

POLSKA AKADEMIA NAUK
KOMITET STATYSTYKI I EKONOMETRII

PRZEGLĄD STATYSTYCZNY

STATISTICAL REVIEW

TOM 63

1

2016

WARSZAWA 2016

WYDAWCA

Komitet Statystyki i Ekonometrii Polskiej Akademii Nauk

RADA REDAKCYJNA

Andrzej S. Barczak, Marek Gruszczyński, Krzysztof Jajuga, Stanisław M. Kot, Tadeusz Kufel,
Władysław Milo (Przewodniczący), Jacek Osiewalski, Sven Schreiber, D. Stephen G. Pollock,
Jaroslav Ramík, Peter Summers, Matti Virén, Aleksander Welfe, Janusz Wywiał

KOMITET REDAKCYJNY

Magdalena Osińska (Redaktor Naczelny)
Marek Walesiak (Zastępca Redaktora Naczelnego, Redaktor Tematyczny)
Michał Majsterek (Redaktor Tematyczny)
Maciej Nowak (Redaktor Tematyczny)
Anna Pajor (Redaktor Statystyczny)
Piotr Fiszedler (Sekretarz Naukowy)

Wydanie publikacji dofinansowane przez Ministra Nauki i Szkolnictwa Wyższego

© Copyright by Komitet Statystyki i Ekonometrii PAN

Strona WWW „Przegląd Statystyczny”:
<http://www.przegladstatystyczny.pan.pl>

Nakład 140 egz.



Przygotowanie do druku:
Dom Wydawniczy ELIPSA
ul. Inflancka 15/198, 00-189 Warszawa
tel./fax 022 635 03 01, 022 635 17 85
e-mail: elipsa@elipsa.pl, www.elipsa.pl

Druk:
Wrocławska Drukarnia Naukowa PAN Sp. z o.o.
im. Stanisława Kulczyńskiego
53-505 Wrocław, ul. Lelewela 4

SPIS TREŚCI

Magdalena Osińska – Od redakcji	5
Marek Walesiak – Wybór grup metod normalizacji wartości zmiennych w skalowaniu wielowymiarowym	7
Maciej Małaczewski – Kapitał fizyczny, produkcja energii a czas wyczerpania się zasobów naturalnych	19
Paulina Małaczewska – Mechanizmy kształtowania się szarej strefy gospodarki. Model teoretyczny	33
Paweł Dykas, Tomasz Misiak – Neoklasyczny model wzrostu gospodarczego z sinusoidalnymi inwestycjami	49
Ewa Szlachetowska, Daniel Kosiorowski, Dominik Mielczarek – Ocena jakości aplikacyjnej odpornego algorytmu analizy skupień TCLUST na przykładzie zbioru danych dotyczących jakości powietrza w Krakowie	67
Mirosław Krzyśko, Agnieszka Majka, Waldemar Wołyński – Ocena zróżnicowania poziomu życia mieszkańców województw w latach 2003–2013 za pomocą składowych głównych dla wielowymiarowych danych funkcjonalnych oraz analizy skupień	81

SPRAWOZDANIA

Anna Jurek, Elżbieta Zalewska – Sprawozdanie z XXXIV międzynarodowej konferencji naukowej Wielowymiarowa Analiza Statystyczna (WAS 2015)	99
--	----

CONTENTS

Magdalena Osińska – From the Editors.....	5
Marek Walesiak – The Choice of Groups of Variable Normalization Methods in Multidimensional Scaling.....	7
Maciej Malaczewski – Physical Capital, Production of the Energy and Time of Natural Resources' Exhaustion	19
Paulina Malaczewska – Driving Forces of the Shadow Economy. Theoretical Model.....	33
Paweł Dykas, Tomasz Misiak – The Neoclassical Growth Model with Sinusoidal Investments... ..	49
Ewa Szlachetowska, Daniel Kosiorowski, Dominik Mielczarek – Evaluation of the Quality of Robust Clustering Algorithm TCLUST on the Example of Dataset of Air Pollutants Emission in Krakow.....	67
Mirosław Krzyśko, Agnieszka Majka, Waldemar Wołyński – Estimation of Diversity of Living Standards in Polish Voivodships in 2003–2013 Using Principal Components for Multidimensional Functional Data and Cluster Analysis	81

REPORTS

Anna Jurek, Elżbieta Zalewska – Report of the XXXIV International Scientific Conference Multivariate Statistical Analysis (WAS 2015).....	99
--	----

OD REDAKCJI

Szanowni Państwo,

w 2015 roku Redakcja Czasopisma „Przegląd Statystyczny” zgodnie z zamierzeniem wydała cztery zeszyty zwykłe z częstotliwością kwartalną, w tym jeden zeszyt w języku angielskim. To już drugi rok, kiedy możliwe było opublikowanie pełnego zeszytu anglojęzycznego, co znacząco zwiększa – dzięki dostępności wersji elektronicznej – liczbę odbiorców tego uznanego w Polsce czasopisma naukowego.

Tendencja wydawania czasopism naukowych w języku angielskim jest obecnie nieuchronna, stąd też Redakcja Przeglądu Statystycznego zachęca autorów do składania swych prac do publikacji właśnie w języku angielskim. W bieżącym roku kalendarzowym, będącym już 79. rokiem działalności czasopisma, najpierw jako organu powołanego przez Polskie Towarzystwo Statystyczne, potem organu Sekcji Statystyki Polskiego Towarzystwa Ekonomicznego, a następnie od 1974 r. do chwili obecnej – czasopisma wydawanego przez Komitet Statystyki i Ekonometrii Polskiej Akademii Nauk zamierzeniem Redakcji jest wydanie dwóch zeszytów w całości w języku angielskim. Uzasadnienie popularyzacji wyników naukowych i prowadzenia dyskusji naukowej w języku angielskim nie jest wyłącznie modą XXI wieku. I nie potrzeba odwoływać się wcale do uniwersalizmu języka łacińskiego w wiekach średnich. Wystarczy powołać się na pierwszy numer Przeglądu Statystycznego wydany w roku 1938, w którym w słowie wstępnym ówczesna Redakcja wskazała wyraźnie, że w Przeglądzie Statystycznym ogłaszane będą m.in. „oryginalne prace teoretyczne autorów obcych i polskich” i „oryginalne prace z zastosowań statystyki, ważne ze względów metodologicznych”, podając jednocześnie nazwę angielską czasopisma (*Statistical Review*), jak również tytuły artykułów w języku angielskim i recenzje publikacji w językach obcych (głównie angielskim i niemieckim), z których to wzorców korzystamy do dzisiaj. Warto przy tym przypomnieć, że twórcami Przeglądu Statystycznego byli wybitni statystycy polscy, którzy jednocześnie tworzą panteon najwybitniejszych postaci statystyki światowej, jak prof. Jerzy Neyman czy prof. Jan Czekanowski.

Przez lata nastąpiły jednak istotne zmiany. Statystyka jako nauka ewoluowała w różnych kierunkach, tworząc podstawy dla ekonometrii, biometrii czy psychometrii, zaś badania statystyczne są coraz częściej stosowane w różnych dziedzinach wiedzy, które są często zalążkiem powstawania nowych metod i technik statystycznych. Tworzy to ogromne pole do publikowania wyników badań, często różnych od profilu

Przeglądu Statystycznego. Liczba czasopism publikujących wyniki najnowszych badań z zakresu badań statystycznych i ekonometrycznych również znacząco wzrosła. Aby się wyróżniać, a jednocześnie realizować swoje cele i zadania, „Przegląd Statystyczny” odwołuje się do najlepszych tradycji statystyki polskiej jako istotnej części statystyki światowej, co udaje się dzięki publikowaniu najlepszych prac nadsyłanych do Redakcji oraz niezwykle rzetelną procedurę recenzencką.

Wyrażam nadzieję, że dążenie do doskonałości publikacji przy danym stanie wiedzy przyświeca całemu środowisku statystyków i ekonometryków polskich, co będzie znajdować odzwierciedlenie w coraz bardziej znaczących publikacjach i wzroście pozycji Przeglądu Statystycznego w światowej literaturze naukowej.

Redaktor Naczelna
Prof. dr hab. Magdalena Osińska

MAREK WALESIAK¹WYBÓR GRUP METOD NORMALIZACJI WARTOŚCI ZMIENNYCH
W SKALOWANIU WIELOWYMIAROWYM

1. WPROWADZENIE

Normalizację przeprowadza się dla macierzy danych metrycznych, tzn. gdy zmienne opisujące obiekty badania mierzone są na skali interwałowej lub ilorazowej. Charakterystykę skal pomiaru zawarto m.in. w pracach (Stevens, 1946; Walesiak, 2011, s. 13–16).

Porównanie metod normalizacji wartości zmiennych może być rozpatrywane z punktu widzenia zastosowania konkretnej metody statystycznej analizy wielowymiarowej. W sytuacji, gdy w badaniu będą wykorzystywane metody analizy skupień, porządkowania liniowego i skalowania wielowymiarowego, zmienne muszą być sprowadzone do porównywalności poprzez transformacje normalizacyjne. Dla analizy skupień badania takie przeprowadzili Milligan, Cooper (1988), Schaffer, Green (1996), Walesiak, Dudek (2016), a dla metod porządkowania liniowego Walesiak (2015), Kukula, Luty (2015)². Inne metody statystycznej analizy wielowymiarowej (analiza regresji, drzewa klasyfikacyjne i regresyjne, *conjoint analysis*, analiza czynnikowa³, analiza dyskryminacyjna, analiza korelacji kanonicznej, analiza wariancji i kowariancji) nie wymagają uprzedniej transformacji normalizacyjnej.

W artykule zaproponowano procedurę badawczą pozwalającą na wyodrębnienie grup metod normalizacji wartości zmiennych prowadzących do zbliżonych wyników skalowania wielowymiarowego. Propozycja dotyczy problemu wyboru metod normalizacji wartości zmiennych w skalowaniu wielowymiarowym bazującym na macierzy danych metrycznych.

¹ Uniwersytet Ekonomiczny we Wrocławiu, Wydział Ekonomii, Zarządzania i Turystyki, Katedra Ekonometrii i Informatyki, ul. Nowowiejska 3, 58-500 Jelenia Góra, Polska, e-mail: marek.walesiak@ue.wroc.pl.

² Badanie tutaj przeprowadzone było szersze, ponieważ obejmowało wybór metody (procedury) porządkowania liniowego. Podobne badania wcześniej przeprowadził Grabiński (1984) stosując inne kryteria wyboru.

³ Metody normalizacyjne będące przekształceniem liniowym (3) nie zmieniają wartości współczynnika korelacji liniowej Pearsona (por. Jajuga, Walesiak, 2000, s. 111), który jest wykorzystywany w analizie czynnikowej.

2. SKALOWANIE WIELOWYMIAROWE NA PODSTAWIE MACIERZY DANYCH METRYCZNYCH – OGÓLNY SCHEMAT POSTĘPOWANIA

Punktem wyjścia skalowania wielowymiarowego jest macierz odległości (niepodobieństw) między obiektami w przestrzeni m -wymiarowej $[\delta_{ik}]$, gdzie $i, k = 1, \dots, n$ oznacza numer obiektu. Wśród metod wyznaczania macierzy odległości $[\delta_{ik}]$ wyróżnia się (por. np. Borg, Groenen, 2005, s. 111–133; Zaborski, 2001, s. 40–50):

1. Bezpośrednie – np. poprzez porównywanie obiektów parami pod względem ich niepodobieństwa przez poszczególnych respondentów.
2. Pośrednie. Punktem wyjścia jest tutaj macierz danych $[x_{ij}]$ (gdzie: x_{ij} – obserwacja j -tej zmiennej w i -tym obiekcie, $j = 1, \dots, m$ – numer zmiennej). Obserwacje na zmiennych uzyskujemy stosując szacowanie na skalach pomocniczych (respondenci oceniają tutaj poszczególne obiekty dla każdej zmiennej) lub z wtórnych źródeł danych. Następnie oblicza się odległości między obiektami z wykorzystaniem miar odległości (dla danych metrycznych stosuje się wcześniej normalizację wartości zmiennych).

Skalowanie wielowymiarowe jest metodą reprezentacji macierzy odległości między obiektami w przestrzeni m -wymiarowej $[\delta_{ik}]$ w macierz odległości między obiektami w przestrzeni q -wymiarowej $[d_{ik}]$ ($q < m$) w celu graficznej prezentacji (wizualizacji) relacji zachodzących między badanymi obiektami oraz określenia (interpretacji) treści q wymiarów. Wymiary q nie są bezpośrednio obserwowalne. Mają one charakter zmiennych ukrytych, które pozwalają na wyjaśnienie podobieństw i różnic między badanymi obiektami. Ze względu na możliwość graficznej prezentacji wyników zazwyczaj q wynosi 2 lub 3.

W algorytmach skalowania wielowymiarowego stosowane są różne miary dopasowania STRESS (ang. *STandardized RESidual Sum of Squares* – standaryzowana suma kwadratów reszt). W monografii (Borg, Groenen, 2005, s. 250–254) prezentowane są m.in. funkcje: STRESS-1 Kruskala, STRESS-2 Kruskala i Carrolla, współczynnik alienacji Guttmana-Lingoesa, S-STRESS Takane, Younga i De Leeuw. Zagadnieniu oceny wartości miar dopasowania, a co za tym idzie wyborowi liczby wymiarów q skalowania Borg, Groenen (2005) poświęcili podrozdział 3.5 (s. 47–55). Mowa jest tutaj m.in. o wykresie osypiska (ang. *scree test*), prostych normach oceny miar dopasowania. Dodatkowym ważnym kryterium jest interpretowalność wymiarów (osi) skalowania wielowymiarowego (zob. Borg, Groenen, 2005, s. 55).

Dla danego zbioru obiektów $A = \{A_1, \dots, A_n\}$ oraz odległości (niepodobieństw) δ_{ik} między obiektami A_i oraz A_k w przestrzeni m -wymiarowej poszukuje się takiego odwzorowania zbioru obiektów w zbiór punktów w przestrzeni q -wymiarowej, aby (Borg, Groenen, 2005, s. 39):

$$d_{ik} \approx \hat{d}_{ik} = f(\delta_{ik}), \quad (1)$$

gdzie:

d_{ik} – odległość między obiektami A_i oraz A_k (punktami $\mathbf{x}_i$ oraz $\mathbf{x}_k$) w przestrzeni q -wymiarowej,

$\hat{d}_{ik}$ – funkcja regresji między d_{ik} a δ_{ik} .

Ogólny schemat postępowania w skalowaniu wielowymiarowym zbioru obiektów przeprowadzanych na podstawie danych metrycznych jest następujący:

$$P \rightarrow A \rightarrow X \rightarrow N \rightarrow S \rightarrow I, \quad (2)$$

gdzie:

P – wybór problemu badawczego,

A – wybór obiektów,

X – dobór zmiennych. Zgromadzenie danych i konstrukcja macierzy danych w przestrzeni m -wymiarowej $[x_{ij}]_{n \times m}$ ($i = 1, \dots, n$ – numer obiektu, $j = 1, \dots, m$ – numer zmiennej),

N – normalizacja wartości zmiennych i konstrukcja macierzy $[z_{ij}]_{n \times m}$ (z_{ij} – znormalizowana wartość j -tej zmiennej dla i -tego obiektu),

S – przeprowadzenie skalowania wielowymiarowego: wybór miary odległości (zob. tab. 2) i konstrukcja macierzy odległości w przestrzeni m -wymiarowej $[\delta_{ik}]$, $f: \delta_{ik} \rightarrow d_{ik}$ – odwzorowanie macierzy odległości w przestrzeni m -wymiarowej $[\delta_{ik}]$ w macierz odległości w przestrzeni q -wymiarowej $[d_{ik}]$ ($q < m$). Iteracyjny schemat postępowania w algorytmie smacof przedstawiono w pracy (Borg, Groenen, 2005, s. 204–205), prezentacja macierzy danych w przestrzeni q -wymiarowej $[x_{ij}]_{n \times q}$,

I – interpretacja wyników skalowania wielowymiarowego (w tym interpretacja osi).

3. NORMALIZACJA WARTOŚCI ZMIENNYCH⁴

Celem normalizacji wartości zmiennych jest doprowadzenie zmiennych do porównywalności poprzez pozbawienie mian wyników pomiaru oraz ujednolicenie ich rzędów wielkości.

Przegląd metod normalizacji wartości zmiennych przedstawia praca Walesiak (2014). Tabela 1 prezentuje metody normalizacyjne dane przekształceniem liniowym (por. Jajuga, Walesiak, 2000, s. 106–107; Zeliaś, 2002, s. 792):

$$z_{ij} = b_j x_{ij} + a_j = \frac{x_{ij} - A_j}{B_j} = \frac{1}{B_j} x_{ij} - \frac{A_j}{B_j} \quad (b_j > 0), \quad (3)$$

gdzie:

x_{ij} – wartość j -tej zmiennej dla i -tego obiektu,

z_{ij} – znormalizowana wartość j -tej zmiennej dla i -tego obiektu,

A_j – parametr przesunięcia do umownego zera dla j -tej zmiennej,

B_j – parametr skali dla j -tej zmiennej,

$a_j = -A_j/B_j$, $b_j = 1/B_j$ – parametry dla j -tej zmiennej określone w tabeli 1.

⁴ Punkt ten opracowano na podstawie artykułu Walesiak (2014).

Tabela 1.

