguta i in. 2007] dowodzi, że właśnie wskaźnik edukacji dorosłych ma bardzo dużą siłę oddziaływania na poziom bezrobocia⁴.

Analiza oddziaływania płacy minimalnej na dynamikę rynku pracy w prezentowanym modelu wskazała, że w długim okresie wzrost płacy minimalnej wpływa dodatnio tylko na wielkość zatrudnienia, zaś bezrobocie, po początkowym okresie wzrostu, stabilizuje się na podobnym poziomie, niezależnie od zmian płacy.

⁴ Pierwotnym celem opracowania [Boguta i in. 2007] było zbadanie wpływu poziomu edukacji na wielkość bezrobocia. Przeprowadzone badania empiryczne wykazały, że to nie edukacja szkolna, ale dalsza edukacja dorosłych jest czynnikiem wyraźnie wpływającym na bezrobocie. Innymi czynnikami istotniejszymi od poziomu edukacji okazały się m.in. dostęp do Internetu oraz nakłady na badania i rozwój.

BIBLIOGRAFIA

- Altonji J. G., Blom E., Meghir C. (2012) Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers, *Annual Review of Economics*, Annual Reviews, Tom 4, str. 185-223.
- Acemoglu D., Autor D. (2012) What Does Human Capital Do? A Review of Goldin and Katz's Race between Education and Technology, *Journal of Economic Literature*, Tom 50, str. 426-463.
- Acemoglu D., Autor D. (2011) Skills, Tasks and Technologies: Implications for Employment and Earnings, *Handbook of Labor Economics*, Tom 4, str. 1043-1171.
- Boguta M., Karbownik K., Wrzosek M. (2007) Influence of the Education on the Level of Unemployment in European Countries, *Roczniki Kolegium Analiz Ekonomicznych SGH*, Zeszyt 17/2007, str. 101-117.
- Card D., DiNardo J. E. (2002) Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles, *Journal of Labor Economics*, Tom 20, str. 733-783.
- Elsby M. W. L., Michaels R., Solon G. (2009) The Ins and Outs of Cyclical Unemployment, *American Economic Journal: Macroeconomics*, American Economic Association, Tom 1, str. 84-110.
- Katz L. F., Margo R. A. (2013) Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective, NBER Working Paper Nr 18752.
- Meer J., West J. (2013) Effects of the minimum wage on employment dynamics, NBER Working Paper Nr 19262.
- Moore M. P., Ranjan P. (2005) Globalisation vs skill-biased technological change: implications for unemployment and wage inequality, *The Economic Journal*, Tom 115, str. 391 – 422.
- Shimer R. (2012) Reassessing the Ins and Outs of Unemployment, *Review of Economic Dynamics*, Elsevier for the Society for Economic Dynamics, Tom 15(2), str. 127-148.
- Weiss M., Garloff A. (2011) Skill-biased technological change and endogenous benefits: the dynamics of unemployment and wage inequality, *Applied Economics*, Tom 43, str. 811-821.

**MODEL OF THE LABOR MARKET DYNAMICS WITH
HETEROGENEOUS LABOR FORCE**

Abstract: The article presents a model of the labor market dynamics with heterogeneous employment structure, e.g. involving workers with different productivity depending on qualifications. This model allows to examine labor market flows and their dependencies on factors such as the minimum wage, whether it is imposed at all, and finally the structure of workers' qualifications. Using dynamical systems to describe the issue enables the analysis of existence, type and stability of the labor market equilibrium. Construction of the theoretical model has been supplemented by computer simulations allowing for proposing and verifying hypotheses about the dynamics of employment and unemployment.

Keywords: labor market, modeling of flows, dynamical system

MIĘDZYRYNKOWE PREMIE ZA WARTOŚĆ, WIELKOŚĆ I MOMENTUM NA GIEŁDACH AKCJI

Adam Zaremba

Uniwersytet Ekonomiczny w Poznaniu
e-mail: adam.zaremba@ue.poznan.pl

Przemysław Konieczka

Szkoła Główna Handlowa w Warszawie
e-mail: przemyslaw.konieczka@doktorant.sgh.waw.pl

Streszczenie: Niniejszy artykuł koncentruje się na charakterystykach międzyrynkowych premii za wartość, wielkość i momentum na giełdach akcji. Przeprowadzone badanie poszerza stan wiedzy naukowej na dwa sposoby. Po pierwsze, dokumentuje funkcjonowanie efektów wskaźnika wartości księgowej do rynkowej, kapitalizacji i momentum na poziomie państw. Po drugie, wykazuje, że opisane efekty wzmacniają się nawzajem pozwalając budować portfele cechujące się ponadprzeciętnymi stopami zwrotu. Obliczenia bazują na notowaniach spółek z 66 państw w latach 2000-2013.