Metody normalizacji wartości zmiennych

Typ	Nazwa metody	Parametr		Skale pomiaru zmiennych	
		b_j	a_j	przed normalizacją	po normalizacji
n1	Standaryzacja	$1/s_j$	$-\bar{x}_j/s_j$	ilorazowa lub interwałowa	interwałowa
n2	Standaryzacja pozycyjna	$1/mad_j$	$-med_j/mad_j$	ilorazowa lub interwałowa	interwałowa
n3	Unitaryzacja	$1/r_j$	$-\bar{x}_j/r_j$	ilorazowa lub interwałowa	interwałowa
n3a	Unitaryzacja pozycyjna	$1/r_j$	$-med_j/r_j$	ilorazowa lub interwałowa	interwałowa
n4	Unitaryzacja zerowana	$1/r_j$	$-\min_i\{x_{ij}\}/r_j$	ilorazowa lub interwałowa	interwałowa
n5	Normalizacja w przedziale $[-1; 1]$	$\frac{1}{\max_i x_{ij}-\bar{x}_j }$	$\frac{-\bar{x}_j}{\max_i x_{ij}-\bar{x}_j }$	ilorazowa lub interwałowa	interwałowa
n5a	Normalizacja pozycyjna w przedziale $[-1; 1]$	$\frac{1}{\max_i x_{ij}-med_j }$	$\frac{-med_j}{\max_i x_{ij}-med_j }$	ilorazowa lub interwałowa	interwałowa
n6	Przekształcenia ilorazowe	$1/s_j$	0	ilorazowa	ilorazowa
n6a		$1/mad_j$	0	ilorazowa	ilorazowa
n7		$1/r_j$	0	ilorazowa	ilorazowa
n8		$1/\max_i\{x_{ij}\}$	0	ilorazowa	ilorazowa
n9		$1/\bar{x}_j$	0	ilorazowa	ilorazowa
n9a		$1/med_j$	0	ilorazowa	ilorazowa
n10		$1/\sum_{i=1}^n x_{ij}$	0	ilorazowa	ilorazowa
n11		$1/\sqrt{\sum_{i=1}^n x_{ij}^2}$	0	ilorazowa	ilorazowa
n12	Normalizacja	$\frac{1}{\sqrt{\sum_{i=1}^n (x_{ij}-\bar{x}_j)^2}}$	$\frac{-\bar{x}_j}{\sqrt{\sum_{i=1}^n (x_{ij}-\bar{x}_j)^2}}$	ilorazowa lub interwałowa	interwałowa
n12a	Normalizacja pozycyjna	$\frac{1}{\sqrt{\sum_{i=1}^n (x_{ij}-med_j)^2}}$	$\frac{-med_j}{\sqrt{\sum_{i=1}^n (x_{ij}-med_j)^2}}$	ilorazowa lub interwałowa	interwałowa
n13	Normalizacja z zerem usytuowanym centralnie	$\frac{1}{r_j/2}$	$-\frac{m_j}{r_j/2}$	ilorazowa lub interwałowa	interwałowa

$\bar{x}_j$ – średnia dla j -tej zmiennej, s_j – odchylenie standardowe dla j -tej zmiennej, r_j – rozstęp dla j -tej zmiennej,

$m_j = \frac{\max_i\{x_{ij}\} + \min_i\{x_{ij}\}}{2}$ – środek rozstępu (ang. *mid-range*), $med_j = med_i(x_{ij})$ – mediana dla j -tej zmiennej, $mad_j = mad_i(x_{ij})$ – medianowe odchylenie bezwzględne dla j -tej zmiennej.

Źródło: Walesiak (2014, s. 364–365).

W artykule normalizację wartości zmiennych przeprowadzono w pakiecie `clusterSim` (zob. Walesiak, Dudek, 2015) programu R (R Development Core Team, 2015) z wykorzystaniem funkcji `data.Normalization`.

4. PROCEDURA BADAWCZA POZWALAJĄCA NA WYODRĘBNIE NIE GRUP METOD NORMALIZACJI PROWADZĄCYCH DO ZBLIŻONYCH WYNIKÓW SKALOWANIA WIELOWYMIAROWEGO

Procedura badawcza pozwalająca na wyodrębnienie grup metod normalizacji wartości zmiennych prowadzących do zbliżonych wyników skalowania wielowymiarowego obejmuje następujące kroki:

1. Przeprowadza się, zgodnie z ogólnym schematem postępowania w skalowaniu wielowymiarowym, następujące etapy $P \rightarrow A \rightarrow X \rightarrow N$. Do normalizacji wartości zmiennych wykorzystuje się wszystkie dopuszczalne metody ujęte w tabeli 1 (dla zmiennych ilorazowych dostępnych jest 18 metod normalizacyjnych, a dla zmiennych przedziałowych – 10 metod normalizacyjnych).
2. Dla wszystkich macierzy danych po normalizacji wartości zmiennych oblicza się odległości między obiektami (zob. tabela 2) i zestawia w macierze odległości $[\delta_{ik}^r]$ (r – numer metody normalizacyjnej). Dla zmiennych ilorazowych otrzymuje się 18 macierzy odległości, a dla zmiennych przedziałowych – 10 macierzy odległości. Miara odległości Canberra nie zależy od parametru skali B_j (Pawełek, 2008, s. 94). Zatem dopuszczalne metody normalizacyjne n6–n11 nie zmieniają wartości tej odległości (zob. tab. 2).

Tabela 2.

Miary odległości dla danych metrycznych (interwałowych, ilorazowych)

Nazwa	Odległość δ_{ik}	Rozstęp	Dozwolone normalizacje
Minkowski ($p \geq 1$)	$\sqrt[p]{\sum_{j=1}^m z_{ij} - z_{kj} ^p}$	$[0; \infty)$	n1–n13
– Manhattan (miejska) ($p = 1$)	$\sum_{j=1}^m z_{ij} - z_{kj} $	$[0; \infty)$	n1–n13
– Euklidesa ($p = 2$)	$\sqrt{\sum_{j=1}^m (z_{ij} - z_{kj})^2}$	$[0; \infty)$	n1–n13
– Czebyszewa (maximum) ($p \rightarrow \infty$)	$\max_j z_{ij} - z_{kj} $	$[0; \infty)$	n1–n13
GDM1 (Walesiak, 2002; Jajuga, Walesiak, Bąk, 2003)	$\frac{1}{2} - \frac{\sum_{j=1}^m (z_{ij} - z_{kj})(z_{kj} - z_{ij}) + \sum_{j=1}^m \sum_{l=1, l \neq i, k}^n (z_{ij} - z_{lj})(z_{kj} - z_{lj})}{2 \left[\sum_{j=1}^m \sum_{l=1}^n (z_{ij} - z_{lj})^2 \cdot \sum_{j=1}^m \sum_{l=1}^n (z_{kj} - z_{lj})^2 \right]^{\frac{1}{2}}}$	$[0; 1]$	n1–n13

Tabela 2. (cd.)

Nazwa	Odległość δ_{ik}	Rozstęp	Dozwolone normalizacje
Bray-Curtis (Bray, Curtis, 1957)*	$\frac{\sum_{j=1}^m z_{ij} - z_{kj} }{\sum_{j=1}^m (z_{ij} + z_{kj})}$	[0;1]	n6–n11
Canberra (Lance, Williams, 1966)	$\sum_{j=1}^m \frac{ z_{ij} - z_{kj} }{(z_{ij} + z_{kj})} = \sum_{j=1}^m \frac{ x_{ij} - x_{kj} }{(x_{ij} + x_{kj})}$	[0;1]	n6–n11

$i, k, l = 1, \dots, n$ – numery obiektów, $j = 1, \dots, m$ – numer zmiennej, m – liczba zmiennych, z_{ij} (z_{kj} , z_{lj}) – znormalizowana wartość j -tej zmiennej dla i -tego (k -tego, l -tego) obiektu.

* Zob. również pracę (Cormack, 1971, s. 367).

Źródło: opracowanie własne.

3. Dla każdej macierzy odległości $[\delta_{ik}^r]$ przeprowadza się skalowanie wielowymiarowe dla ustalonej liczby wymiarów q otrzymując macierz odległości między obiektami w przestrzeni q -wymiarowej $[d_{ik}^r]$.
4. Otrzymane, dla dopuszczalnych metod normalizacyjnych, macierze odległości $[d_{ik}^r]$ porównuje się z wykorzystaniem odległości miejskiej d_{rs} :

$$d_{rs} = \sum_{\substack{i,k=1 \\ i < k}}^{\frac{n(n-1)}{2}} \left| \frac{d_{ik}^r}{\max_{i,k} \{d_{ik}^r\}} - \frac{d_{ik}^s}{\max_{i,k} \{d_{ik}^s\}} \right|, \quad (4)$$

gdzie: r, s – numery metod normalizacyjnych.

W celu sprowadzenia macierzy odległości $[d_{ik}^r]$ do porównywalności we wzorze (4) podzielono odległości w każdej macierzy odległości przez wartość maksymalną. Po tej operacji odległości w każdej macierzy odległości zawarte będą w przedziale [0;1]. Im mniejszą wartość przyjmuje miara d_{rs} o postaci (4), tym większe jest podobieństwo wyników skalowania wielowymiarowego dla metod normalizacyjnych o numerach r oraz s .

5. Na podstawie macierzy odległości $[d_{rs}]$ przeprowadza się analizę skupień, która pozwala wyodrębnić grupy metod normalizacji wartości zmiennych prowadzących do zbliżonych wyników skalowania wielowymiarowego. Można zastosować tutaj jedną z wielu metod klasyfikacji (zob. np. Everitt i in., 2011; Gordon, 1999). W artykule zastosowano hierarchiczną metodę aglomeracyjną najdalszego sąsiada. Dla miary odległości GDM1 oraz odległości Braya-Curtisa można sformułować spostrzeżenia odnośnie metod normalizacyjnych ujęte w tabeli 3.

Tabela 3.

Grupy metod normalizacyjnych prowadzących do identycznych odległości w macierzy odległości wyznaczonej za pomocą miary GDM1 oraz odległości Braya-Curtisa

Grupy metod	Metody normalizacyjne	
	odległość GDM1	odległość Braya-Curtisa
A	n1, n6, n12	–
B	n2, n6a	–
C	n3, n3a, n4, n7, n13	–
D	n9, n10	n9, n10

Źródło: opracowanie własne.

Identyczne macierze odległości dla grup metod A, B, C i D wynikają z tego, że miara GDM1 nie zależy od parametru przesunięcia A_j stosowanego w metodach normalizacyjnych. Ponadto przemnożenie wartości znormalizowanych przez stałą nie zmienia odległości GDM1 i Braya-Curtisa (Walesiak, 2015):

– dla metody n13 stała równa się 2:

$$z_{ij} = \frac{x_{ij}}{r_j / 2} - \frac{m_j}{r_j / 2} = 2 \cdot \left(\frac{x_{ij}}{r_j} - \frac{m_j}{r_j} \right), \quad (5)$$

– dla metody n12 stała równa się $\sqrt{\frac{1}{n-1}}$:

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}} - \frac{\bar{x}_j}{\sqrt{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}} = \sqrt{\frac{1}{n-1}} \cdot \left(\frac{x_{ij}}{s_j} - \frac{\bar{x}_j}{s_j} \right), \quad (6)$$

– dla metody n10 stała równa się $1/n$:

$$z_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} = \frac{1}{n} \cdot \left(\frac{x_{ij}}{\bar{x}_j} \right). \quad (7)$$

W pracy (Pawełek, 2008, s. 94) wykazano, że wartości miar odległości Minkowskiego (miejska, euklidesowa, Czebyszewa) nie zależą od parametru przesunięcia A_j stosowanego w metodach normalizacyjnych. Zatem identyczne macierze odległości otrzymuje się dla grup metod normalizacyjnych ujętych w tabeli 4.

Tabela 4.

Grupy metod normalizacyjnych prowadzących do identycznych macierzy odległości dla odległości Minkowskiego

Grupy metod	Metody normalizacyjne	
	D1	D2
A	n1, n6	n1, n6, n12*
B	n2, n6a	n2, n6a
C	n3, n3a, n4, n7	n3, n3a, n4, n7, n13*
D	–	n9, n10*

D2 – po podzieleniu odległości w każdej macierzy odległości przez wartość maksymalną.

* – dla tej metody normalizacji macierz odległości jest przemnożona przez stałą (zob. wzory (5)–(7)).

Źródło: opracowanie własne.

5. WYNIKI BADANIA EMPIRYCZNEGO

W badaniu empirycznym wykorzystane zostaną dane statystyczne z artykułu (Gryszel, Walesiak, 2014) dotyczące poziomu atrakcyjności turystycznej 29 powiatów Dolnego Śląska. Ocenę poziomu atrakcyjności turystycznej powiatów Dolnego Śląska przeprowadzono z wykorzystaniem 16 zmiennych metrycznych (mierzonych na skali ilorazowej):

- x1 – miejsca noclegowe w obiektach na 1 km² powierzchni powiatu,
- x2 – liczba noclegów turystów rezydentów (Polaków) przypadających dziennie na 1 tys. mieszkańców powiatu,
- x3 – liczba noclegów turystów zagranicznych przypadających dziennie na 1 tys. mieszkańców powiatu,
- x4 – emisja zanieczyszczeń gazowych w tonach na 1 km² powierzchni powiatu,
- x5 – liczba przestępstw o charakterze kryminalnym oraz przestępstw przeciwko życiu i zdrowiu na 1 tys. mieszkańców powiatu,
- x6 – liczba przestępstw przeciwko mieniu na 1 tys. mieszkańców powiatu,
- x7 – liczba obiektów zabytkowych na 100 km² powierzchni powiatu,
- x8 – lesistość powiatu w %,
- x9 – udział obszarów prawnie chronionych w powierzchni powiatu w %,
- x10 – liczba imprez oraz wydarzeń kulturalnych i turystycznych w powiecie,
- x11 – liczba pomników przyrody w przeliczeniu na 1 km² powierzchni powiatu,
- x12 – liczba podmiotów gospodarki turystycznej na 1 tys. mieszkańców powiatu (osoby fizyczne i prawne),
- x13 – wydatki gmin i powiatów na turystykę, kulturę i ochronę dziedzictwa narodowego oraz kulturę fizyczną na 1 mieszkańca powiatu w zł,
- x14 – widzowie w kinach na 1 tys. mieszkańców powiatu,
- x15 – zwiedzający muzea na 1 tys. mieszkańców powiatu,

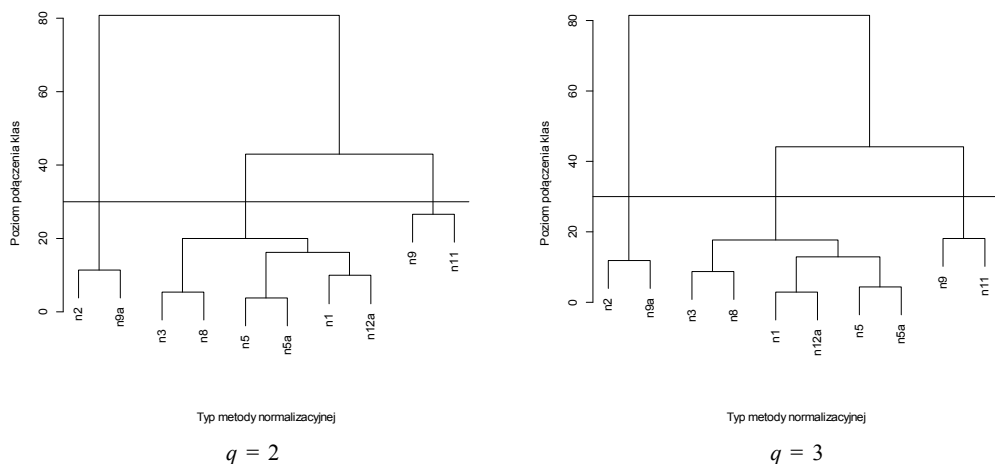
x16 – liczba pozwoleń na budowę (hotele i budynki zakwaterowania, budynki handlowo-usługowe, budynki transportu i łączności, obiekty inżynierii lądowej i wodnej) wydanych w powiecie w latach 2011–2012 na 1 km² powierzchni powiatu.

Dane statystyczne pochodzą z roku 2012 z Banku Danych Lokalnych (BDL), jedynie dane dla zmiennej x7 pochodzą od wojewódzkiego konserwatora zabytków.

W artykule zastosowano skrypt programu R przygotowany zgodnie z procedurą badawczą z sekcji 4, która pozwala na wyodrębnienie grup metod normalizacji wartości zmiennych prowadzących do zbliżonych wyników skalowania wielowymiarowego.

Pomiar zmiennych na skali ilorazowej dopuszcza wszystkie metody normalizacyjne (badaniem objęto zatem 18 metod). Z uwagi na to, że grupy metod normalizacyjnych A, B, C i D dają dla skalowania wielowymiarowego identyczne wyniki dalszej analizie poddano pierwsze metody ze wskazanych grup (n1, n2, n3, n9) oraz pozostałe metody (n5, n5a, n8, n9a, n11, n12a).

Skalowanie wielowymiarowe 29 powiatów Dolnego Śląska ze względu na poziom atrakcyjności turystycznej przeprowadzono z wykorzystaniem funkcji `smacofSym` pakietu `smacof` (Mair i in., 2015). Jako miarę odległości zastosowano odległość Euklidesa. Grupy metod normalizacji wartości zmiennych prowadzących do zbliżonych wyników skalowania wielowymiarowego prezentuje dendrogram na rysunku 1.



Rysunek 1. Dendrogram podobieństwa metod normalizacji w skalowaniu wielowymiarowym 29 powiatów Dolnego Śląska ze względu na poziom atrakcyjności turystycznej

Źródło: opracowanie własne z wykorzystaniem programu R.

Na podstawie dendrogramu wyróżniono trzy grupy metod normalizacyjnych (zarówno w dwóch, jak w trzech wymiarach) prowadzących do zbliżonych wyników skalowania wielowymiarowego w sensie macierzy odległości $[d_{ik}^r]$ oraz rozmieszczenia

obiektów w przestrzeni q -wymiarowej (w nawiasach przedstawiono metody normalizacyjne dające identyczne wyniki skalowania wielowymiarowego):

grupa 1 (3 metody): (n2, n6a), n9a,

grupa 2 (12 metod): (n1, n6, n12), (n3, n3a, n4, n7, n13), n5, n5a, n8, n12a,

grupa 3 (3 metody): (n9, n10), n11.

Do wyboru liczby klas można wykorzystać tutaj indeksy oceny jakości klasyfikacji przedstawione w pakietach `NbClust` (Charrad i in., 2014; Charrad i in., 2015) oraz `clusterSim` (Walesiak, Dudek, 2015).

W analizowanym przypadku istotne różnice między wynikami skalowania wielowymiarowego pojawiają się dla metod normalizacji wartości zmiennych z różnych grup. W dotychczasowej praktyce, nie uwzględniając zaproponowanej procedury badawczej, dokonując wyboru metody normalizacji wartości zmiennych w skalowaniu wielowymiarowym dla danych metrycznych mieliśmy do wyboru 18 propozycji (tabela 1). Rozważania ujęte w tab. 3 i 4 zmniejszają tę liczbę do 10 metod normalizacji. Wybór nadal staje się arbitralny i trudny do uzasadnienia. Zaproponowane podejście nie rozwiązuje całkowicie problemu, ale pozwala wyodrębnić grupy metod normalizacji prowadzące do zbliżonych wyników skalowania wielowymiarowego. W analizowanym przykładzie mamy już do wyboru *de facto* 3 metody normalizacji (metody normalizacji znajdujące się w tych samych grupach dają identyczne lub zbliżone wyniki skalowania wielowymiarowego). Zatem przedstawiona propozycja pozwala ograniczyć problem wyboru metody normalizacyjnej.

6. PODSUMOWANIE

W artykule przedstawiono propozycję procedury badawczej pozwalającą na wyodrębnienie grup metod normalizacji wartości zmiennych prowadzących do zbliżonych wyników skalowania wielowymiarowego. Propozycja pozwala ograniczyć problem wyboru formuły normalizacji wartości zmiennych w skalowaniu wielowymiarowym. Istotne różnice między wynikami skalowania wielowymiarowego pojawiają się dla formuł normalizacyjnych z różnych grup.

Wskazano dla miar odległości GDM1 oraz Braya-Curtisa metody normalizacyjne dające identyczne odległości w macierzy odległości. Analogiczne spostrzeżenia sformułowano dla miar odległości Minkowskiego (miejska, euklidesowa, Czebyszewa).

Wyniki badawcze zobrazowano przykładem empirycznym dotyczącym zastosowania funkcji `smacofSym` pakietu `smacof` w celu przeprowadzenia skalowania wielowymiarowego 29 powiatów Dolnego Śląska ze względu na poziom atrakcyjności turystycznej z wykorzystaniem 18 formuł normalizacyjnych.

LITERATURA

- Borg I., Groenen P. J. F., (2005), *Modern Multidimensional Scaling. Theory and Applications*, 2nd Edition, Springer Science+Business Media, New York.
- Bray J. R., Curtis J. T., (1957), An Ordination of the Upland Forest Communities of Southern Wisconsin, *Ecological Monographs*, 27 (4), 325–349.
- Charrad M., Ghazzali N., Boiteau V., Niknafs A., (2014), NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set, *Journal of Statistical Software*, 61 (6), 2–36.
- Charrad M., Ghazzali N., Boiteau V., Niknafs A., (2015), NbClust Package for Determining the Best Number of Clusters. R Package Version 3.0, URL <http://CRAN.R-project.org/package=NbClust>.
- Cormack R. R., (1971), A Review of Classification, *Journal of the Royal Statistical Society, Series A*, 134 (3), 321–367.
- Everitt B. S., Landau S., Leese M., Stahl D., (2011), *Cluster Analysis*, John Wiley & Sons, Chichester.
- Gordon A. D., (1999), *Classification*, 2nd Edition, Chapman & Hall/CRC, London.
- Grabiński T., (1984), *Wielowymiarowa analiza porównawcza w badaniach dynamiki zjawisk ekonomicznych*, Zeszyty Naukowe Akademii Ekonomicznej w Krakowie, Seria specjalna: Monografie nr 61.
- Gryszel P., Walesiak M., (2014), Zastosowanie uogólnionej miary odległości GDM w ocenie atrakcyjności turystycznej powiatów Dolnego Śląska, *Folia Turistica*, 31, 127–147.
- Jajuga K., Walesiak M., (2000), Standardisation of Data Set under Different Measurement Scales, w: Decker R., Gaul W., (red.), *Classification and Information Processing at the Turn of the Millennium*, 105–112. Springer-Verlag, Berlin, Heidelberg.
- Jajuga K., Walesiak M., Bąk A., (2003), On the General Distance Measure, w: Schwaiger M., Opitz O., (red.), *Exploratory Data Analysis in Empirical Research*, 104–109, Springer-Verlag, Berlin, Heidelberg.
- Kukuła K., Luty L., (2015), Propozycja procedury wspomagającej wybór metody porządkowania liniowego, *Przegląd Statystyczny*, 62 (2), 219–231.
- Lance G. N., Williams W. T., (1966), Computer Programs for Hierarchical Polythetic Classification (“Similarity Analyses”), *The Computer Journal*, 9 (1), 60–64.
- Mair P., De Leeuw J., Borg I., Groenen P. J. F., (2015), smacof: Multidimensional Scaling. R Package Version 1.7-0, URL <http://CRAN.R-project.org/package=smacof>.
- Milligan G. W., Cooper M. C., (1988), A Study of Standardization of Variables in Cluster Analysis, *Journal of Classification*, 5 (2), 181–204.
- Pawełek B., (2008), *Metody normalizacji zmiennych w badaniach porównawczych złożonych zjawisk ekonomicznych*, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.
- R Development Core Team (2015), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, URL <http://www.R-project.org>.
- Schaffer C. M., Green P. E., (1996), An Empirical Comparison of Variable Standardization Methods in Cluster Analysis, *Multivariate Behavioral Research*, 31 (2), 149–167.
- Stevens S. S., (1946), On the Theory of Scales of Measurement, *Science*, 103 (2684), 677–680.
- Walesiak M., (2002), *Uogólniona miara odległości w statystycznej analizie wielowymiarowej*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
- Walesiak M., (2011), *Uogólniona miara odległości GDM w statystycznej analizie wielowymiarowej z wykorzystaniem programu R*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław.
- Walesiak M., (2014), Przegląd formuł normalizacji wartości zmiennych oraz ich własności w statystycznej analizie wielowymiarowej, *Przegląd Statystyczny*, 61 (4), 363–372.
- Walesiak M., (2015), The Results of Linear Ordering of the Set of Objects via Synthetic Measures and the Choice of Normalization Formula, *Statistics in Transition – new series*, w recenzji.
- Walesiak M., Dudek A., (2015), clusterSim: Searching for Optimal Clustering Procedure for a Data Set. R package version 0.44-2, URL <http://CRAN.R-project.org/package=clusterSim>.

- Walesiak M., Dudek A., (2016), The Choice of Variable Normalization Method in Cluster Analysis with clusterSim Package and R Environment, w przygotowaniu.
- Zaborski A., (2001), *Skalowanie wielowymiarowe w badaniach marketingowych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
- Zeliaś A., (2002), Some Notes on the Selection of Normalisation of Diagnostic Variables, *Statistics in Transition*, 5 (5), 787–802.

WYBÓR GRUP METOD NORMALIZACJI WARTOŚCI ZMIENNYCH W SKALOWANIU WIELOWYMIAROWYM

Streszczenie

W skalowaniu wielowymiarowym przeprowadzanym na podstawie macierzy danych metrycznych (przedziałowych, ilorazowych) jednym z etapów jest wybór metody normalizacji wartości zmiennych. W badaniu zastosowano funkcję `data.Normalization` pakietu `clusterSim` programu R. Funkcja ta zawiera 18 różnych metod normalizacyjnych.

W artykule zaproponowano procedurę badawczą pozwalającą na wyodrębnienie grup metod normalizacji wartości zmiennych prowadzących do zbliżonych wyników skalowania wielowymiarowego. Propozycja pozwala ograniczyć problem wyboru metody normalizacji wartości zmiennych w skalowaniu wielowymiarowym. Wyniki zilustrowano przykładem empirycznym.

Słowa kluczowe: normalizacja zmiennych, skalowanie wielowymiarowe, miary odległości, program R, pakiet `clusterSim`

THE CHOICE OF GROUPS OF VARIABLE NORMALIZATION METHODS IN MULTIDIMENSIONAL SCALING

Abstract

In multidimensional scaling carried out on the basis of metric data matrix (interval, ratio) one of the stages is the choice of the variable normalization method. The R package `clusterSim` with `data.Normalization` function has been developed for that purpose. It provides 18 data normalization methods.

In this paper the proposal of procedure which allows to isolate groups of normalization methods that lead to similar multidimensional scaling results were presented. The proposal can reduce the problem of choosing the normalization method in multidimensional scaling. The results are illustrated via empirical example.

Keywords: normalization of variables, multidimensional scaling, distance measures, R program, `clusterSim` package

MACIEJ MALACZEWSKI¹KAPITAŁ FIZYCZNY, PRODUKCJA ENERGII
A CZAS WYCZERPANIA SIĘ ZASOBÓW NATURALNYCH²

1. WSTĘP

Zasoby naturalne stanowią jeden z najważniejszych czynników wzrostu gospodarczego (Stern, Cleveland, 2004; Lee i inni, 2008; Constantini, Martini, 2010; Stern, 2011). Ich rola w procesach produkcyjnych jest podwójna – z jednej strony stanowią substrat w produkcji rozmaitych dóbr przemysłowych, z drugiej – źródło energii. O ile w pierwszym przypadku możliwy jest częściowy recykling lub częściowa substytucja z innymi, odnawialnymi i dostępnymi w dużo większym stopniu materiałami, o tyle wciąż nie ma dobrej alternatywy dla ropy naftowej, gazu ziemnego i węgla w produkcji energii i zasilaniu silników maszyn ciężkich. Około 81,7% energii produkowanej dziś na świecie pochodzi z nieodnawialnych zasobów naturalnych³, a angażowany w procesie produkcji kapitał fizyczny wykorzystuje głównie paliwo pochodzące najczęściej z ropy naftowej lub gazu ziemnego, które występują na kuli ziemskiej w relatywnie dużej, ale jednak skończonej ilości.

W literaturze wzrostu gospodarczego problematyka nieodnawialnych zasobów naturalnych jest dość często omawiana. Dominujące podejście polega na ujęciu zasobów naturalnych jako czynnika produkcji w funkcji produkcji typu Cobba-Douglassa substytucyjnego z innymi czynnikami produkcji – kapitałem fizycznym, ludzkim, pracą (np. Dasgupta, Heal, 1974; Solow, 1974; Stiglitz, 1974; Scholz, Ziemes, 1999; Groth, Shou, 2002, 2007; Grimaud, Rougé, 2008, 2014; da Silva, 2008; Pittel, Bretschger, 2010; Neustroev, 2013; Malaczewski, 2014 i wiele innych). Substytucyjność ta uzasadniana jest głównie postępem technicznym (Stiglitz, 1974) oraz możliwością zmniejszenia strumienia wykorzystywanych zasobów naturalnych w przypadku wykorzystania lepszej generacji kapitału fizycznego (Solow, 1974). Interesującą ale trudną do zaakceptowania konsekwencją takiego podejścia jest wniosek, iż zasoby naturalne nieod-

¹ Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Katedra Ekonometrii, ul. Rewolucji 1905 r. nr 41/43, 90-214 Łódź, Polska, e-mail: mmalaczewski@uni.lodz.pl.

² Autor chciałby w tym miejscu podziękować kolegom i koleżankom z Katedry Ekonometrii, uczestnikom dorocznej konferencji ekonomii matematycznej w Poznaniu oraz dwóm anonimowym recenzentom za liczne wartościowe uwagi służące poprawie jakości niniejszego artykułu. Jednocześnie w całości bierze na siebie winę za wszystkie błędy i niedociągnięcia.

³ Źródło danych: International Energy Agency.

nawialne nigdy nie wyczerpią się, ich złoża będą jedynie wykładniczo kurczyć się, osiągając zero w nieskończoności. Z rzadka pojawiają się podejścia inne, rozważające zasoby naturalne jako czynnik niedoskonale substytucyjny lub komplementarny brutto (np. Constanza, Daly, 1992; Smulders, de Nooij, 2003; Growiec, Schumacher, 2008; Bretschger, 2013; Stuermer, Schwerhoff, 2013). Tymczasem, przez wzgląd na rzeczywistą rolę zasobów naturalnych w procesie produkcyjnym wydaje się, że prawidłowym sposobem modelowania jest ujęcie zasobów naturalnych jako komplementarnych względem kapitału fizycznego. W literaturze takie ujęcie jest dość rzadko spotykane, nielicznymi wyjątkami są prace Georgescu-Roegen (1979), Constanza, Daly (1992), Daly (1997, 1999), van Zon, Yetkiner (2003), Amigues i inni (2004). W przypadku, gdy zasoby naturalne stanowią czynnik produkcji komplementarny względem kapitału fizycznego, można oczekiwać ich całkowitego wyczerpania w skończonym czasie. Wówczas, w dłuższej perspektywie, jedynym dostępnym w procesie produkcji kapitałem fizycznym pozostanie ten, który nie będzie wymagał zasilania energią wytwarzaną z zasobów naturalnych. Powstaje naturalne pytanie o politykę ekonomiczną, jaka powinna być zrealizowana w takiej sytuacji.

Celem niniejszej pracy jest analiza teoretyczna długookresowego wzrostu gospodarczego w świetle problematyki wyczerpujących się zasobów naturalnych. W proponowanym modelu zakładamy, inaczej niż w klasycznej literaturze, że zasoby naturalne są komplementarne względem tej części kapitału fizycznego, którego uruchomienie jako czynnika produkcji wymaga określonych nakładów energii wytwarzanej z nieodnawialnych surowców naturalnych. W modelowanej gospodarce występuje także drugi rodzaj kapitału fizycznego niewymagający energii pochodzącej z zasobów naturalnych. Prześledzimy optymalne reguły funkcjonowania takiej gospodarki w stanie stacjonarnym, wyznaczmy krytyczny moment całkowitego wyczerpania zasobów naturalnych oraz omówimy pewne rekomendacje dla polityki gospodarczej zorientowane na oddalenie tego momentu w czasie.

Struktura pracy jest następująca. W punkcie 2 zaprezentujemy model wzrostu gospodarczego. W punkcie 3 wyznaczmy stan stacjonarny oraz prześledzimy jego niektóre własności. W punkcie 4 zajmujemy się stanem stacjonarnym, w którym gospodarka zapewnia maksymalny poziom konsumpcji *per capita*. Całość kończy podsumowanie.

2. MODEL

W konstrukcji proponowanego modelu stosujemy klasyczne podejście, w duchu prac Solowa (1956) oraz Mankiwa i inni (1992)⁴. Rozważamy gospodarkę zamkniętą, bez widocznego udziału państwa. W gospodarce występują dwa sektory – sektor pro-

⁴ Podobne podejście do modelowania zużycia zasobów naturalnych w świetle modeli wzrostu gospodarczego jak w niniejszym artykule można znaleźć w pracach Dalgaard, Strulik (2007) i da Silva (2008).

dukcji dóbr i usług oraz sektor produkcji energii zasilającej pewną część kapitału fizycznego. Gospodarka jest zaludniona przez L jednostek, stanowiących jednocześnie zasób pracy⁵, którego wielkość rośnie w czasie ze stałą stopą n :

$$\dot{L} = nL \Rightarrow L = L^0 e^{nt}, \quad (1)$$

gdzie L^0 stanowi startową wielkość zasobów sił pracy w gospodarce. Dla uproszczenia obliczeń zakładamy, że $L^0 = 1$.

W gospodarce mamy dwa substytucyjne rodzaje kapitału fizycznego. Pierwszy z nich, K_1 , tworzą te wszystkie postaci kapitału fizycznego, które nie wymagają w procesie produkcji zużycia energii wytwarzanej z zasobów naturalnych. W jego skład wchodzi zatem np. narzędzia ręczne, urządzenia napędzane siłą mięśni, budynki, biurka itp., ale także wszelkie maszyny napędzane w inny sposób, niż silnikami spalinowymi. Drugi rodzaj kapitału, K_2 , wymaga do efektywnego wykorzystania w procesie produkcji energii wytwarzanej z zasobów naturalnych. Są to m.in. samochody, samoloty, ciągniki, ale też np. piły łańcuchowe. Cały zasób kapitału jest dzielony w gospodarce pomiędzy oba sektory – sektor produkcji dóbr i usług oraz sektor produkcji energii.⁶

Zależność między produkcją Y i nakładami czynników K_1 , K_2 , L opisuje funkcja typu Cobb-Douglasa o stałych efektach skali:

$$Y = A(uK_1)^\alpha (vK_2)^\beta L^{1-\alpha-\beta}, \quad (2)$$

gdzie $u \in (0; 1)$ oznacza udział kapitału fizycznego K_1 w sektorze produkcyjnym, $v \in (0; 1)$ – udział kapitału fizycznego K_2 w sektorze produkcyjnym, $\alpha, \beta \in (0; 1)$ są elastycznościami produkcji względem (odpowiednio) kapitału K_1 i kapitału K_2 . Parametr A charakteryzuje poziom zaawansowania technologicznego gospodarki związany z produkcją dóbr. A ma również interpretację TFP. Zakładamy też, że $1 - \alpha - \beta \in (0; 1)$. Funkcja produkcji (2) spełnia warunki Inady oraz prawo malejącej produktywności krańcowej (por. Tokarski, 2011). Czynniki produkcji są względem siebie wzajemnie substytucyjne z elastycznością substytucji równą 1⁷.

W gospodarce pewną część produktu przeznacza się na inwestycje w obie formy kapitału. Ewolucje kapitałów fizycznych K_1 i K_2 dane są następującymi wzorami:

$$\dot{K}_1 = s_1 Y - \lambda K_1, \quad (3)$$

⁵ Pomimo tego, że *explicite* nie jest to pokazane, wszystkie zmienne makroekonomiczne są różniczkowalnymi funkcjami czasu.