Słowa kluczowe: premia za wartość, premia za wielkość, efekt momentum, analiza przekrojowa stóp zwrotu, międzynarodowe rynki finansowe, efekt wartości księgowej do rynkowej, efekt małej spółki.

WSTĘP

Premie za wielkość, wartość i momentum zostały dobrze udokumentowane na dojrzałych i wschodzących rynkach akcji [Fama & French 2012, Asness et al. 2013]. Wprawdzie ich pochodzenie wciąż budzi kontrowersje, jednak sam fakt, że istnieją, jest powszechnie akceptowany. Premie za wartość, wielkość i momentum są obecnie elementem wielu powszechnie akceptowanych modeli wyceny aktywów, jak choćby model trójczynnika modelu Famy-Frencha lub czteroczynnikowego modelu Carharta [Fama & French 1993, Carhart 1997]. Są one często wykorzystywane w zarządzaniu portfelem, ocenie wyników

inwestycyjnych, a nawet znajdują zastosowanie w praktyce prawnej przy ocenie szkód [Mitchell & Netter 1994] lub transakcji fuzji i przejęć [Beverley 2007].

Efekty wartości, wielkości i momentum są zwykle wykorzystywane jako objaśnienie przekrojowej zmienności w stopach zwrotu z pojedynczych spółek. Czy jednak istnieją pewne analogie na poziomie makro? Czy trzy współczynniki mogłyby zostać zastosowane do prognozowania stóp zwrotu z całych krajowych rynków akcji i wyjaśnić międzyrynkową zmienność w stopach zwrotu? Niniejszy artykuł daje przekonującą odpowiedź na to pytanie.

Opracowanie ma na celu poszerzyć wiedzę akademicką na dwa sposoby. Po pierwsze, przeprowadzone badanie w sposób jednoznaczny dokumentuje występowanie międzyrynkowych premii za wartość, wielkość i momentum przy wykorzystaniu najszerzej znanej nam dotychczas próby. Po drugie, wykazane zostają interakcje pomiędzy poszczególnymi efektami, które sprawiają, że premie wzmacniają się nawzajem.

Artykuł składa się z trzech zasadniczych części. W pierwszej kolejności zaprezentowany zostaje krótki przegląd dotychczasowych badań. Następnie przedstawiona jest charakterystyka metod badawczych i źródeł danych. Trzecia sekcja zawiera omówienie wyników analiz empirycznych. Obliczenia bazują na notowaniach spółek z 66 państw w latach 2000-2013, co stanowi najszerzą znaną nam próbę w literaturze światowej w badaniach nad zagadnieniem, które jest przedmiotem rozważań tego opracowania. Artykuł wieńczy podsumowanie i wskazanie kierunków dla dalszych badań.

PODSTAWY TEORETYCZNE

Niniejsze badanie koncentruje się na trzech czynnikach wyceny aktywów, które są dobrze udokumentowane w literaturze przedmiotu.

Czynnik wielkości związany jest z efektem małych spółek zaobserwowanym przez Banz [1981] i Reinganuma [1981]. Banz przeprowadził swoje badania na spółkach amerykańskich notowanych na NYSE w latach 1936-1975. Banz zauważył, że 20 proc. najmniejszych spółek na giełdzie wypracowują średnio o 5 proc. wyższą stopę zwrotu niż pozostałe firmy. Innymi słowy, spółki o niskiej kapitalizacji (małe spółki) mają tendencję do wypracowywania przeciętnie wyższych stóp zwrotu. Mimo, że małe spółki charakteryzowały się najczęściej wysokimi współczynnikami beta, model CAPM nie był w stanie wyjaśnić tak wysokich stóp zwrotu. Reinganum [1981] oraz Cook i Roseff [1982] potwierdzili efekt wielkości spółki przy wykorzystaniu większej próby i portfeli złożonych z decyli spółek. W późniejszym czasie występowanie efektu wielkości zostało dostrzeżone na giełdach amerykańskich i innych rynkach zagranicznych przez wielu badaczy [Herrera i Lockwood 1994, Heston i in. 1999, Rouwenhorst 1999, Horowitz i in. 2000, Fama i French 2008, Michou i in. 2010].