⁶ Analogicznie rozdysponowują kapitał fizyczny Antony (2007), da Silva (2008) oraz Pittel, Bretschger (2010).

⁷ Funkcja typu Cobb-Douglasa jest szczególnym przypadkiem funkcji produkcji typu CES dla elastyczności substytucji równej 1, por. Acemoglu (2009).

$$\dot{K}_2 = s_2 Y - \lambda K_2, \quad (4)$$

gdzie $s_i \in (0; 1)$, $i = 1, 2$, oznacza stopę inwestycji w kapitał i -ty. Zakładamy, że $s_1 + s_2 \in (0; 1)$; Stopa deprecjacji kapitału fizycznego $\lambda \in (0; 1)$ jest jednakowa dla obu form kapitału⁸. W momencie początkowym gospodarka dysponuje zasobami K_1^0 i K_2^0 kapitału każdego rodzaju. Oznacza to, że w momencie $t = 0$ produkcja w gospodarce jest równa $Y^0 = A(uK_1^0)^\alpha(vK_2^0)^\beta$.

Produkcja jest przeznaczana na inwestycje powiększające oba rodzaje kapitału fizycznego oraz na konsumpcję C :

$$Y = s_1 Y + s_2 Y + C, \quad (5)$$

skąd wynika, że

$$C = (1 - s_1 - s_2)Y. \quad (6)$$

Gospodarka posiada pewne złoża zasobów naturalnych, które stanowią w naszym modelu niezbędny czynnik produkcji energii. Przez S oznaczać będziemy wielkość posiadanych złóż zasobów naturalnych, a przez R – strumień zasobów naturalnych, który w danym momencie jest zużywany (spalany) w procesie produkcji energii (niezbędny do uruchomienia kapitału K_2). Mamy zatem:

$$\dot{S} = -R \Rightarrow S(t) = S^0 - \int_0^t R dt, \quad (7)$$

gdzie przez S^0 oznaczamy wielkość złóż zasobów naturalnych w momencie początkowym.

Sektor drugi do produkcji energii wykorzystuje pozostałą część obu rodzajów kapitału fizycznego oraz strumień zasobów naturalnych, które są spalane i bezpowrotnie tracone. Wielkość produkowanej energii dana jest następującym wzorem:

$$E = B((1 - u)K_1)^\gamma((1 - v)K_2)^\delta R^{1-\gamma-\delta}, \quad (8)$$

gdzie $(1-u) \in (0; 1)$ oznacza udział kapitału fizycznego K_1 w sektorze produkcji energii, $(1-v) \in (0; 1)$ – udział kapitału fizycznego K_2 w sektorze produkcji energii; $\gamma, \delta \in (0; 1)$ są elastycznościami produkcji energii względem (odpowiednio) kapitału K_1 i kapitału K_2 . Parametr B charakteryzuje poziom zaawansowania technologicznego sektora produkcji energii. Zakładamy też, że $1 - \gamma - \delta \in (0; 1)$. Nietrudno zauważyć,

⁸ Można założyć, że obie formy kapitału zużywają się w różnym tempie, ale brak wyraźnych przesłanek empirycznych dla tego typu twierdzenia. Założenie to nie ma większego wpływu na uzyskane wnioski.

że strumień zasobów naturalnych jest substytucyjny względem obu form kapitału fizycznego, ale jednocześnie niezbędny w całym procesie.

Kapitał fizyczny K_2 wymaga pewnej ilości energii, zależnej od jego energochłonności. Zakładamy, że popyt na energię jest zatem proporcjonalny do wielkości tego kapitału:

$$E^d = dK_2, \quad (9)$$

gdzie parametr d mierzy energochłonność kapitału fizycznego K_2 , tj. jednostka kapitału fizycznego K_2 wymaga d jednostek energii do zasilenia.

Większość parametrów w powyższym modelu to współczynniki techniczne, które, jako takie, mają charakter egzogeniczny. W gestii podmiotów decyzyjnych pozostaje wysokość stóp inwestycji oraz sposób podziału obu rodzajów kapitału między sektory przy założeniu długookresowej równowagi wzrostu.

3. RÓWNOWAGA

Równowaga w gospodarce (i efektywność gospodarowania zasobami) wymaga, by popyt na energię równy był jej produkcji. W przeciwnym wypadku, jeśli popyt na energię nie jest zaspokojony, gospodarka nie może wykorzystać całego zasobu kapitału K_2 , a jeśli popyt jest zbyt niski – produkcję energii można zmniejszyć przez przesunięcie nadwyżkowego kapitału fizycznego (w jednej lub obu jego postaciach) do sektora produkcji.

Przyrównując prawe strony równań (8) i (9) otrzymujemy:

$$K_2 = \left(\frac{B(1-u)^\gamma(1-v)^\delta}{d} \right)^{\frac{1}{1-\delta}} K_1^{\frac{\gamma}{1-\delta}} R^{\frac{1-\gamma-\delta}{1-\delta}}, \quad (10)$$

co pozwala wyznaczyć wzór na wielkość zasobów naturalnych potrzebną do wyprodukowania energii zasilającej cały kapitał fizyczny K_2 :

$$R = \left(\frac{K_2^{1-\delta} d}{K_1^\gamma B(1-u)^\gamma(1-v)^\delta} \right)^{\frac{1}{1-\gamma-\delta}}. \quad (11)$$

Ze wzoru (11) wynika, że strumień spalanych w procesie produkcji energii zasobów naturalnych jest:

- tym wyższy, im większy jest zasób kapitału fizycznego K_2 ; bezpośrednio oznacza to, że te dwa czynniki są względem siebie komplementarne, bowiem jak słusznie zauważają Constanza, Daly (1992), większym zasobom kapitału fizycznego towarzyszy większe spalanie zasobów naturalnych w celu dostarczenia odpowiednio większej ilości energii;
- tym wyższy, im wyższa jest energochłonność kapitału fizycznego;

- tym niższy, im wyższy jest poziom zaawansowania technologii (parametr B) w sektorze produkcji energii;
- tym niższy, im większy jest zasób kapitału fizycznego K_1 , nie wymagającego zasilania energią, a będącego substytucyjnym względem kapitału K_2 .

Gospodarka może zmniejszyć zużycie zasobów naturalnych rezygnując z pewnej części zasobu K_2 na rzecz, zwiększonego zasobu kapitału fizycznego K_1 , prowadząc prace badawczo-rozwojowe nakierowane na zmniejszenie energochłonności kapitału fizycznego lub zwiększając produktywność w sektorze energetycznym.

Równania (2), (3) i (4), wyrażone *per capita*, prowadzą do następującego prawa ruchu obu rodzajów kapitału fizycznego:

$$\dot{k}_1 = s_1 A(u k_1)^\alpha (v k_2)^\beta - (\lambda + n) k_1, \quad (12)$$

$$\dot{k}_2 = s_2 A(u k_1)^\alpha (v k_2)^\beta - (\lambda + n) k_2, \quad (13)$$

gdzie $k_1 = K_1/L$ oraz $k_2 = K_2/L$ ⁹. Układ równań dynamicznych (12) i (13) ma jeden nietrywialny punkt równowagi:

$$k_1^* = \left(\frac{s_1^{1-\beta} s_2^\beta A u^\alpha v^\beta}{\lambda + n} \right)^{\frac{1}{1-\alpha-\beta}}, \quad (14)$$

$$k_2^* = \left(\frac{s_1^\alpha s_2^{1-\alpha} A u^\alpha v^\beta}{\lambda + n} \right)^{\frac{1}{1-\alpha-\beta}}, \quad (15)$$

gdzie przez k_1^* i k_2^* oznaczamy wielkość kapitału, odpowiednio pierwszego i drugiego rodzaju na jednostkę w równowadze. Zauważmy też, że $k_1^* = k_2^* \cdot s_1/s_2$. Można udowodnić, że otrzymany w ten sposób punkt równowagi jest przyciągającym punktem długookresowej równowagi stabilnej¹⁰. Jeżeli tak, to bez względu na warunki początkowe układu gospodarka zmierza w stronę punktu równowagi stabilnej, znajdując się w skończonym czasie w pewnym jego odpowiednio bliskim otoczeniu i osiągając go w czasie nieskończonym. W związku z tym dalej zakładamy, iż gospodarka znajduje się w równowadze wzrostu. W równowadze takiej, analogicznie jak w modelu Mankiwa-Romera-Weila, wszystkie zmienne makroekonomiczne rosną ze stałą stopą równą n .

Gospodarka w równowadze, zgodnie z równaniem (11), zużywa w każdym momencie strumień zasobów naturalnych dany następującym wzorem:

⁹ W tym przypadku przekształcenia są klasyczne, dla oszczędności miejsca pomijamy je. Zainteresowany czytelnik znajdzie je np. w pracy Tokarski (2011).

¹⁰ Dowód pomijamy. Przebiega podobnie jak w modelu Mankiwa-Romera-Weila, patrz Tokarski (2011). Dynamika tych dwóch modeli jest bowiem identyczna.

$$\begin{aligned}
 R^*(t) &= \left(\frac{(k_2^*)^{1-\delta} d}{(k_1^*)^\gamma B(1-u)^\gamma (1-v)^\delta} \right)^{\frac{1}{1-\gamma-\delta}} e^{nt} = \\
 &= \left(\frac{d}{\left(\frac{s_1}{s_2}\right)^\gamma B(1-u)^\gamma (1-v)^\delta} \right)^{\frac{1}{1-\gamma-\delta}} e^{nt} k_2^*,
 \end{aligned} \tag{16}$$

a zatem:

$$R^*(t) = \left(\frac{d}{\left(\frac{s_1}{s_2}\right)^\gamma B(1-u)^\gamma (1-v)^\delta} \right)^{\frac{1}{1-\gamma-\delta}} \left(\frac{s_1^\alpha s_2^{1-\alpha} A u^\alpha v^\beta}{\lambda + n} \right)^{\frac{1}{1-\alpha-\beta}} e^{nt}. \tag{17}$$

Strumień zużywanych zasobów naturalnych $R^*(t)$ rośnie w czasie *ceteris paribus* ze stopą równą stopie wzrostu zasobów pracy. To oznacza, że zużycie zasobów naturalnych *per capita* jest stałe w czasie i jedynie wielkość populacji (wielkość zasobów pracy) wyznacza całkowite zużycie zasobów naturalnych w czasie.

Wielkość niezużytych złóż zasobów naturalnych (w gospodarce znajdującej się w równowadze) jest równa:

$$\begin{aligned}
 S^*(t) &= S^0 - \int_0^t R^* dt = \\
 &= S^0 - \int_0^t \left(\frac{d}{\left(\frac{s_1}{s_2}\right)^\gamma B(1-u)^\gamma (1-v)^\delta} \right)^{\frac{1}{1-\gamma-\delta}} \left(\frac{s_1^\alpha s_2^{1-\alpha} A u^\alpha v^\beta}{\lambda + n} \right)^{\frac{1}{1-\alpha-\beta}} e^{n\tau} d\tau,
 \end{aligned} \tag{18}$$

$$S^*(t) = S^0 - \frac{1}{n} \left(\frac{d}{\left(\frac{s_1}{s_2}\right)^\gamma B(1-u)^\gamma (1-v)^\delta} \right)^{\frac{1}{1-\gamma-\delta}} \left(\frac{s_1^\alpha s_2^{1-\alpha} A u^\alpha v^\beta}{\lambda + n} \right)^{\frac{1}{1-\alpha-\beta}} (e^{nt} - 1). \tag{19}$$

Wyczerpanie złóż zasobów naturalnych nastąpi w chwili $\tilde{t}$, w której $S^*(\tilde{t}) = 0$:

$$\tilde{t} = \ln \left(\frac{n S^0}{\left(\frac{d}{\left(\frac{s_1}{s_2}\right)^\gamma B(1-u)^\gamma (1-v)^\delta} \right)^{\frac{1}{1-\gamma-\delta}} \left(\frac{s_1^\alpha s_2^{1-\alpha} A u^\alpha v^\beta}{\lambda + n} \right)^{\frac{1}{1-\alpha-\beta}} + 1} \right)^{\frac{1}{n}}. \tag{20}$$

Od momentu $\tilde{t}$ zasoby naturalne są całkowicie wyczerpane, a cały zasób kapitału K_2 staje się bezużyteczny. Racjonalne wydaje się zastąpienie go w całości kapitałem typu K_1 . Przy przyjętych postaciach funkcji produkcji, zarówno dóbr i usług jak i energii, niezbędna jest jednak chociażby najmniejsza ilość kapitału K_2 . Istotnie, $K_2 = 0$ implikuje $Y = 0$. W gospodarce z funkcją produkcji typu Cobba-Douglasa nie

jest zatem możliwe całkowite zrezygnowanie z kapitału fizycznego wykorzystującego zasoby naturalne jako źródło energii.

Czy dla takiej gospodarki rozwiązaniem byłoby zaprzestanie inwestycji w kapitał K_2 ? Wówczas, na mocy (4), zasób kapitału K_2 będzie wykładniczo malał, osiągając wartość zerową dopiero w nieskończoności. Oznacza to stopniowy, wykładniczy spadek wielkości produkcji (*ceteris paribus*). Nietrudno zauważyć także, że $s_2 = 0$ implikuje z kolei $\tilde{t} = +\infty$, zatem zasoby naturalne wyczerpią się dopiero w nieskończoności, pozwalając utrzymać niezerową produkcję przez cały ten czas. Wydaje się, że rozwiązaniem jest tak duże inwestowanie w kapitał K_1 , które pozwoli na zrekomensowanie niedostatku kapitału K_2 . Malejąca produktywność krańcowa kapitału K_1 sprawia, że będzie on przynosić coraz mniejszy przyrost produkcji przy liniowo rosnącej jego deprecjacji. *Per capita* oznaczać to będzie stopniowy spadek wielkości produkcji w czasie, aż do zera w nieskończoności. Ewentualny postęp techniczny (który w naszym modelu tradycyjnie może przejawiać się w stałym, wykładniczym wzroście TFP, czyli A) może jedynie spowolnić ten proces.

4. OPTIMALIZACJA KONSUMPCJI

Gospodarka może sterować udziałami u i v obu form kapitału fizycznego w poszczególnych sektorach. Na mocy (2) i (10) funkcja produkcji, po wyeliminowaniu K_2 , przyjmuje postać:

$$Y = AK_1^{\alpha + \frac{\beta\gamma}{1-\delta}} \left(\frac{B}{d}\right)^{\frac{\beta}{1-\delta}} R^{\frac{\beta(1-\gamma-\delta)}{1-\delta}} L^{1-\alpha-\beta} u^\alpha (1-u)^{\frac{\beta\gamma}{1-\delta}} v^\beta (1-v)^{\frac{\beta\delta}{1-\delta}}. \quad (21)$$

Równanie (21) ukazuje wpływ udziałów u i v na wielkość produkcji. Wielkość konsumpcji *per capita*, na mocy (6), dana jest wzorem:

$$c = (1 - s_1 - s_2) AK_1^{\alpha + \frac{\beta\gamma}{1-\delta}} \left(\frac{B}{d}\right)^{\frac{\beta}{1-\delta}} R^{\frac{\beta(1-\gamma-\delta)}{1-\delta}} L^{-\alpha-\beta} u^\alpha (1-u)^{\frac{\beta\gamma}{1-\delta}} v^\beta (1-v)^{\frac{\beta\delta}{1-\delta}}. \quad (22)$$

Rozważmy maksymalizację konsumpcji *per capita* względem udziałów u i v przy ustalonych K_1 , R i L oraz s_1 i s_2 . Warunki pierwszego rzędu dane są następującymi wzorami:

$$\begin{aligned} \frac{\partial c}{\partial u} &= (1 - s_1 - s_2) AK_1^{\alpha + \frac{\beta\gamma}{1-\delta}} \left(\frac{B}{d}\right)^{\frac{\beta}{1-\delta}} R^{\frac{\beta(1-\gamma-\delta)}{1-\delta}} L^{-\alpha-\beta} v^\beta (1-v)^{\frac{\beta\delta}{1-\delta}} \times \\ &\quad \times (\alpha u^{\alpha-1} (1-u)^{\frac{\beta\gamma}{1-\delta}} - \frac{\beta\gamma}{1-\delta} u^\alpha (1-u)^{\frac{\beta\gamma}{1-\delta}-1}) = 0, \end{aligned} \quad (23)$$

$$\begin{aligned} \frac{\partial c}{\partial v} = (1 - s_1 - s_2) A K_1^{\alpha + \frac{\beta\gamma}{1-\delta}} \left(\frac{B}{d} \right)^{\frac{\beta}{1-\delta}} R^{\frac{\beta(1-\gamma-\delta)}{1-\delta}} L^{-\alpha-\beta} u^\alpha (1-u)^{\frac{\beta\gamma}{1-\delta}} \times \\ \times (\beta v^{\beta-1} (1-v)^{\frac{\beta\delta}{1-\delta}} - \frac{\beta\delta}{1-\delta} v^\beta (1-v)^{\frac{\beta\delta}{1-\delta}-1}) = 0. \end{aligned} \quad (24)$$

Rozwiązanie układu równań $\frac{\partial c}{\partial u} = 0$ i $\frac{\partial c}{\partial v} = 0$ prowadzi do wniosku, że maksymalizujące konsumpcję udziały obu form kapitału w poszczególnych procesach produkcyjnych są następujące: $u = \frac{\alpha(1-\delta)}{\gamma\beta + \alpha(1-\delta)}$ i $v = 1 - \delta$.¹¹ Oczywiście $u, v \in (0; 1)$. Zauważmy, że gdy elastyczności produkcji i produkcji energii względem K_1 i K_2 są równe (tj. $\alpha = \beta$ i $\gamma = \delta$) to $u = v = 1 - \delta$. Zauważmy też, że udziały te są niezależne od wielkości zmiennych makroekonomicznych, zależą jedynie od parametrów gospodarki. Jeżeli zatem parametry te nie zmieniają się w czasie, to reguły wyboru optymalnego poziomu konsumpcji pozostają takie same, zarówno w stanie równowagi, jak i poza nim.

Załóżmy teraz, że w gospodarce oba rodzaje kapitału są w obu sektorach ulokowane w sposób maksymalizujący konsumpcję *per capita*. Wówczas konsumpcja *per capita* w takiej gospodarce, na mocy równań (2), (6), (10), (14) i (15), dana jest następującym wzorem:

$$\begin{aligned} c^* = (1 - s_1 - s_2) \left(A \left(\frac{\alpha(1-\delta)}{\gamma\beta + \alpha(1-\delta)} \right)^\alpha (1-\delta)^\beta \right)^{\frac{1}{1-\alpha-\beta}} \times \\ \times \left(\frac{1}{\lambda + n} \right)^{\frac{\alpha+\beta}{1-\alpha-\beta}} s_1^{\frac{\alpha}{1-\alpha-\beta}} s_2^{\frac{\beta}{1-\alpha-\beta}}. \end{aligned} \quad (25)$$

Postawmy zadanie maksymalizacji konsumpcji *per capita* ze względu na obie stopy inwestycji. Warunki pierwszego rzędu w tym przypadku dane są następującymi wzorami:

$$\begin{aligned} \frac{\partial c^*}{\partial s_1} = \left(A \left(\frac{\alpha(1-\delta)}{\gamma\beta + \alpha(1-\delta)} \right)^\alpha (1-\delta)^\beta \right)^{\frac{1}{1-\alpha-\beta}} \left(\frac{1}{\lambda + n} \right)^{\frac{\alpha+\beta}{1-\alpha-\beta}} s_2^{\frac{\beta}{1-\alpha-\beta}} \times \\ \times \left(-s_1^{\frac{\alpha}{1-\alpha-\beta}} + (1 - s_1 - s_2) \frac{\alpha}{1 - \alpha - \beta} s_1^{\frac{\alpha}{1-\alpha-\beta}-1} \right) = 0, \end{aligned} \quad (26)$$

$$\begin{aligned} \frac{\partial c^*}{\partial s_2} = \left(A \left(\frac{\alpha(1-\delta)}{\gamma\beta + \alpha(1-\delta)} \right)^\alpha (1-\delta)^\beta \right)^{\frac{1}{1-\alpha-\beta}} \left(\frac{1}{\lambda + n} \right)^{\frac{\alpha+\beta}{1-\alpha-\beta}} s_1^{\frac{\alpha}{1-\alpha-\beta}} \times \\ \times \left(-s_2^{\frac{\beta}{1-\alpha-\beta}} + (1 - s_1 - s_2) \frac{\beta}{1 - \alpha - \beta} s_2^{\frac{\beta}{1-\alpha-\beta}-1} \right) = 0. \end{aligned} \quad (27)$$

¹¹ Bez trudu można sprawdzić, że warunki dostateczne istnienia maksimum funkcji (22) są spełnione.

Rozwiązanie układu dwóch powyższych równań prowadzi do wniosku, że $s_1^* = \alpha$ i $s_2^* = \beta^{12}$, co jest zgodne z klasycznymi wynikami tego typu dla modelu Mankiwa-Romera-Weila¹³ i oznacza, że w długookresowej równowadze wzrostu optymalne stopy inwestycji s_1^* , s_2^* , przy których gospodarka zapewnia maksymalną konsumpcję *per capita* są równe odpowiednim elastycznościom produkcji względem K_1 i K_2 .

Podstawiając optymalne udziały poszczególnych form kapitału oraz złote stopy akumulacji do (20) otrzymujemy następujący wzór na $\tilde{t}^*$:

$$\tilde{t}^* = \ln \left(\frac{nS^0}{\left(\left(\frac{d}{(\frac{\alpha}{\beta})^\gamma B \left(\frac{\gamma\beta}{\gamma\beta + \alpha(1-\delta)} \right)^\gamma \delta^\delta} \right)^{\frac{1}{1-\gamma-\delta}} \left(\frac{\alpha^\alpha \beta^{1-\alpha} A \left(\frac{\alpha(1-\delta)}{\gamma\beta + \alpha(1-\delta)} \right)^\alpha (1-\delta)^\beta}{\lambda + n} \right)^{\frac{1}{1-\alpha-\beta}} + 1} \right)^{\frac{1}{n}}}. \quad (28)$$

Stopień skomplikowania powyższego wzoru sprawia, że analiza wpływu poszczególnych parametrów makroekonomicznych gospodarki na czas wyczerpania się zasobów naturalnych jest utrudniona – znaki odpowiednich pochodnych zależą od układu wartości poszczególnych parametrów. Obliczenia są dużo łatwiejsze w przypadku A , B , d , λ i S^0 :

$$\frac{\partial \tilde{t}^*}{\partial A} < 0; \quad \frac{\partial \tilde{t}^*}{\partial B} > 0; \quad \frac{\partial \tilde{t}^*}{\partial d} < 0; \quad \frac{\partial \tilde{t}^*}{\partial \lambda} > 0; \quad \frac{\partial \tilde{t}^*}{\partial S^0} > 0. \quad (29)$$

Prowadzi to do następujących wniosków:

- Postęp techniczny, wyrażający się we wzroście TFP w sektorze produkcji dóbr i usług (we wzroście wartości parametru A) wpływa negatywnie na oczekiwany czas wyczerpania się zasobów naturalnych. Jeżeli zatem odkrycia technologiczne spowodują, że z jednakowego nakładu czynników produkcji powstanie więcej produktu, to jednocześnie, przy stałych w czasie stopach inwestycji, większe będą nakłady inwestycyjne na zasób kapitału K_2 . Zwiększanie w czasie rozmiarów kapitału wymagającego energii produkowanej z zasobów naturalnych implikuje konieczność produkcji większej ilości energii, co powoduje wzrost zużycia zasobów naturalnych i przyspiesza moment ich całkowitego wyczerpania.
- Jeżeli jednak postęp techniczny przekłada się na wzrost wartości parametru B (tj. wzrost TFP w sektorze produkcji energii), wówczas moment wyczerpania się zasobów naturalnych oddala się w czasie. Przy wyższym B tę samą ilość energii można uzyskać przy mniejszych nakładach wszystkich czynników produkcji energii, w tym przede wszystkim przy mniejszym nakładzie zasobów naturalnych, co oczywiście wydłuża czas ich eksploatacji i ma wpływ na zwiększenie się $\tilde{t}^*$.

¹² Tu także bez trudu można sprawdzić, że warunki dostateczne istnienia maksimum funkcji (25) są spełnione.

¹³ Por. Tokarski (2011).

- Postęp techniczny może też powodować zmniejszanie się współczynnika energochłonności kapitału fizycznego d . Jeżeli energochłonność zmaleje, kapitał fizyczny K_2 będzie wymagał mniejszego zużycia energii. Niższa produkcja energii oznacza mniejszą eksploatację zasobów naturalnych. To z kolei oddala moment wyczerpania się zasobów.
- Wzrost stopy deprecjacji kapitału λ , która, jak założyliśmy, jest identyczna dla obu form kapitału, wymaga większych nakładów inwestycyjnych na ich odtworzenie (większe inwestycje restytucyjne). To z kolei, wraz z malejącą produktywnością krańcową, implikuje niższe poziomy zasobów obu rodzajów kapitału fizycznego w stanie równowagi. Niższy zasób kapitału K_2 wymaga mniejszej ilości energii, co ogranicza zużycie zasobów naturalnych i oddala w czasie moment ich całkowitego wyczerpania.
- Wreszcie, im większe są całkowite złoża zasobów naturalnych tym dłuższy jest czas, w jakim pozostaną one dostępne. Zwiększyć S^0 można jednak wyłącznie przez poszukiwania nowych złóż zasobów naturalnych.

5. PODSUMOWANIE

W niniejszym artykule zaprezentowano model wzrostu gospodarczego, w którym zasoby naturalne stanowią źródło energii dla (przynajmniej) części kapitału fizycznego. Zasoby naturalne są komplementarne względem (części) kapitału fizycznego – im wyższy zasób kapitału wymagającego energii, tym większy wymagany nakład zasobów naturalnych, i na odwrót – wraz ze zwiększonym wykorzystaniem zasobów naturalnych możliwe staje się użycie większej ilości kapitału fizycznego. Wobec ograniczonej i nieodnawialności zasobów naturalnych będących źródłami energii oznacza to ich wyczerpanie w skończonym czasie – zgodnie z intuicją ekonomiczną, a wbrew klasycznym rezultatom płynącym z modeli, w których zakłada się substytucyjność zasobów naturalnych i kapitału fizycznego. Pomimo tego, pewien rodzaj substytucyjności między zasobami naturalnymi a kapitałem fizycznym w analizowanym modelu występuje – w produkcji energii te czynniki produkcji są do pewnego stopnia wzajemnie zastępowalne, jednakże oba są w niej niezbędne.

Szczególną uwagę poświęcono gospodarce w równowadze wzrostu, w której wybrane parametry (decyzyjne) mogą podlegać wyborom. Za kryterium decyzyjne przyjęto maksymalizację konsumpcji na zatrudnionego. Pozwoliło to na wyznaczenie optymalnego podziału obu rodzajów kapitału fizycznego między dwa sektory gospodarki oraz na wyznaczenie złotych stóp oszczędności/inwestycji. Obliczenia wykazały, że rozdysponowanie kapitału fizycznego pomiędzy sektorem produkcji dóbr i usług a sektorem produkcji energii jest zależne przede wszystkim od roli, jaką w procesie produkcji energii pełni kapitał fizyczny wymagający energii, a konkretnie od poziomu elastyczności produkcji energii względem tegoż kapitału. Wynik ten jest niezależny od czasu i stanu gospodarki, jedynie od makroekonomicznych parametrów technologii produkcyjnych w obu sektorach.

Złote reguły akumulacji obu form kapitału w równowadze są równe elastycznościom produkcji dóbr i usług względem odpowiednich kapitałów. Wynik ten jest zgodny z klasycznymi wynikami (patrz Tokarski, 2011).

Model pozwala na wyznaczenie czasu wyczerpania zasobów naturalnych. Czas ten zależy od wszystkich parametrów makroekonomicznych modelowanej gospodarki, jednak większość z nich ma na niego wpływ zarówno pozytywny jak i negatywny. Ich sumaryczny wpływ jest zatem zależny od układu wartości parametrów, co może być przedmiotem dalszej, szczegółowej analizy, w tym także numerycznej. Jedynie wpływ kilku parametrów ma charakter jednoznaczny, na co wskazują znaki odpowiednich pochodnych. Wskazówki dla polityki ekonomicznej gospodarki, która chce jak najdłużej dysponować zasobami naturalnymi, są następujące. Postęp techniczny powinien mieć ukierunkowaną formę, mianowicie powinien być silniej zorientowany na zwiększenie TFP w sektorze produkcji energii, gdyż postępowi technicznemu w sektorze produkcji dóbr i usług towarzyszy zwiększone zużycie zasobów naturalnych wskutek zwiększonej akumulacji kapitału fizycznego absorbującego więcej energii, co implikuje wzrost zapotrzebowania na zasoby naturalne. Istotnym zatem jest, by prace naukowo-badawcze nastawione na poprawę jakości produkcji w sektorze produkcji energii były intensywniejsze niż w sektorze dóbr i usług. Pomocna będzie także troska o zmniejszenie energochłonności kapitału fizycznego, np. poprzez kreowanie nowej generacji mniej energochłonnego kapitału fizycznego bądź poprawę jakości aktualnie funkcjonującego kapitału i eliminowanie strat energii. Należy jednocześnie poszukiwać nowych istniejących złóż zasobów naturalnych celem zwiększenia ich dostępnej ilości.

Zaprezentowany model nie jest oczywiście doskonały i wymaga dalszej pracy. Można spróbować rozszerzyć go o jawny postęp techniczny, o elementy optymalizacji dynamicznej, czy też o zwiększenie stopnia substytucyjności między obiema formami kapitału. Wydaje się, że jakościowe wyniki będą w tym przypadku podobne, choć oczywiście wymaga to dogłębnych analiz.

LITERATURA

- Acemoglu D., (2009), *Introduction to Modern Economic Growth*, MIT press.
- Amigues J. P., Grimaud A., Moreaux M., (2004), Optimal Endogenous Sustainability with an Exhaustible Resource Through Dedicated R&D, *Les Cahiers du LERNA*, 4, 154.
- Antony J., (2007), Depletion of Non-Renewable Resources and Endogenous Technical Change (No. 291), *Volkswirtschaftliche Diskussionsreihe/Institut für Volkswirtschaftslehre der Universität Augsburg*.
- Bretschger L., (2013), Population Growth and Natural Resource Scarcity: Long-Run Development under Seemingly Unfavorable Conditions, *The Scandinavian Journal of Economics*, 115 (3), 722–755.
- Costanza R., Daly H. E., (1992), Natural Capital and Sustainable Development, *Conservation Biology*, 6 (1), 37–46.
- Costantini V., Martini C., (2010), The Causality between Energy Consumption and Economic Growth: A Multi-Sectoral Analysis using Non-stationary Cointegrated Panel Data, *Energy Economics*, 32 (3), 591–603.
- Dalgaard C. J., Strulik H., (2007), Rediscovering the Solow Model: An Energy Network Approach, Discussion Papers, Department of Economics, University of Copenhagen.

- Daly H. E., (1997), Georgescu-Roegen versus Solow/Stiglitz, *Ecological Economics*, 22 (3), 261–266.
- Daly H. E., (1999), How Long Can Neoclassical Economists Ignore the Contributions of Georgescu-Roegen? *Bioeconomics and Sustainability: Essays in Honour of Nicholas Georgescu-Roegen*, 13–24.
- Dasgupta P., Heal G., (1974), The Optimal Depletion of Exhaustible Resources, *The Review of Economic Studies*, 41 (5), 3–28.
- da Silva A. S., (2008), Growth with Exhaustible Resource and Endogenous Extraction Rate, *Economic Modelling*, 25 (6), 1165–1174.
- Georgescu-Roegen N., (1979), Energy Analysis and Economic Valuation, *Southern Economic Journal*, 1023–1058.
- Grimaud A., Rougé L., (2008), Environment, Directed Technical Change and Economic Policy, *Environmental and Resource Economics*, 41 (4), 439–463.
- Grimaud A., Rougé L., (2014), Carbon Sequestration, Economic Policies and Growth, *Resource and Energy Economics*, 36 (2), 307–331.
- Groth C., Schou P., (2002), Can Non-renewable Resources Alleviate the Knife-edge Character of Endogenous Growth? *Oxford Economic Papers*, 54 (3), 386–411.
- Groth C., Schou P., (2007), Growth and Non-renewable Resources: the Different Roles of Capital and Resource Taxes, *Journal of Environmental Economics and Management*, 53 (1), 80–98.
- Growiec J., Schumacher I., (2008), On Technical Change in the Elasticities of Resource Inputs, *Resources Policy*, 33 (4), 210–221.
- Lee C. C., Chang C. P., Chen P. F., (2008), Energy-Income Causality in OECD Countries Revisited: The Key Role of Capital Stock, *Energy Economics*, 30 (5), 2359–2373.
- Malaczewski M., (2014), *Zasoby naturalne, postęp techniczny a długookresowy wzrost gospodarczy*, WUŁ, Łódź.
- Mankiw N. G., Romer D., Weil D. N., (1992), A Contribution to the Empirics of Economic Growth, *The Quarterly Journal of Economics*, 107 (2), 407–437.
- Neustroev D., (2013), The Uzawa-Lucas Growth Model with Natural Resources (No. 52937), University Library of Munich, Germany.
- Pittel K., Bretschger L., (2010), The Implications of Heterogeneous Resource Intensities on Technical Change and Growth, *Canadian Journal of Economics/Revue canadienne d'économie*, 43 (4), 1173–1197.
- Scholz C. M., Ziemes G., (1999), Exhaustible Resources, Monopolistic Competition, and Endogenous Growth, *Environmental and Resource Economics*, 13 (2), 169–185.
- Smulders S., De Nooij M., (2003), The Impact of Energy Conservation on Technology and Economic Growth, *Resource and Energy Economics*, 25 (1), 59–79.
- Solow R. M., (1956), A Contribution to the Theory of Economic Growth, *The Quarterly Journal of Economics*, 70 (1), 65–94.
- Solow R. M., (1974), Intergenerational Equity and Exhaustible Resources, *The Review of Economic Studies*, 41 (5), 29–45.
- Stern D. I., Cleveland C. J., (2004), Energy and Economic Growth, *Encyclopedia of energy*, 2, 35–51.
- Stern D. I., (2011), The Role of Energy in Economic Growth, *Annals of the New York Academy of Sciences*, 1219 (1), 26–51.
- Stiglitz J., (1974), Growth with Exhaustible Resources: Efficient and Optimal Growth Paths, *Review of Economic Studies*, (symposium volume), 123–137.
- Stuermer M., Schwerhoff G., (2013), Technological Change in Resource Extraction and Endogenous Growth, *Bonn Econ Discussion Papers*, 12.
- Tokarski T., (2011), *Ekonomia matematyczna. Modele Makroekonomiczne*, PWE, Warszawa.
- Van Zon A., Yetkiner I. H., (2003), An Endogenous Growth Model with Embodied Energy-saving Technical Change, *Resource and Energy Economics*, 25 (1), 81–103.

KAPITAŁ FIZYCZNY, PRODUKCJA ENERGII A CZAS WYCZERPANIA SIĘ ZASOBÓW NATURALNYCH

Streszczenie

Celem pracy jest analiza długookresowego wzrostu gospodarczego gospodarki z nieodnawialnymi zasobami naturalnymi. W proponowanym modelu zakładamy, że zasoby naturalne są głównym źródłem energii niezbędnej do zasilenia kapitału fizycznego. Jednocześnie postulujemy występowanie drugiego rodzaju kapitału fizycznego, niewymagającego energii. Rozważamy optymalne postępowanie takiej gospodarki, gdzie celem jest maksymalizacja konsumpcji *per capita*, a także analizujemy moment wyczerpania się zasobów naturalnych.