Czynnik wartości związany jest z tak zwanym efektem spółek nisko wycenianych lub inaczej efektem wartości. Efekt ten to tendencja spółek

niskowartościowych do wypracowywania wyższych stóp zwrotu skorygowanych o ryzyko niż spółki wzrostowe. W ramach literatury spółki dzielone są na niskowartościowe i wzrostowe biorąc pod uwagę zazwyczaj wskaźnik wartości księgowej do rynkowej (B/M). Formalne dowody statystyczne na istnienie efektu wartości zostały dostarczone przez Stattmana [1980] oraz Rosenberga i in. [1985]. Obaj autorzy wykorzystali stosunek ceny do wartości księgowej jako wskaźnik dzielący spółki na niskowartościowe i wzrostowe. Fama i French [1992] pokazali, że w latach 1963 – 1990 na rynku amerykańskim wskaźnik B/M miał większą zdolność wyjaśniania stóp zwrotu niż współczynnik beta i kapitalizacja. Fama i French [1992, 1993, 1995, 1996] uważają, że wyższe stopy zwrotu z akcji spółek o wysokim wskaźniku B/M wynikają z występowania dodatkowego czynnika ryzyka systematycznego. Davis [1994] potwierdził efekt wartości także na rynku amerykańskim, a Chan i in. [1991] i Capaul i in. [1993] potwierdzili efekt wielkości na innych rynkach zagranicznych. Efekt wartości został zaobserwowany w stopach zwrotu również przez Famę i Frencha [2012], Rouwenhorsta [1999], Chui i in. [2010] oraz Asness i in. [2013].

Trzecie analizowane zjawisko wiąże się z tak zwanym efektem momentum. Efekt momentum związany jest głównie z występowaniem autokorelacji krótkoterminowych stóp zwrotu. Pierwsze badania w zakresie autokorelacji stóp zwrotu zawierała praca DeBondta i Thalera [1985]. Znaleźli oni dowody na to, że spółki, które uzyskiwały dobre historyczne wyniki, przynoszą kiepskie stopy zwrotu w ciągu kolejnych 3-5 lat. Jagadeesh i Titman [1993, 2001] przeprowadzili analizy podobne do DeBondta i Thalera [1985], jednak koncentrując się na krótkoterminowym horyzoncie inwestycyjnym tzn. od 3 do 12 miesięcy. Przeprowadzili oni symulacje 32 różnych strategii inwestycyjnych w oparciu o spółki notowane na NYSE i AMEX w latach 1965 – 1989. Jagadeesha i Titmana [1993] zaobserwowali, że akcje z wysokimi historycznymi stopami zwrotu [zwycięzcy] wypracowują w kolejnych okresach lepsze wyniki niż akcje z niskimi historycznymi stopami zwrotu [przegrani]. Dowody na istnienie efektu momentum w stopach zwrotu z akcji na rynkach międzynarodowych zostały zaprezentowane między innymi przez Asnessa [1994], [2012], Grinblatt i Moskowitz [2004], Chui i in. [2010], Asnessa i in. [2013]. Istnieją także dowody, że czynnik momentum może być skutecznie łączony z czynnikami wartości i wielkości [Fama i French 2012; Asness i in. 2013].

Czynniki wartości, wielkości i momentum są tradycyjnie wykorzystywane do objaśniania przekrojowej zmienności w stopach zwrotu z pojedynczych spółek. Niemniej jednak, wydaje się, że można zaobserwować pewne analogie również na poziomie całych państw i związanych z nimi krajowych rynków akcji. Wprawdzie dowody nie są zbyt obfite, jednak powstałe dotychczas publikacje zdają się potwierdzać opisane zjawisko [Kouwenberg i Salomons 2005, Bhojraj and Swaminathan 2006, Kim 2012]. Garff [2013] analizuje próbę od 18 do 41 państw i dostrzega efekt wartości na poziomie krajów, jednak w badaniu brak jest formalnego wnioskowania statystycznego. Asness i in. [2013] z kolei

przedstawiają bardziej złożoną analizę statystyczną, jednak bazują jedynie na od 8 do 18 rynków w próbie.

ŹRÓDŁA DANYCH I METODY BADAWCZE

Zagadnienie wykorzystania premii za wielkość, wartość i momentum jako objaśnienia dla przekrojowej zmienności w krajowych stopach zwrotu zostało zbadane przy wykorzystaniu danych z 66 różnych państw¹ oraz dwóch typów indeksów. Skorzystano z indeksów MSCI jako indeksów reprezentacji poszczególnych rynków akcji. Wszelkie dane dotyczące cen i zmiennych fundamentalnych zostały przeliczone na dolary amerykańskie² i pochodzą z serwisu Bloomberg.

W pierwszej kolejności przeanalizowane zostają stopy zwrotu z inwestowania w wartość i wzrost na rynkach międzynarodowych. Portfele spółek z poszczególnych państw zostają posortowane na podstawie czterech czynników: wielkości, wartości, długoterminowego momentum i krótkoterminowego momentum. Czynnik wartości (V) to stosunek wartości księgowej do rynkowej (B/M) portfela krajowego w danym państwie ważony według metodologii stosownego indeksu. Czynnik wielkości (S) to łączna kapitalizacja wszystkich spółek w danym portfelu krajowym. Czynniki długo- i krótkoterminowego momentum (LTM i STM) to odpowiednio 12- i 1-miesięczne stopy zwrotu za okres poprzedzający formowanie portfela. Wybrany rynek zostaje włączony w danym okresie do próby wyłącznie w sytuacji, gdy możliwe jest jednoczesne wyznaczenie obu wskaźników. Łączna liczba indeksów w próbie zwiększa się wraz z rozwojem globalnych rynków kapitałowych z 47 na początku badania do 66 pod koniec. Wykorzystane zostają pełne miesięczne szeregi czasowe za okres 05/31/2000-11/30/2013, przy czym wszystkie portfele rekonstruowane są zawsze ostatniego dnia miesiąca. Wcześniejsze dane zostają pominięte z obawy, że zbyt mała liczba rynków w próbie mogłaby zaburzyć ostateczne wyniki.

Na podstawie czynników wartości, wielkości i momentum, skonstruowane zostają trzy odrębne portfele dla każdego czynnika, które obejmują kolejno 30% rynków z najniższymi wskaźnikami, 30% rynków z najwyższymi wskaźnikami

¹ Argentyna, Australia, Austria, Bahrajn, Belgia, Brazylia, Bułgaria, Kanada, Chile, Chiny, Kolumbia, Cypr, Czechy, Dania, Egipt, Estonia, Finlandia, Francja, Niemcy, Wielka Bryt., Grecja, Hong Kong, Węgry, Islandia, Indie, Indonezja, Irlandia, Włochy, Japonia, Jordania, Kuwejt, Łotwa, Liban, Litwa, Luxemburg, Malezja, Meksyk, Maroko, Holandia, Nowa Zelandia, Norwegia, Oman, Pakistan, Peru, Filipiny, Polska, Portugalia, Katar, Rumunia, Rosja, Arabia Saudyjska, Singapur, Słowenia, RPA, Korea Płd., Hiszpania, Szwecja, Szwajcaria, Tajwan, Tajlandia, Turcja, Ukraina, ZEA, USA, Wenezuela, Wietnam.

² Niemniej jednak, przeprowadzone obliczenia wskazują, że wyniki są odporne także na konwersję na euro i jeny japońskie.

oraz pozostałe 40% „średkowych” rynków. Portfele są równoważone. Przyjęte zostaje założenie, że z punktu widzenia pojedynczego inwestora całe rynki (lub instrumenty je odwzorowujące) są wystarczające płynne, aby uzasadnić takie podejście. Równoległe ze stopami zwrotu z portfeli czynnikowych, obliczane są także stopy zwrotu z portfela rynkowego. Jako reprezentacja portfela rynkowego wykorzystana zostaje średnia ważona kapitalizacją rynkową ze wszystkich rynków znajdujących się w próbie w analizowanym okresie.

Następnie, skonstruowane zostają w pełni zabezpieczone portfele *long/short* (L/S), które naśladują zachowanie czynników wartości, wielkości i momentum. Budowa portfeli L/S przybiera taką formę, że posiadają one dodatnią ekspozycję na rynki o wyższej oczekiwanej stopie zwrotu, a ujemną na spółki o niższej. Ujmując rzecz inaczej, portfele zawsze zawierają 100% pozycji długiej w 30% rynków o najwyższym V, LTM lub STM, lub też najniższym S oraz 100% pozycji krótkiej w 30% rynków o najniższym V, LTM lub STM, lub też najwyższym S. Z uwagi na fakt, że tak skonstruowany portfel nie wymaga angażowania kapitału, portfele zostają dodatkowo całkowicie zabezpieczone przy pomocy długiej pozycji w instrumencie wolnym od ryzyka, przy czym za reprezentację stopy zwrotu z instrumentu wolnego od ryzyka przyjęto 1-miesięczne kwotowania *bid* na stawki BBA Libor USD. Ponownie, podobnie jak w poprzednim przypadku, portfele są równoważone.

Ostatecznie, stopy zwrotu z portfeli *long/short* zostały przetestowane względem trzech różnych modeli: modelu zerowego, CAPM oraz trójczynnika modelu Fama-Frencha³. Testy bazują na logarytmicznych stopach zwrotu. Pierwszy model (ang. *zero model*) zakłada, że oczekiwana stopa zwrotu wynosi zero, więc test odpowiada na pytanie, czy stopy zwrotu są statystycznie różne od zera. Drugi wykorzystany model to Model Wyceny Aktywów Kapitałowych (CAPM, ang. *Capital Asset Pricing Model*) [Sharpe 1964, 1966, Lintner 1965, Mossin 1966]. Zgodnie z tym modelem, przeprowadzona zostaje regresja nadwyżkowych stóp zwrotu z portfeli L/S względem nadwyżkowych stóp zwrotu z portfela rynkowego. Ostatni użyty model to trójczynnika model Fama i Frencha [Fama & French 1992, 1993]. Model ten uwzględnia również efekty ceny księgowej do rynkowej oraz wielkości spółki, które są reprezentowane przez stopy zwrotu z zerokosztowych portfeli SMB i HML. Portfel SMB zawiera długą pozycję w amerykańskich małych spółkach i krótką w dużych, natomiast portfel HML długą pozycję w spółkach o wysokim wskaźniku wartości księgowej do rynkowej, a krótką w spółkach o niskim wskaźniku. Wykorzystane zostają obliczone wcześniej dane, które dostępne są na stronie internetowej Kennetha Frencha⁴.