Słowa kluczowe: nieodnawialne zasoby naturalne, wzrost gospodarczy, zasoby naturalne jako źródło energii, komplementarność zasobów naturalnych i kapitału fizycznego

PHYSICAL CAPITAL, PRODUCTION OF THE ENERGY AND TIME OF NATURAL RESOURCES' EXHAUSTION

Abstract

The aim of this paper is to analyze long-run economic growth of the economy endowed with natural resources. In the model we assume that natural resources are the main source of the energy necessary to power physical capital. We also assume existence of second type of physical capital that does not need energy. We consider optimal consumption per capita – maximizing behaviour of the economy, and also analyze the time of exhaustion of natural resources.

Keywords: non-renewable natural resources, economic growth, natural resources as an energy source, complementarity between natural resources and physical capital

PAULINA MALACZEWSKA¹MECHANIZMY KSZTAŁTOWANIA SIĘ SZAREJ STREFY GOSPODARKI.
MODEL TEORETYCZNY²

1. WPROWADZENIE

Pod pojęciem szarej strefy³ rozumieć należy wszelkie aktywności gospodarcze, które przyczyniają się do tworzenia PKB, a które pozostają niezarejestrowane (za Schneider, Enste, 2000, s. 78). Tak szeroka definicja szarej strefy gospodarki sprawia, że zaliczyć do niej należy działalność całkowicie nielegalną, przestępczą, taką jak produkcja i sprzedaż dóbr całkowicie zakazanych, kradzieże, handel ludźmi, przemysł, prostytutka. Oprócz tego w skład szarej strefy wchodzi też działania ze sfery legalnej, które polegają na unikaniu całości bądź części obciążeń podatkowych, parafiskalnych i świadczeń społecznych, zaniżaniu swoich dochodów, zysków, unikaniu przepisów biurokratycznych, regulacji rynku pracy (np. wymogów BHP) itp. W związku z tym, w niniejszej pracy zawężono ogólną definicję szarej strefy zgodnie z sugestiami zawartymi w pracy Raczkowski, Schneider (2013, s. 1). I tak też, pod pojęciem gospodarka nieoficjalna (gospodarka cienia) rozumieć należy wszelkie aktywności gospodarcze niezarejestrowane, które podzielić można na dwie części – gospodarkę całkowicie nielegalną, czyli czarny rynek oraz gospodarkę półlegalną, czyli szarą strefę. Dodatkowo z tak ograniczonej definicji szarej strefy wykluczona została jeszcze produkcja i usługi „wewnętrzne” gospodarstw domowych, takie jak sprzątanie i gotowanie w celu zaspokojenia potrzeb członków gospodarstwa domowego, robienie weków czy nauczanie własnych dzieci (Schneider, Williams, 2013, s. 25). Ostatecznie, otrzymano definicję szarej strefy gospodarki, która jest zgodna z terminologią stosowaną przez Główny Urząd Statystyczny w Polsce⁴. W związku z tym, w niniejszej pracy przyjęto, iż szara strefa gospodarki obejmuje aktywności

¹ Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Katedra Ekonometrii, ul. Rewolucji 1905 r. nr 41/43, 90-214 Łódź, Polska, e-mail: pmalaczewska@uni.lodz.pl.

² Autorka serdecznie dziękuje prof. W. Milo, prof. J.J. Sztudyngierowi oraz Recenzentom za liczne uwagi, które znacząco przyczyniły się do poprawienia jakości niniejszej pracy.

³ Niezwykle bogatą i dogłębną analizę różnych definicji szarej gospodarki, które pojawiają się w literaturze ekonomicznej znaleźć można w pracy Mróz (2002, s. 11–23).

⁴ Wdrożenie Europejskiego Systemu Rachunków Narodowych i Regionalnych w Unii Europejskiej (ESA2010) do polskich rachunków narodowych. Zmiany metodologiczne oraz ich wpływ na główne agregaty makroekonomiczne, Notatka informacyjna, GUS, Warszawa, 2014, s. 12–13.

gospodarcze, które są legalne w danym kraju, ale które nie zostały w całości bądź części zarejestrowane ze względu na:

- (1) dążenia do uniknięcia obciążeń podatkowych;
- (2) dążenia do uniknięcia składek na ubezpieczenie społeczne;
- (3) nieprzestrzeganie wymogów prawa pracy, takich jak: warunki pracy, płaca minimalna, czas pracy;
- (4) dążenia do uniknięcia procedur biurokratycznych.

Kraciuk (2013, s. 87–88) powołując się na wyniki uzyskane przez Schneidera wyróżnia dwa obszary szarej strefy – pracę nierejestrowaną (tzw. pracę na czarno) oraz działalność handlową, usługową, w której niezarejestrowaniu ulega część prowadzonej aktywności gospodarczej. Według różnych szacunków pierwszy z tych obszarów stanowi około dwie trzecie działalności w szarej strefie i dotyczy przede wszystkim pracy w budownictwie, handlu detalicznym, usługach hotelarskich i restauracyjnych, natomiast drugi obszar obejmuje m.in. małe sklepy, bary, firmy taksówkowe.

Celem niniejszego badania jest stworzenie teoretycznego modelu kształtowania się szarej strefy gospodarki. Uwzględnione w nim zostaną różne determinanty szarej aktywności, będące wynikiem dotychczasowych badań teoretycznych i empirycznych nad tą problematyką, m.in. obciążenia podatkowe, korupcja urzędników państwowych, nadmierna biurokracja, efektywność organów ścigania i stawka kar za działalność nierejestrowaną, moralność podatkowa społeczeństwa, jakość instytucji publicznych. Po wyznaczeniu stanu równowagi gospodarki przeprowadzona zostanie analiza jego wrażliwości.

Układ niniejszej pracy jest następujący. Artykuł rozpoczyna krótki przegląd literatury ekonomicznej dotyczącej determinant szarej gospodarki, pochodzący z różnych prac teoretycznych i empirycznych. Następnie przybliżone zostaną wybrane modele teoretyczne kształtowania się strefy nieoficjalnej. Kolejna część zawierać będzie autorską propozycję modelu teoretycznego szarej gospodarki. Całość zamknie analiza wrażliwości skonstruowanego modelu, wyciągnięcie wniosków i podsumowanie.

2. DETERMINANTY SZAREJ STREFY GOSPODARKI

W niemal wszystkich badaniach nad czynnikami wpływającymi na wielkość szarej strefy wymienia się obciążenia podatkowe i paropodatkowe (m.in. Schneider, Williams, 2013; Kraciuk, 2013). Wraz ze wzrostem obciążeń podatkowych korzyść z działalności nierejestrowanej, wolnej od tego typu ciężarów, rośnie. Spodziewać się zatem należy wzrostu rozmiarów szarej gospodarki. Tę intuicyjną zależność potwierdziło wiele badań empirycznych, m.in. Ihrig, Moe (2004). Tym niemniej, Gołębiowski (2007, s. 6) sugeruje, że „Zbyt prostym byłoby jednak założenie, że wystarczy obniżyć podatki. (...) Stąd wzrost szarej strefy powinien skłaniać do spojrzenia na cały sektor publiczny i prowokować do szeroko rozumianej konsolidacji fiskalnej”. Autor przytacza dodatkowo wyniki badań przeprowadzonych w Nowej Zelandii, zgodnie z którymi istnieje optymalny poziom przychodów podatkowych, poniżej którego obniżanie obciążeń

podatkowych nie przynosi już pożądanych rezultatów w zmniejszaniu rozmiarów szarej strefy. Co więcej, rolę odgrywa tutaj nie tylko obniżenie samego opodatkowania, ale i zmiany strukturalne w formach opodatkowania, polegające np. na zwiększeniu udziału opodatkowania pośredniego w opodatkowaniu ogółem.

Niejednoznaczny kierunek i siłę wpływu poszczególnych form opodatkowania akcentuje w swojej pracy również Cichocki (2009). Na podstawie badania przeprowadzonego dla polskiej gospodarki zauważył, że impuls w postaci podniesienia przeciętnego opodatkowania pośredniego prowadzi do wzrostu rozmiarów szarej gospodarki, zaś podniesienie opodatkowania bezpośredniego – do jej spadku. Ten zaskakujący rezultat w przypadku podatków bezpośrednich Cichocki tłumaczy niedoskonałością pomiaru wielkości szarej strefy i sugeruje konieczność uwzględniania w dalszych badaniach nie samej wielkości szarej strefy, ale dynamiki jej zmian.

Dodatkowy problem w przypadku obciążeń podatkowych wiąże się z dwoma zagadnieniami behawioralnymi, które wpływają na odbiór przez społeczeństwo wielkości rzeczywistych obciążeń podatkowych – tzw. moralnością podatkową (ang. *tax morale*) i iluzją fiskalną. W przypadkach społeczeństwa cechującego się niską moralnością podatkową brak jest zaufania do organów sprawujących władzę, iż wartość pobranych podatków zostanie oddana społeczeństwu pod postacią dóbr i usług publicznych. W rezultacie nawet relatywnie niskie stawki opodatkowania mogą być postrzegane jako zbyt duże i prowadzić do przeniesienia aktywności gospodarczej do strefy nierejestrowanej, na co zwraca uwagę m.in. Schneider, Williams (2013), Walewski (2009). Natomiast w przypadku iluzji fiskalnej możliwe jest ukrycie rzeczywistych ciężarów podatkowych lub też prawdziwych beneficjentów systemu podatkowego. Badania nad związkiem szarej strefy i iluzji fiskalnej znaleźć można w pracy Buehn i inni (2012).

Wpływ obciążeń podatkowych na wielkość szarej strefy poddaje pod wątpliwość Aruoba (2010). Na podstawie analizy korelacji dla 118 państw zwraca uwagę na fakt, iż obecnie przestała to być kluczowa determinanta wielkości szarej gospodarki. Twierdzi, że należy brać pod uwagę czynniki, które wpływają na otoczenie gospodarcze i swobodę prowadzenia działalności gospodarczej, takie jak jakość instytucji publicznych (ang. *institutions*⁵). Jako jej wskaźnik przyjmuje tzw. miernik praworządności (ang. *Rule of Law*). Podobny czynnik, obok opodatkowania, wymienia w swoich badaniach Schneider (m.in. Schneider, Williams, 2013). Określa go jako jakość i efektywność rządu. Dodatkowo zwraca uwagę też na rolę organów kontroli i narzędzi odstraszenia przed działalnością nierejestrowaną (np. kar, grzywien)⁶, choć nie udało się tego jeszcze potwierdzić za pomocą badań empirycznych ze względu na brak danych.

⁵ Dokładnie, Aruoba definiuje *institutions* jako zbiór zasad, które określają sposób i swobodę prowadzenia działalności gospodarczej.

⁶ Na co też zwraca uwagę Beloded (2005).

Wśród innych czynników Kabaj (2009, s. 7–8) wymienia sytuację panującą na rynku pracy, której miernikiem jest wielkość bezrobocia i regulacje rynku pracy. Wraz ze wzrostem stopy bezrobocia należy się spodziewać, że część dotychczas zatrudnionych osób, z konieczności zdobycia środków niezbędnych do życia i utrzymania rodziny, zdecyduje się na pracę nierejestrowaną, co doprowadzi do wzrostu ogólnego poziomu szarej strefy. Kabaj przytacza jednak badania, które potwierdzają tę zależność, jak i takie, które poddają ją pod wątpliwość, por. m.in. Cichocki, Tyrowicz (2011)⁷. W przypadku natomiast ogólnego poziomu biurokracji (w tym także regulacji na rynku pracy) jej wzrost, poziom skomplikowania, nieczytelności i niska jakość przepisów prawnych stanowi niewątpliwie jeden z kluczowych czynników skłaniających podmioty gospodarcze do prowadzenia działalności w szarej strefie, na co zwracają też uwagę Schneider, Enste (2000).

W literaturze ekonomicznej często zwraca się również uwagę na zależność pomiędzy szarą strefą gospodarki a poziomem korupcji, m.in. Bilotkach (2006), Choi, Thum (2005), Friedman i inni (2000). W pracy Dreher, Schneider (2010) autorzy podsumowali niejednoznaczne wyniki kilku badań empirycznych z różnych krajów stawiając hipotezę, że zależność pomiędzy korupcją i szarą gospodarką jest zróżnicowana i zależy od poziomu rozwoju gospodarczego analizowanych krajów. W przypadku krajów rozwiniętych szara strefa i korupcja są względem siebie substytucyjne, natomiast w krajach rozwijających się – komplementarne. Wynika to, zdaniem autorów, z odmiennego charakteru szarej strefy w tych krajach i zróżnicowanej jakości instytucji publicznych. Przeprowadzone badanie nie potwierdziło jednoznacznie powyższej hipotezy, jego główną słabością był brak odporności wyników na różne mierniki korupcji.

Kolejną z determinant szarej strefy jest poziom nierówności płacowych. Im nierówności płacowe są większe, tym większe jest poczucie niesprawiedliwości społecznej i mniejsze zaufanie do organów sprawujących władzę. Co więcej, podmioty gospodarcze nie wierzą w możliwość dogonienia zamożniejszej od siebie grupy społecznej, i, jak należy przypuszczać, chętniej prowadzić będą działalność gospodarczą poza sektorem formalnym. Badania nad tym zjawiskiem na przykładzie gospodarek transformujących przedstawiono w pracy Rosser i inni (2000). Wśród rzadziej analizowanych sił napędowych szarej strefy wymienić można poziom IQ społeczeństwa, np. Salahodjaev (2015), religijność i rodzaj wyznania, np. Heinemann, Schneider (2011), kapitał społeczny, np. Giza-Poleszczuk (2009).

Podsumowanie wyników kilkunastu różnych badań empirycznych znaleźć można w pracy Schneider, Williams (2013). Stwierdzono, iż istnieje kilka determinant szarej strefy gospodarki, które objaśniają 84–98% jej zmienności. Należą do nich obciążenia podatkowe i parapodatkowe (35–38%), jakość instytucji państwowych (10–12%), regulacje rynku pracy (7–9%), transfery społeczne (5–7%), usługi publiczne (5–7%) oraz moralność podatkowa (22–25%).

⁷ Łączy się z tym kolejny problem – czy szara strefa jest procykliczna czy antycykliczna?

3. MODEL TEORETYCZNY

W literaturze ekonomicznej znaleźć można różne propozycje teoretycznych modeli kształtowania się szarej strefy. Najczęściej akcentuje się w nich pojedynczą determinantę szarej gospodarki, np. korupcję, opodatkowanie.

Friedman i inni (2000) w swoim, uznawanym za fundamentalny, artykule proponują statyczny model kształtowania się szarej strefy jako rezultatu występowania korupcji (nadmiernych opłat biurokratycznych) w sferze oficjalnej gospodarki. Choi, Thum (2005) również stworzyli model, w którym pojawiają się skorumpowani urzędnicy w strefie oficjalnej. Płynie z niego interesujący wniosek – w sytuacji, w której występuje korupcja wśród urzędników państwowych istnienie szarej strefy gospodarki jest zjawiskiem pożądanym. Firmy, przed podjęciem decyzji o rozpoczęciu działalności gospodarczej kalkulują czy stać je na opłacenie kosztów wejścia na rynek. Jeżeli koszty te są zwiększane przez żądania łapówkarskie urzędników, to może się okazać, że przedsiębiorstwa, które w warunkach uczciwego rynku rozpoczęłyby swoją działalność, w sytuacji korupcji nie będą w stanie tego uczynić. Mogą natomiast wejść do szarej strefy gospodarki (wolnej od tych dodatkowych kosztów) i tam prowadzić swoją działalność, co w rezultacie przyczyni się do zmniejszenia żądań łapówkarskich i redukcji korupcji w gospodarce.

Nieco inne podejście do zależności między korupcją i szarą gospodarką znaleźć można w pracy Bilotkach (2006). Przedstawiono w niej grę w postaci symultanicznej pomiędzy przedsiębiorcą i kontrolującym go urzędnikiem. Przedsiębiorca może zataić część swojej produkcji (przenieść ją do szarej strefy), zaś urzędnik może z nim współpracować – i przyjąć łapówkę, lub też nie dać się przekupić i nałożyć dodatkową karę na przedsiębiorcę. Natomiast Johnson i inni (1997) stworzyli model, w którym oprócz rządu i przedsiębiorstw występuje jeszcze jeden podmiot – mafia. Rząd ustala politykę gospodarczą w sferze oficjalnej, mafia zaś zasady funkcjonowania w szarej strefie, i w rezultacie prowadzi konkurencję na rynku próbując zachęcić przedsiębiorstwa do prowadzenia aktywności w ich sektorze.

Beloded (2005) proponuje dynamiczny model szarej gospodarki, w którym uwzględnia dodatkowo rynek finansowy i możliwość zaciągania pożyczek na prowadzenie aktywności gospodarczej od lokalnych pożyczkodawców lub od międzynarodowych instytucji. Pożyczki lokalne nie wymagają formalnych umów (nie podlegają zabezpieczeniu), w związku z czym są objęte dużo wyższym oprocentowaniem, które stanowi rekompensatę ryzyka. Model Ihrig, Moe (2004) koncentruje się wokół decyzji jednostek związanych z akumulacją kapitału i wyborem czasu pracy w szarej strefie i strefie oficjalnej. W szarej gospodarce przedsiębiorstwa nie płacą podatków (chyba, że zostaną przyłapane na takiej aktywności), ale w zamian za to wykorzystują mniej efektywne metody produkcji.

Jędrzejowicz (1995) konstruuje model, w którym szara gospodarka powstaje jako rezultat polityki skarbowej państwa – ustalonej wielkości stawki opodatkowania, kary za działalność nierejestrowaną oraz wysokości nakładów budżetowych na wykrycie

szarej aktywności. Natomiast Amaral, Quintin (2006) proponują model, w którym nie występują bariery wejścia do strefy oficjalnej dla mniej wykwalifikowanych pracowników i jednostki mogą swobodnie przepływać pomiędzy obiema strefami.

Pomimo licznych prac teoretycznych brakuje ciągle kompleksowego modelu, który uwzględniłby możliwie wiele sił napędowych szarej gospodarki, wynikających z dotychczasowych badań empirycznych. Próbę stworzenia takiego modelu podjęto w niniejszej pracy.

3.1. ZAŁOŻENIA I KONSTRUKCJA MODELU

Proponowany model jest statyczny. Występują w nim dwie grupy podmiotów – rząd i gospodarstwa domowe. Gospodarstwa domowe dysponują zasobem czasu pracy $\bar{L}$, który mogą poświęcić na pracę w sektorze rejestrowanym gospodarki (L_r) lub też na pracę nierejestrowaną (L_s). Normując łączny czas pracy do jednostki otrzymujemy $L_r = 1 - L_s$.

Działalność gospodarcza w szarej strefie nie podlega opodatkowaniu. Co więcej, zwolniona jest ona również z konieczności spełniania przepisów prawa pracy, m.in. wymogów BHP, maksymalnego czasu pracy, posiadania zezwoleń na prowadzenie działalności gospodarczej. W związku z tym wiąże się z brakiem kosztów biurokratycznych (korupcyjnych⁸) b . W rezultacie, gospodarstwa domowe wybierając aktywność w szarej gospodarce czerpią korzyści związane z brakiem opodatkowania i opłat biurokratycznych.

Przeniesienie działalności gospodarczej do szarej gospodarki nie przynosi tylko korzyści, ale związane jest również z pewnymi stratami. Przede wszystkim obarczone jest ono ryzykiem wykrycia, co wiąże się z koniecznością uiszczenia dodatkowej kary P . Kara P zależy od prawdopodobieństwa wykrycia działalności nierejestrowanej, na które składa się efektywność organów kontroli p oraz rozmiary szarej strefy – im aktywność nierejestrowana jest większa, tym trudniej ją ukryć ($\frac{\partial P}{\partial L_s} > 0$ i $\frac{\partial^2 P}{\partial L_s^2} > 0$). W literaturze ekonomicznej można znaleźć różne propozycje postaci funkcji P , m.in. Jędrzejowicz (1995), Choi, Thum (2005), Beloded (2005), Malaczewska (2014). Zgodnie z sugestiami Friedman i inni (2000) ostatecznie przyjęto funkcję P postaci:

$$P = paL_s^2, \quad (1)$$

gdzie $p \in (0,1)$ jest efektywnością organów kontroli, zaś $a \in (0,\infty)$ wysokością kary. Konsekwencją zaproponowanej postaci funkcji P jest to, iż łatwo jest ukryć małe rozmiary aktywności nierejestrowanej. Wraz z jej wzrostem rosną koszty jej ukrycia,

⁸ Koszty biurokratyczne możemy potraktować również jako opłaty korupcyjne, czyli łapówki dla urzędników, które gospodarstwa domowe zmuszone są zapłacić w celu pomyślnego lub przyspieszonego rozstrzygnięcia interesujących ich kwestii związanych z ich aktywnością gospodarczą, np. otrzymanie zezwolenia, wygranę przetargu, por. m.in. Friedman i inni (2000).

co więcej – kolejne przyrosty L_s są coraz bardziej kosztowne ze względu na rosnące ryzyko dostrzeżenia i wykrycia.

Co więcej, gospodarstwa domowe w szarej strefie nie mogą korzystać ze wszystkich dóbr i usług publicznych. Przykładowo, dostęp każdej jednostki do drogi publicznej jest taki sam, bez względu na to, czy pracuje w szarej strefie czy strefie oficjalnej. Natomiast podmiot pracujący „na czarno”, który czuje się pokrzywdzony przez swojego pracodawcę, ma ograniczone możliwości dochodzenia swoich praw przed sądem, ze względu na brak formalnej umowy i ryzyko poniesienia kary za działalność w szarej strefie. Przez x oznaczono stan jakości państwa, czyli jakość zasobu dóbr i usług publicznych oferowanych przez rząd. Zasób ów obejmuje między innymi administrację, organizację prawną, infrastrukturę, sądownictwo⁹. Stan jakości państwa x zwielokrotnia dochody z działalności gospodarczej – aktywność prowadzona w gospodarce dostarczającej dobra i usługi publiczne wysokiej jakości będzie bardziej produktywna, aniżeli w sytuacji gospodarki, w której przepisy prawne, infrastruktura itp. są o niskim standardzie.

W szarej strefie możliwość korzystania z dóbr i usług publicznych jest ograniczona, co odzwierciedla parametr $\lambda \in (0,1)$, który mierzy dostępność usług instytucji państwowych w strefie nieformalnej. Gdy $\lambda = 0$, gospodarstwa domowe w szarej strefie nie mają w ogóle dostępu do dóbr i usług publicznych, gdyby $\lambda = 1$, dostępność byłaby taka sama jak w sferze oficjalnej¹⁰.

Wynagrodzenie w sferze rejestrowanej i szarej strefie gospodarki jest różne i wynosi odpowiednio w_r i w_s . Dochody uzyskiwane w strefie oficjalnej ulegają opodatkowaniu stawką $\tau \in (0,1)$. Dodatkowo w sferze oficjalnej gospodarstwa domowe ponoszą pewne obciążenia biurokratyczne b , związane z koniecznością prowadzenia działalności gospodarczej zgodnie z obowiązującymi przepisami i normami prawa pracy.

Obciążenia biurokratyczne $b \in (1,\infty)$ wprowadzone zostały do modelu razem z opodatkowaniem, ze względu na podobny mechanizm ich oddziaływania¹¹. Założono, że opłaty biurokratyczne są zależne od wielkości dochodu gospodarstw domowych. Niech Y_r oznacza dochód wypracowany w sferze rejestrowanej. Wówczas $(1 - \tau b)Y_r$ stanowi część dochodu, która trafia do gospodarstw domowych, zaś τY_r odpowiada opodatkowaniu pobieranemu przez rząd. Pozostała część dochodu $\tau Y_r(b - 1)$ znika

⁹ Stan jakości państwa można utożsamić z *institutions*, które opisał Aruoba (2010).

¹⁰ Ten przypadek nie będzie rozważany ze względu na założenie o ograniczonym dostępie do dóbr i usług publicznych w szarej strefie.

¹¹ Z postaci równania (2) czytelnik może odnieść wrażenie, że autor niniejszej pracy traktuje podatki jako łapówkę, którą społeczeństwo jest zobowiązane wręczać rządowi. Tymczasem, co wynika z postaci funkcji użyteczności gospodarstw domowych (4), zapłacone podatki wracają do nich pod postacią dóbr i usług publicznych świadczonych przez rząd. Łapówki płacone skorumpowanym urzędnikom nie mają takiej własności. Zatem w modelu występuje wyraźne rozróżnienie stawki opodatkowania od stawki obciążeń biurokratycznych (korupcyjnych), mimo iż występują wspólnie w równaniu (2).

w gospodarce w ramach kosztów biurokratycznych, np. jako czas zmarnowany w urzędach na stanie w kolejkach, który można by było poświęcić na pracę lub czas wolny¹².

Uzyskane dochody gospodarstwo domowe poświęca na konsumpcję C i zapłatę oczekiwanych kar P za prowadzenie działalności nierejestrowanej. Dla uproszczenia założono, że budżet gospodarstw domowych jest zbilansowany, tj. nie mogą one pożyczać ani oszczędzać. W ten sposób otrzymane zostało ograniczenie budżetowe gospodarstw domowych:

$$w_r(1 - L_s)(1 - \tau b)(1 + x) + w_s L_s(1 + \lambda x) = C + P. \quad (2)$$

W modelu zakładamy, że społeczeństwo cechuje się pewnym poziomem *tax morale*. W zależności od jego wielkości, działalność w szarej strefie, jako aktywność nielegalna, może ciążyć na sumieniach gospodarstw domowych i powodować dysużyteczność. Parametr $m \in \langle 0, \infty \rangle$ określa jak bardzo doskwiera moralnie gospodarstwom domowym działalność nierejestrowana, niezadowolone to zależy również od skali szarej aktywności – im jest większa, tym większy jest spadek użyteczności. Ów czynnik moralny jest postaci $-mL_s$ i został wprowadzony do funkcji użyteczności gospodarstw domowych. Gdy $m = 0$, to czynnik moralny nie ma wpływu na decyzje gospodarstw domowych.

Rząd czuwa, poprzez odpowiednie organy kontroli, nad sposobem i legalnością prowadzenia działalności gospodarczej. W przypadku wykrycia aktywności w szarej strefie nakłada kary P . Otrzymuje zatem dochody z dwóch źródeł – z wpływów z oczekiwanych kar z szarej strefy oraz z podatków zapłaconych przez gospodarstwa domowe działające w sferze rejestrowanej. Uzyskane przychody rząd przeznacza w całości na wydatki rządowe G , nie ma możliwości zaciągania pożyczek ani oszczędzania. Ograniczenie budżetowe rządu jest zatem postaci:

$$\tau w_r(1 - L_s) + P = G. \quad (3)$$

Wydatki rządowe G mogą służyć zaspokojeniu potrzeb wyłącznie rządu (są to tzw. nieużyteczne społecznie wydatki rządowe; ang. *wasteful government expenditure*) lub też przyczyniać się do wytworzenia dóbr i usług publicznych (użyteczne społecznie wydatki rządowe). Wartość parametru $\psi \in \langle 0, 1 \rangle$ określa, jaka część zapłaconych przez gospodarstwa domowe podatków i kar wraca do nich pod postacią użytecznych społecznie wydatków rządowych, a zatem wartość ψ świadczy o efektywności rządu ze społecznego punktu widzenia. Im udział wydatków użytecznych w wydatkach rządowych ogółem jest większy, tym bardziej zadowolone są gospodarstwa domowe i wyższy jest ich poziom użyteczności.

¹² W przypadku obciążeń korupcyjnych $\tau Y_{L_r}(b - 1)$ stanowi łapówki, które należy zapłacić skorumpowanemu urzędnikom.

Zarówno rząd, jak i gospodarstwa domowe, dążą do maksymalizacji własnego zadowolenia, odzwierciedlanego przez ich funkcje użyteczności. Obie z nich przyjmują klasyczną postać funkcji CRRA¹³ (ang. *Constant Relative Risk Aversion*). W przypadku gospodarstw domowych, użyteczność zależy od trzech czynników: konsumpcji C (+), efektywności rządu określanej przez użyteczne społecznie wydatki rządowe ψG (+) oraz od poziomu moralności podatkowej społeczeństwa mL_s (–):

$$U_h = \frac{(C + \psi G - mL_s)^{1-\delta} - 1}{1-\delta}, \quad (4)$$

gdzie $\delta \in (0, \infty) \setminus \{1\}$. Dla $\delta = 1$ funkcja użyteczności CRRA przyjmuje postać:

$$U_h = \ln(C + \psi G - mL_s). \quad (5)$$

Natomiast użyteczność rządu zależy od wielkości wydatków rządowych G (+) oraz od poziomu poparcia społecznego S (+):

$$U_g = \frac{(G + S)^{1-\beta} - 1}{1-\beta}, \quad (6)$$

gdzie $\beta \in (0, \infty) \setminus \{1\}$ ¹⁴. W modelu założono, że jednym z dążeń rządu jest utrzymanie się przy władzy. Musi zatem dbać o swoje szanse na reelekcję, które odzwierciedla funkcja S , mierząca poparcie gospodarstw domowych dla polityki gospodarczej rządu. W modelu przyjęto, że poparcie społeczne zależy od dwóch czynników – stawki opodatkowania¹⁵ τ (–) oraz stopy biurokracji (korupcji) b (–). Funkcja S jest postaci:

$$S = -D(\tau b)^2, \quad (7)$$

gdzie $D \in (0, \infty)$ oraz $\frac{\partial S}{\partial \tau} < 0$ i $\frac{\partial^2 S}{\partial \tau^2} < 0$ oraz $\frac{\partial S}{\partial b} < 0$ i $\frac{\partial^2 S}{\partial b^2} < 0$.

Zmienną decyzyjną gospodarstw domowych jest wielkość szarej strefy L_s , zaś rządu stawka opodatkowania τ . Ostatecznie po wstawieniu równań (1) i (7) do (2), (3), (4), (6) zagadnienie optymalizacyjne przyjmuje postać:

¹³ Jest to funkcja o stałej względnej awersji do ryzyka równej δ , do czego w proponowanym deterministycznym modelu nie należy przywiązywać uwagi. Kluczowe są natomiast jej inne własności, funkcja CRRA jest różniczkowalna, rosnąca i wklęsła (spełnia prawo malejącej użyteczności krańcowej), a parametr $-\delta$ ma wygodną interpretację elastyczności użyteczności krańcowej. Funkcja ta zatem posiada wszystkie niezbędne własności funkcji użyteczności. Z tego powodu jest szeroko stosowana w literaturze ekonomicznej, funkcję typu CRRA przyjmuje też w swoim modelu m.in. Beloded (2005), Orsi i inni (2014).

¹⁴ Podobnie jak w przypadku funkcji użyteczności gospodarstw domowych, dla $\beta = 1$ funkcja użyteczności rządu ma postać funkcji logarytmicznej.

¹⁵ M.in. Kneebone, McKenzie (2001) pokazują, że w latach wyborów wstrzymywane są podwyżki podatków.

$$\left\{ \begin{array}{l} U_g = \frac{[G-D(\tau b)^2]^{1-\beta}-1}{1-\beta} \xrightarrow{\tau} \max, \\ U_h = \frac{(C+\psi G-mL_s)^{1-\delta}-1}{1-\delta} \xrightarrow{L_s} \max, \\ w_r(1-L_s)(1-\tau b)(1+x) + w_s L_s(1+\lambda x) = C + paL_s^2, \\ \tau w_r(1-L_s) + paL_s^2 = G, \end{array} \right. \quad (8)$$

gdzie $L_s, \psi, x, \tau \in \langle 0,1 \rangle, \lambda \in \langle 0,1 \rangle, p \in (0,1), \alpha, \beta, \delta, C, G, D, w_s, w_r, m \in (0,\infty), b \in (1,\infty)$.

3.2. ROZWIĄZANIE ZADANIA OPTIMALIZACYJNEGO

Algorytm rozwiązania zadania optymalizacyjnego (8) jest następujący¹⁶. Najpierw wyznaczone zostanie maksimum funkcji U_h pod względem zmiennej L_s (przy stałej wartości τ), oznaczone jako $\tilde{L}_s$. Następnie, w analogiczny sposób, znalezione zostanie maksimum funkcji U_g pod względem zmiennej τ (przy stałej wartości L_s), oznaczone jako $\tilde{\tau}$. Otrzymane w ten sposób $\tilde{L}_s$ i $\tilde{\tau}$ dadzą tak zwane krzywe najlepszej reakcji¹⁷, które pokazują optymalną odpowiedź na daną strategię innego gracza. W punkcie ich przecięcia otrzymane zostanie rozwiązanie optymalne całego modelu, oznaczane jako L_s^* i τ^* . Będzie to równowaga Nasha, bowiem każda z grup podmiotów wyznacza najlepszą odpowiedź na najlepszą strategię innych jednostek. Wzór na L_s^* otrzymany zostanie zatem wstawiając równanie $\tilde{\tau}$ do $\tilde{L}_s$, τ^* zaś podstawiając równanie $\tilde{L}_s$ do $\tilde{\tau}$.

Poszukując maksimum funkcji U_h i funkcji U_g otrzymano warunki konieczne postaci:

$$\left\{ \begin{array}{l} \frac{\partial U_h}{\partial L_s} = \frac{-w_r(1-\tau b)(1+x) + w_s(1+\lambda x) - 2L_s pa - m - \psi \tau w_r + 2L_s pa \psi}{[w_r(1-L_s)(1-\tau b)(1+x) + w_s(1+\lambda x)L_s - paL_s^2 - mL_s + \psi(\tau w_r(1-L_s) + paL_s^2)]^\delta} = 0, \\ \frac{\partial U_g}{\partial \tau} = \frac{w_r(1-L_s) - 2D\tau b^2}{[\tau w_r(1-L_s) + paL_s^2 - D(\tau b)^2]^\beta} = 0. \end{array} \right. \quad (9)$$

Z pierwszego równania można wyliczyć punkt stacjonarny $\tilde{L}_s$:

$$\tilde{L}_s = \frac{w_s(1+\lambda x) - w_r(1-\tau b)(1+x) - m - \psi \tau w_r}{2pa(1-\psi)}. \quad (10)$$

¹⁶ Sposób rozwiązywania zadania optymalizacyjnego przypomina wyznaczanie równowagi w modelu Cournot, por. m.in. Varian (2005, s. 497–501), Tokarski (2011, s. 256–269).

¹⁷ Przykładowo dla rządu, krzywa najlepszej reakcji pokazuje, jaką stawkę opodatkowania powinien ustalić rząd, by maksymalizować swoją funkcję użyteczności przy danych rozmiarach szarej strefy, którą wybrały gospodarstwa domowe.

Wyznacza on jednocześnie krzywą najlepszej reakcji gospodarstw domowych. Wzór (10) pozwala zatem obliczyć rozmiary szarej strefy, którą powinny wybrać gospodarstwa domowe przy danym poziomie stopy podatkowej τ tak, aby maksymalizować swoją użyteczność. Podobnie, z równania drugiego wyznaczono punkt stacjonarny i krzywą reakcji rządu $\tilde{\tau}$:

$$\tilde{\tau} = \frac{w_r(1-L_s)}{2Db^2}. \quad (11)$$

Oczywiście zachodzi $\frac{\partial \tilde{L}_s}{\partial \tau} > 0$ oraz $\frac{\partial \tilde{\tau}}{\partial L_s} < 0$.