³ Szczegółowy przegląd i charakterystyka modeli oczekiwanych stóp zwrotu zawarta jest na przykład w pracach Cambella, Lo i MacKinlaya [1997] oraz Cochrane'a [2005].

⁴ www://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

We wszystkich modelach hipoteza zerowa zakłada, że portfele nie przynoszą ponadprzeciętnych stóp zwrotu, wobec hipotezy alternatywnej, że ponadprzeciętne stopy zwrotu są istotnie różne od zera. Parametry równań estymowane są klasyczną metodą najmniejszych kwadratów i oceniane przy pomocy statystyk parametrycznych.

Na koniec badania przeprowadzona została także analiza, której celem było zbadanie interakcji pomiędzy poszczególnymi czynnikami fundamentalnymi determinującymi stopę zwrotu. W tym celu, zbudowane portfele podwójnie sortowane (ang. *double-sorted portfolios*) według par czynników V, S, LTM i STM, przy czym za progi graniczne ponownie przyjęto wartości 30% i 70%. W dalszej kolejności, skonstruowano portfele zabezpieczone long/short dla każdej z kombinacji. Przesłanki tworzenia konkretnych portfeli long/short bazowały na dotychczasowych badaniach empirycznych, a przyświecała im maksymalizacja oczekiwanej ponadprzeciętnej stopy zwrotu. W tej części badania, przyjęto, że wszystkie portfele są równoważone. Po ich skonstruowaniu, przetestowano je względem trzech scharakteryzowanych wcześniej modeli, zgodnie z identyczną procedurą, jak przedstawiona wcześniej.

WYNIKI BADAŃ

Tabela 1. obrazuje wyniki inwestycyjne zabezpieczonych portfeli czynnikowych typu *long/short* oraz ich istotność statystyczną.

Tabela 1. Jednoczynnikowe zabezpieczone portfele L/S

	Model zerowy		CAPM		Model Famy-Frencha			
	Średnia	Odch.	β	α	HML	SMB	β	α
V	0,64%	2,93%	0,09	0,45%	0,22	0,21	0,01	0,27%
	(2,77)		(2,05)	(1,99)	(3,46)	(2,44)	(0,24)	(1,25)
S	0,64%	3,36%	-0,23	0,45%	0,17	0,10	-0,28	0,35%
	(2,43)		(-4,68)	(1,83)	(2,35)	(1,03)	(-5,27)	(1,39)
LTM	0,79%	3,84%	-0,15	0,60%	-0,29	-0,18	-0,06	0,80%
	(2,63)		(-2,63)	(2,04)	(-3,57)	(-1,62)	(-1,02)	(2,75)
STM	0,12%	3,27%	-0,10	-0,07%	-0,11	-0,01	-0,07	-0,02%
	(0,45)		(-1,91)	(-0,28)	(-1,49)	(-0,12)	(-1,31)	(-0,09)

Źródło: opracowanie własne. „Średnia” oznacza średniomiesięczną logarytmiczną stopę zwrotu, „Odch.” jej odchylenie standardowe, natomiast β , α , HML i SML to parametry poszczególnych modeli. V, S, LTM i STM odnoszą się do premii za wartość, wielkość oraz długo- i krótkoterminowe momentum. Wartości w nawiasach oznaczają statystyki testowe t określające istotność statystyczną estymowanych parametrów

Wszystkie analizowane czynniki wypracowują dodatnie stopy zwrotu, choć w przypadku STM są one bliskie zera. Ponadprzeciętne stopy zwrotu pozostają statystycznie różne od zera w przypadku, V, S i LTM po skorygowaniu o ryzyko

zgodnie z modelem CAPM. Interesujące wnioski przynosi także wyznaczenie ponadprzeciętnych stóp zwrotu w oparciu o model trójczynnikowy Famy-Frecha. Po pierwsze, międzyrynkowa premia za wartość jest w znacznym stopniu objaśniana przez czynniki HML i SMB. Po drugie, premia za wielkość może być również po części wytłumaczona jednak przy pomocy czynnika HML. Może to się wydawać dość zaskakujące, bowiem intuicyjne rozumowanie wskazywałoby, że powinna tu występować znacznie silniejsza korelacja z SML. Po trzecie, zależność czynnika LTM jest relatywnie silna jednak ujemna. Ostatecznie, po skorygowaniu wyników o trójczynnikowy model FF, jedynie czynnik długoterminowego momentum przynosi dodatnie statystycznie istotne ponadprzeciętne stopy zwrotu. Premię za wielkość i wartość również pozostają dodatnie, jednak brak im istotności statystycznej, natomiast krótkoterminowe momentum nie wypracowuje premii.

Tabela 2. Prezentuje wyników dwuczynnikowych, które były budowane na podstawie spółek sortowanych według par czynników.

Tabela 2. Dwuczynnikowe zabezpieczone portfele L/S

	Model zerowy		CAPM		Model Famy-Frecha			
	Średnia	Odch.	β	α	HML	SMB	β	α
V+LTM	1,36%	9,11%	-0,18	1,17%	-0,34	-0,20	-0,08	1,39%
	(1,90)		(-1,30)	(1,64)	(-1,63)	(-0,71)	(-0,53)	(1,91)
V+STM	0,85%	5,40%	-0,15	0,66%	0,05	0,13	-0,19	0,58%
	(2,00)		(-1,89)	(1,58)	(0,40)	(0,79)	(-2,08)	(1,36)
V+S	0,53%	4,75%	-0,13	0,35%	0,46	0,31	-0,27	0,03%
	(1,43)		(-1,75)	(0,94)	(4,65)	(2,31)	(-3,69)	(0,08)
S+LTM	1,34%	6,98%	-0,53	1,15%	0,02	-0,12	-0,52	1,20%
	(2,44)		(-5,40)	(2,28)	(0,14)	(-0,58)	(-4,73)	(2,31)
S+STM	1,12%	5,57%	-0,34	0,94%	0,17	-0,06	-0,36	0,90%
	(2,57)		(-4,21)	(2,25)	(1,37)	(-0,37)	(-4,06)	(2,12)

Źródło: opracowanie własne. „Średnia” oznacza średniomiesięczną logarytmiczną stopę zwrotu, „Odch.” jej odchylenie standardowe, natomiast β , α , HML i SML to parametry poszczególnych modeli. V, S, LTM i STM odnoszą się do premii za wartość, wielkość oraz długo- i krótkoterminowe momentum. Wartości w nawiasach oznaczają statystyki testowe t określające istotność statystyczną estymowanych parametrów

Analiza wyników przedstawionych w Tabeli 2. prowadzi do kilku interesujących wniosków. Po pierwsze, kombinacje czynników z momentum notowały znacząco wyższe stopy zwrotu niż pojedyncze czynniki, choć towarzyszyła temu również wyższa zmienność. Po drugie, te same portfele okazywały się przynosić statystycznie istotne ponadprzeciętne stopy zwrotu w ujęciu modelu CAPM, chociaż należy podkreślić, że poziom tej istotności był różny. Po trzecie, warto zauważyć, że także kombinacja z STM implikowała podwyższone stopy zwrotu, chociaż samodzielnie ten czynnik nie charakteryzował się tego typu własnościami. Po czwarte wreszcie, trzy kombinacje okazały się

wyjątkowo interesujące. Wielkość w połączeniu z krótko- lub długoterminowym momentum oraz wartość z długoterminowym momentum generowały statystycznie istotne ponadprzeciętne stopy zwrotu nawet po korekcie o ryzyka z modelu Famy-Frencha.

Podsumowując, wyniki badania pokazują, że pomiędzy czynnikami widoczne są pewne synergie, które sprawiają, że premie czynnikowe związane z momentum ulegają amplifikacji w połączeniu z wartością lub wielkością.

PODSUMOWANIE

Zaprezentowane badanie koncentrowało się na analizie międzyrynkowych premii za wielkość, wartość i momentum. Pozwoliło ono udokumentować, że na globalnych rynkach akcji funkcjonują międzyrynkowe premie za wielkość, wartość i momentum. Dodatkowo, badanie pozwoliło odkryć, że efekty te na poziomie państw wzmacniają się nawzajem. Podwójnie sortowane portfele czynnikowe charakteryzują się bardzo wysokimi ponadprzeciętnymi stopami zwrotu.

Niniejszy artykuł może stać się przyczółkiem do dalszych badań, które powinny koncentrować się w kilku obszarach. Po pierwsze, bardzo użyteczne mogłyby się okazać przeniesieni modeli wyceny aktywów, jak model Carharta [1997] czy model Famy-Frencha [1993] do skali makro. Modele takie mogłyby zostać wykorzystane do oceny funduszy inwestycyjnych z globalnym mandatem inwestycyjnym lub też do analizy zdarzeń na międzynarodowych rynkach akcji. Po drugie, warto byłoby przeprowadzić formalne testy przesunięcia granicy efektywnej (ang. *mean-variance spanning*) w celu oceny zasadności alokacji aktywów w portfelu w strategii ilościowe oparte o międzyrynkowe efekty wartości, wielkości i momentum. Po trzecie wreszcie, przyszłe analizy powinny skoncentrować się na znalezieniu źródeł funkcjonowania premii z wartością i wielkością na poziomie całych państw. Pytanie, dlaczego efekty wielkości i wartości na poziomie państw właściwie działają, pozostaje otwarte.

BIBLIOGRAFIA

- Asness, C. S., Moskowitz, T. J. & Pedersen L. H. (2013) Value and momentum everywhere. *Journal of Finance*, 68 (3) 929–985.
- Asness, C.S. (1994) Variables that Explain Stock Returns. Ph.D. Dissertation, University of Chicago.
- Banz, R.W. (1981) The relation between return and market value of common stocks. *Journal of Financial Economics*, 9, s. 3–18.
- Beverley, L. (2007) Stock market event studies and competition commission inquiries. Mimeo, U.K. Competition Commission.
- Bhojraj, S. & B. Swaminathan (2006) Macromomentum, s. Returns Predictability in International Equity Indices, *Journal of Business*, 79(1), s. 429-451.

- Cambell, J.Y., A.W. Lo and A.C. MacKinlay (1997) *The Econometrics of Financial Markets*. Princeton University Press, Princeton, New Jersey, USA.
- Capaul, C., Rowley, I. & Sharpe, W. (1993) International value and growth stock returns. *Financial Analysts Journal*, 49, s. 27-36.
- Carhart, M.M. (1997) On Persistence in Mutual Fund Performance. *Journal of Finance*, 52, s. 57-82.
- Chan, L.K.C, Hamao, Y. & Lakonishok, J. (1991) Fundamentals and stock returns in Japan. *Journal of Finance*, 46, s. 1739-1764.
- Chui, A. C.W., Titman, S. & Wei, K.C. J. (2010) Individualism and momentum around the world. *Journal of Finance*, 65, s. 361-392.
- Cochrane, J.C., (2005) *Asset Pricing*. Princeton University Press, Princeton, New Jersey, USA.
- Cook T., J. & Roseff, M. (1982) Size Dividend Yield and Co-Skewness Effects on Stock Returns: Some Empirical Tests. Iowa City, s. University of Iowa, Working Paper Series 18, s. 82-20.
- Davis, L. J. (1994) The cross-section of realized stock returns, The pre-COMPUSTAT evidence. *Journal of Finance*, 49, s. 1579-1593.
- De Bondt, W. & Thaler, R. (1985) Does the Stock Market Overreact? *Journal of Finance*, 40 (3), s. 793-805.
- Fama, E.F. & French, K.R. (1992) The cross-section of expected stock returns. *Journal of Finance*, 47, s. 427-466.
- Fama, E.F. & French, K.R. (1993) Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, s. 3-56.
- Fama, E. F. & French, K.R. (1995) Size and book-to-market factors in earnings and returns. *Journal of Finance*, 50, s. 131-156.
- Fama, E. F. & French, K.R. (1996) Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 51, s. 55-84.
- Fama, E.F. & French, K. R. (2008) Dissecting anomalies. *Journal of Finance*, 63 (4), s. 1653-1678.
- Fama, E.F. & French, K.R. (2012) Size, Value, and Momentum in International Stock Returns. *Journal of Financial Economics*, 105(3), s. 457-472.
- Garff, D., 2013. Multi-Style Global Equity Investing. A Statistical Study on Combining Fundamentals, Momentum, Risk and Valuation for Improved Performance. Working paper, available at SSRN, s. [http, s.//ssrn.com/abstract=2367400](http://ssrn.com/abstract=2367400).
- Grinblatt, M. & Moskowitz, T. J. (2004) Predicting Stock Price Movements from Past Returns, The Role of Consistency and Tax-Loss Selling. *Journal of Financial Economics*, 71, s. 541-579.
- Herrera, M. J. & Lockwood, L.J. (1994) The size effect in the Mexican stock market. *Journal of Banking and Finance*, 18, s. 621-632.
- Heston, S. L., Rouwenhorst, K.G. & Weessels, R. E. (1999) The role of beta and size in the cross-section of European stock returns. *European Financial Management*, 5, s. 9 -27.
- Horowitz, J. L., Loughran, T. & Savin, N.E. (2000) The disappearing size effect. *Research in Economics*, 54 (1), s. 83-100.
- Jegadeesh, N. & Titman, S. (1993) Returns to Buying Winners and Selling Losers. Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1), s. 65-91.

- Jegadeesh, N. & Titman, S. (2001) Profitability of Momentum Strategies. An Evaluation of Alternative Explanations. *The Journal of Finance*, 56(2), s. 699–720.
- Kim, D. (2012) Value Premium Across Countries. *Journal of Portfolio Management* 38 (4), 75-86.
- Kouwenberg, R. & R. Salomons (2005) The Value Premium in Emerging Equity Markets and Local Macroeconomic Conditions. Working paper, available online at: <s://people.few.eur.nl/kouwenberg/Value%20EM%20Paper%20v2.5b.pdf>.
- Lintner, J. (1965) The Valuation of Risk Assets and The Selection of Risky Investments in Stock Portfolios and Capital Budget. *Review of Economics and Statistics*, 47, s. 13-37.
- Michou, M., Mouselli, S. & Stark A. (2010) Fundamental analysis and the modelling of normal returns in the UK. Working paper, SSRN.
- Mitchell, M.L. & Netter, J.M. (1994) The role of financial economics in securities fraud cases. Applications at the Securities and Exchange Commission. *Business Lawyer*, 49 (2), s. 545-590.
- Mossin, J. (1966) Equilibrium in a Capital Asset Market, *Econometrica*, 34(2), s. 768-783.
- Reinganum, M.R (1981) Misspecification of Capital Asset Pricing. Empirical Anomalies based on earning' yield and market values. *Journal of Financial Economics*, 9(1), s. 19-46.
- Rosenberg, B., Reid, K. & Lanstein, R. (1985) Persuasive evidence of market inefficiency. *Journal of Portfolio Management*, 11, s. 9–17.
- Rouwenhorst, K.G. (1999) Local returns factors and turnover in emerging stock market. *Journal of Finance*, 54, s. 1439-1464.
- Sharpe, W.F. (1964) Capital Asset Pricing. A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, 19, s. 425-441.
- Stattman D. (1980) Book values and stock returns. *The Chicago MBA, Journal of Selected Papers*, 4, s. 25-45.

INTER-COUNTRY VALUE, SIZE AND MOMENTUM PREMIUMS ACROSS THE STOCK MARKETS

Abstract: The study examines the characteristics of inter-country value, size and momentum premiums. We contribute to the asset-pricing literature in two ways. First, we deliver evidence on value, size and momentum premiums across countries. Second, we demonstrate, that the country-level value, size and momentum premiums tend to strengthen each other in double-sorted portfolios. We investigate stock markets in 66 countries 2000 and 2013.

Keywords: value premium, size premium, momentum effect, cross-section of stock returns, international financial markets

REVIEWERS COOPERATING WITH THE JOURNAL IN 2014

Hanna Adamkiewicz-Drwiłło
Aleksandra Baszczyńska
Jarosław Becker
Jacek Bednarz
Piotr Bilski
Zbigniew Binderman
Dariusz Błaszczuk
Lucyna Błażejczyk-Majka
Bolesław Borkowski
Jadwiga Bożek
Ryszard Budziński
Mariola Chrzanowska
Wiesław Dębski
Robert W. Dimand
Hanna Dudek
Marcin Dudziński
Gema Fernández-Avilés Calderón
Jennifer P. Foo
Peter Friedrich
Konrad Furmańczyk
Paolo Gajo
Stanisław Gędek
Vasile Glavan
Marek Gruszczyński
Urszula Grzybowska
Mariusz Hamulczuk
Stanisław Jaworski
Andrzej Karpio
Krzysztof Karpio
Marek Karwański
Stanisław Kasiewicz
Joanna Kisielińska
Stephan Klasen
Paweł Kliber
Marek Kociński
Krzysztof Kompa
Yuriy Kondratenko
Vassilis Kostoglou
Grzegorz Koszela
Stanisław Kot
Robert Kragler
Monika Krawiec
Tadeusz Kufel
Justyna Kujawska
Karol Kukuła
Tadeusz Kwater
Joanna Landmesser
Barbara Z. Liberda
Andrzej Łodziński
Piotr Łukasiewicz
Wanda Marcinkowska-Lewandowska
Jerzy Marzec
Aleksandra Matuszewska-Janica
Jose-María Montero Lorenzo
Rafik Nafkha
Magdalena Osińska
Maria Parlińska
Mariola Piłatowska
Marian Podstawka
Iwona Pomianek
Artur Prędko
Alexander Prokopenya
Mariusz Próchniak
Ryszarda Rempała
Janina Sawicka
Yochanan Shachmurove
Mirbulat B. Sikhov
Agnieszka Sompolska-Rzechuła
Stanisław Stańko
Aleksander Strasburger
Jacek Strojny
Ewa Syczewska
Wiesław Szczesny
Waldemar Tarczyński
Grażyna Trzpiot
Beata Turyna-Pułaska
Ewa Wędrowska
Andrzej Wiatrak
Marek Wierzbicki
Dorota Witkowska
Bartosz Witkowski
Michał Zasada
Jan Zawadzki
Jarosław Zawadzki
Tomasz Ząbkowski
Beata Zduniak
Monika Zielińska-Sitkiewicz
Wojciech Zieliński
Krzysztof Zmarzłowski
Dorota Żebrowska-Suchodolska