Warunek dostateczny istnienia ekstremum funkcji U_h i U_g w obu przypadkach jest spełniony¹⁸, bowiem $\frac{\partial^2 U_h}{\partial L_s^2}(\tilde{L}_s) < 0$ oraz $\frac{\partial^2 U_g}{\partial \tau^2}(\tilde{\tau}) < 0$, a zatem funkcja U_h posiada w punkcie $\tilde{L}_s$ maksimum lokalne, podobnie funkcja U_g w punkcie $\tilde{\tau}$.

W punkcie przecięcia krzywych reakcji, wstawiając równanie (10) do (11) lub (11) do (10), otrzymujemy optymalne rozmiary szarej strefy L_s^* i stopy podatkowej τ^* , czyli równowagę Nasha modelu postaci:

$$\begin{cases} L_s^* = \frac{2Db^2[w_s(1+\lambda x) - w_r(1+x) - m] + w_r^2[b(1+x) - \psi]}{4pa(1-\psi)Db^2 + w_r^2[b(1+x) - \psi]}, \\ \tau^* = \frac{2pa(1-\psi)w_r - w_s w_r(1+\lambda x) + w_r^2(1+x) + m w_r}{4pa(1-\psi)Db^2 + w_r^2[b(1+x) - \psi]}, \end{cases} \quad (12)$$

przy czym zakładamy, że $L_s^* \in \langle 0,1 \rangle$ i $\tau^* \in \langle 0,1 \rangle$.¹⁹

4. STATYKA PORÓWNAWCZA I WNIOSKI

Analiza wrażliwości polega na określeniu, co się stanie z rozwiązaniem optymalnym modelu L_s^* i τ^* na skutek zmiany wartości jednego z parametrów: λ , w_s , w_r , m , D , p , a , b , x , ψ . Pozwala to porównać dwie prawie identyczne gospodarki, które róż-

¹⁸ Obliczenia te ze względu na ich rozmiar zostały pominięte w niniejszej pracy. Są jednak dostępne u autorki na życzenie.

¹⁹ Można zauważyć, że przedziały zmienności niektórych parametrów dopuszczają sytuację, gdy równowagowe wartości obu zmiennych wykraczają poza przedział $\langle 0,1 \rangle$, np. dzieje się tak gdy parametr m jest bardzo duży w porównaniu do innych parametrów. Wówczas $L_s^* < 0$ a $\tau^* > 1$. Tak duża wartość parametru oznacza jednak wyjątkowo niską skłonność gospodarstw domowych do uchylania się od podatków, co oczywiście oznacza, *de facto*, że szara strefa nie będzie w ogóle występowała w gospodarce ($L_s^* = 0$). Zatem w przypadku gdy zmienne decyzyjne wykraczają będą poza założony przedział zmienności przyjmować będziemy dla nich wartości skrajne (0 lub 1). Oczywiście zastosowanie różnych skali wartości (różnych jednostek) dla poszczególnych parametrów pozwala ustalić wartości zmiennych decyzyjnych na ekonomicznie akceptowalnych poziomach.

nią się wyłącznie wartością jednego z tych parametrów. Po obliczeniu odpowiednich pochodnych cząstkowych otrzymano zależności zaprezentowane w tabeli 1.

Tabela 1.

Analiza wrażliwości stanu równowagi Nasha zaproponowanego modelu na skutek zmiany wartości jego poszczególnych parametrów

		Parametry						
		λ	w_s	m	D	p	a	b
Zmienne	L_s^*	$\frac{\partial L_s^*}{\partial \lambda} > 0$	$\frac{\partial L_s^*}{\partial w_s} > 0$	$\frac{\partial L_s^*}{\partial m} < 0$	$\frac{\partial L_s^*}{\partial D} < 0$	$\frac{\partial L_s^*}{\partial p} < 0$	$\frac{\partial L_s^*}{\partial a} < 0$	$\frac{\partial L_s^*}{\partial b} < 0$
	τ^*	$\frac{\partial \tau^*}{\partial \lambda} < 0$	$\frac{\partial \tau^*}{\partial w_s} < 0$	$\frac{\partial \tau^*}{\partial m} > 0$	$\frac{\partial \tau^*}{\partial D} < 0$	$\frac{\partial \tau^*}{\partial p} > 0$	$\frac{\partial \tau^*}{\partial a} > 0$	$\frac{\partial \tau^*}{\partial b} < 0$

Źródło: obliczenia własne.

Analiza wyników zaprezentowanych w Tabeli 1 prowadzi do następujących wniosków:

- 1) Gdy dostępność usług instytucji państwowych w szarej strefie wzrośnie (wzrost λ), to spodziewać się należy zwiększenia rozmiarów szarej gospodarki. Spowodowane jest to wzrostem dochodów z działalności nierejestrowanej, co zachęca gospodarstwa domowe do ukrywania swojej aktywności gospodarczej. W rezultacie, w celu zatrzymania ucieczki gospodarstw domowych do szarej strefy rząd obniży stawkę opodatkowania. Podobny mechanizm występuje w przypadku wzrostu wynagrodzenia w szarej strefie w_s .
- 2) Wzrost poziomu *tax morale* społeczeństwa (m) przyczynia się do zwiększenia dysużyteczności gospodarstw domowych z tytułu wykonywania działalności nierejestrowanej i uchylania się od opodatkowania. Jest to związane z dyskomfortem moralnym oraz presją społeczną odczuwaną przez gospodarstwa domowe. W rezultacie, prowadzi to do zmniejszenia ich aktywności nierejestrowanej. Rząd może wówczas pozwolić sobie na podniesienie stopy opodatkowania, w celu zwiększenia przychodów budżetowych.
- 3) Podniesienie wielkości kary a za działalność nierejestrowaną (analogicznie – efektywności działania organów kontroli p), zmniejsza oczekiwane zyski gospodarstw domowych z pracy w szarej strefie. Obniża to atrakcyjność tego sektora i w rezultacie prowadzi do spadku rozmiarów szarej strefy, co skutkuje podniesieniem opodatkowania przez rząd.
- 4) Zwiększenie opłat biurokratycznych b (żądań łapówkarskich skorumpowanych urzędników), prowadzi do wzrostu niezadowolenia gospodarstw domowych. Rząd, który zabiega o reelekcję, zmuszony jest obniżyć opodatkowanie. Korzyści płynące ze spadku opodatkowania przewyższają straty wynikające z konieczności płacenia

wyższych opłat biurokratycznych w strefie oficjalnej – a zatem szara gospodarka zmaleje. Podobny mechanizm zaobserwować można w przypadku spadku wartości wskaźnika D .

W przypadku x , w_s i ψ znaki pochodnych zależą od kombinacji wartości innych parametrów. Jest to związane z jednoczesnym działaniem kilku efektów o przeciwnym kierunku wpływu. Wzrost wartości każdego z wymienionych parametrów związany jest ze zwiększeniem atrakcyjności prowadzenia działalności gospodarczej w sferze rejestrowanej oraz jednoczesną odpowiedzią rządu w postaci zmiany stawki opodatkowania. Który z tych efektów ostatecznie przeważy, zależy od konkretnych wartości pozostałych parametrów gospodarki. Przykładowo, poprawienie stanu jakości państwa x , zwielokrotnia dochody uzyskiwane w obu sektorach, ale ich wzrost w szarej strefie jest ograniczony przez parametr λ . Tym samym praca nierejestrowana staje się relatywnie mniej atrakcyjna. Odpływ do sfery oficjalnej jest hamowany przez jednoczesną podwyżkę stopy podatkowej. To, które z tych efektów przeważy zależy od relacji pomiędzy płacami w obu sektorach oraz od dostępności dóbr i usług publicznych w szarej strefie λ .

5. ZAKOŃCZENIE

W niniejszej pracy przedstawiono autorski model kształtowania się szarej strefy gospodarki, która powstaje jako wynik świadomej kalkulacji oczekiwanych zysków i strat gospodarstw domowych na skutek określonej polityki gospodarczej rządu. Wśród korzyści z aktywności gospodarczej w szarej strefie uwzględniono możliwość uniknięcia opodatkowania i paraopodatkowania oraz zwolnienie z konieczności przestrzegania regulacji prawa pracy i ponoszenia związanych z tym kosztów biurokratycznych (korupcyjnych). Działalność w szarej gospodarce nie jest jednak wolna od strat – jednostki muszą liczyć się z możliwością przyłapania i koniecznością poniesienia dodatkowych kar oraz z ograniczonym dostępem do dóbr i usług publicznych. Co więcej, praca nierejestrowana może wzbudzać wątpliwości moralne i prowadzić do spadku zadowolenia jednostek. Rząd natomiast prowadzi politykę fiskalną państwa (ustala stawkę opodatkowania) oraz ustanawia organy kontrolne czuwające nad sposobem prowadzenia aktywności gospodarczej. Jednocześnie troszczy się o swoje poparcie społeczne i szanse na reelekcję. Gospodarstwa domowe i rząd podejmują decyzje tak, by maksymalizować własne funkcje użyteczności, i, w wyniku ich wzajemnych interakcji, zostaje ustalony równowagowy poziom szarej strefy i opodatkowania w modelowanej gospodarce.

Badanie wrażliwości stanu równowagi Nasha prezentowanego modelu pozwala określić, jak będą różniły się optymalne poziomy szarej strefy i stawki opodatkowania w gospodarkach, które różnią się jedynie wartością jednego konkretnego parametru. Przykładowo, w gospodarce o wyższej stopie korupcji wśród urzędników państwowych (i identycznych pozostałych makroparametrach) spodziewać się należy niższego

optymalnego poziomu szarej gospodarki i stawki opodatkowania. Na podstawie analizy wrażliwości można też wysnuć wnioski dla polityki gospodarczej państwa, które dąży do ograniczenia szarej aktywności. Oprócz polityki fiskalnej, rząd może posłużyć się szeregiem innych, często długookresowych, działań: ograniczyć dostępność dóbr i usług publicznych w szarej strefie, akcentować jej nieetyczny charakter budując kręgosłup moralny społeczeństwa, zwiększyć efektywność organów ścigania i podnieść stawki kar, ograniczyć biurokrację (korupcję), próbować zmniejszyć wynagrodzenia w strefie nieoficjalnej, zadbać o wzrost poparcia dla swojej polityki gospodarczej.

Proponowany model może ulec dalszym ulepszeniom, np. poszerzeniu o bardziej realistyczny rynek pracy, bezrobocie i transfery społeczne. W następnych badaniach planowana jest kalibracja parametrów modelu dla wybranych krajów na świecie.

LITERATURA

- Amaral P. S., Quintin E., (2006), A Competitive Model of the Informal Sector, *Journal of Monetary Economics*, 53 (7), 1541–1553.
- Aruoba B., (2009), Informal Sector, Government Policy and Institutions, in 2010 Meeting Papers, No. 324, Society for Economic Dynamics (working paper).
- Beloded O., (2005), Shadow Economy of Ukraine: the Case of Financial Constraints (Doctoral dissertation, Economics Education and Research Consortium).
- Bilotkach V., (2006), A Tax Evasion – Bribery Game: Experimental Evidence from Ukraine, *European Journal of Comparative Economics*, 3, 31–49.
- Buehn A., Dell’Anno R., Schneider F., (2012), Fiscal Illusion and the Shadow Economy: Two Sides of the Same Coin?, Working Paper, 39944.
- Choi J. P., Thum M., (2005), Corruption and the Shadow Economy, *International Economic Review*, 46 (3), 817–836.
- Cichocki S., (2009), Shadow Economy and Its Relations with Tax System and State Budget in Poland, *Ekonomia*, 24, 94–111.
- Cichocki S., Tyrowicz J., (2011), Determinanty zatrudnienia nierejestrowanego w Polsce w okresach wysokiej i niskiej koniunktury gospodarczej, *Gospodarka Narodowa*, 3, 1–27.
- Dreher A., Schneider F., (2010), Corruption and the Shadow Economy: An Empirical Analysis, *Public Choice*, 144 (1), 215–238.
- Friedman E., Johnson S., Kaufmann D., Zoido-Lobaton P., (2000), Dodging the Grabbing Hand: the Determinants of Unofficial Activity in 69 Countries, *Journal of Public Economics*, 76 (3), 459–493.
- Giza-Poleszczuk A., (2009), Jak zmierzyć zjawisko pracy nierejestrowanej?, *Polityka Społeczna*, (10), 20–23.
- Gołębiowski G., (2007), Zjawisko szarej strefy z uwzględnieniem gospodarki polskiej, *Współczesna Ekonomia*, 1 (1), 17–28.
- Heinemann F., Schneider F., (2011), Religion and the Shadow Economy, *ZEW-Centre for European Economic Research Discussion Paper*, 11-038.
- Ihrig J., Moe K. S., (2004), Lurking in the Shadows: The Informal Sector and Government Policy, *Journal of Development Economics*, 73 (2), 541–557.
- Jędrzejowicz P., (1995), Problem określenia polityki skarbowej wobec zjawiska nierejestrowanych dochodów z pracy, w: Grabowski M., (red.), *Szara strefa w transformacji gospodarki*, Gdańsk, 19–38.
- Johnson S., Kaufmann D., Shleifer A., (1997), The Unofficial Economy in Transition, *Brookings Papers on Economic Activity*, 159–239.

- Kabaj M., (2009), Praca nierejestrowana we współczesnej literaturze ekonomicznej, *Polityka Społeczna*, (10), 3–10.
- Kneebone R. D., McKenzie K. J., (2001), Electoral and Partisan Cycles in Fiscal Policy: An Examination of Canadian Provinces, *International Tax and Public Finance*, 8 (5–6), 753–774.
- Kraciuk J., (2013), Szara strefa w Europie, *Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie. Problemy Rolnictwa Światowego*, 13 (1).
- Malaczewska P., (2014), Analiza zjawiska szarej strefy jako gry niekooperacyjnej, w: Jurek W., (red.), *Matematyka i informatyka na usługach ekonomii. Rozważania ogólne. Modele*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, 84–96.
- Mról B., (2002), *Gospodarka nieoficjalna w systemie ekonomicznym*, Monografie i Opracowania, Szkoła Główna Handlowa, 509.
- Orsi R., Raggi D., Turino F., (2014), Size, Tend, and Policy Implications of the Underground Economy, *Review of Economic Dynamics*, 17 (3), 417–436.
- Raczkowski K., Schneider F., (2013), Sfera nieoficjalna w gospodarce, *Infos*, nr 21 (158), Biuro Analiz Sejmowych.
- Rosser J. B., Rosser M. V., Ahmed E., (2000), Income Inequality and the Informal Economy in Transition Economies, *Journal of Comparative Economics*, 28 (1), 156–171.
- Salahodjaev R., (2015), Intelligence and Shadow Economy: A Cross-country Empirical Assessment, *Intelligence*, 49, 129–133.
- Schneider F., Enste D., (2000), Shadow Economies: Size, Causes, and Consequences, *Journal of Economic Literature*, 38 (1), 77–114.
- Schneider F., Williams C., (2013), *The Shadow Economy*, The Institute of Economic Affairs, London.
- Tokarski T., (2011), *Ekonomia matematyczna: modele mikroekonomiczne*, Polskie Wydawnictwo Ekonomiczne.
- Varian H. R., (2005), *Mikroekonomia: kurs średni – ujęcie nowoczesne*, Wydawnictwo Naukowe PWN.
- Walewski M., (2009), Zatrudnienie nierejestrowane od strony pracowników, *Polityka Społeczna*, (10), 8–11.
- Wdrożenie Europejskiego Systemu Rachunków Narodowych i Regionalnych w Unii Europejskiej (ESA2010) do polskich rachunków narodowych. Zmiany metodologiczne oraz ich wpływ na główne agregaty makroekonomiczne, Notatka informacyjna, GUS, Warszawa, 2014, 12–13.

MECHANIZMY KSZTAŁTOWANIA SIĘ SZAREJ STREFY GOSPODARKI. MODEL TEORETYCZNY

Streszczenie

Przedstawiony zostanie model teoretyczny kształtowania się szarej strefy gospodarki. Uwzględnione w nim będą zarówno czynniki typowo ekonomiczne, takie jak obciążenia podatkowe i parapodatkowe czy wysokość kar za działalność nierejestrowaną, oraz czynniki kształtujące otoczenie prawno-organizacyjne działalności gospodarczej, takie jak poziom biurokracji (korupcji), jakość instytucji publicznych, czy moralność podatkowa społeczeństwa. W proponowanym modelu szara strefa powstaje jako wynik interakcji pomiędzy podmiotami gospodarczymi – rządem i gospodarstwami domowymi. Wyznaczony zostanie poziom szarej strefy będący równowagą Nasha, po czym przeprowadzona zostanie jego statyka porównawcza.

Słowa kluczowe: szara strefa gospodarki, obciążenia podatkowe, jakość instytucji publicznych, *tax morale*, korupcja, biurokracja, statyka porównawcza, równowaga Nasha

DRIVING FORCES OF THE SHADOW ECONOMY. THEORETICAL MODEL

Abstract

We present theoretical model of shadow economy. The model includes not only economic factors, such as tax burden or fines for unregistered activity, but also legal and organizational background factors, such as level of bureaucracy (corruption), quality of public sector services and institutions, or tax morale. In the proposed model shadow economy arises as a result of interactions between economic entities – government and households. The Nash Equilibrium level of shadow economy is derived and comparative statics analysis is presented.

Keywords: shadow economy, tax burden, institutions, tax morale, corruption, bureaucracy, comparative statics, Nash equilibrium

PAWEŁ DYKAS¹, TOMASZ MISIAK²NEOKLASYCZNY MODEL WZROSTU GOSPODARCZEGO
Z SINUSOIDALNYMI INWESTYCJAMI³

1. WPROWADZENIE

Celem prezentowanego artykułu jest próba rozszerzenia neoklasycznego modelu wzrostu gospodarczego Solowa poprzez uwzględnienie fluktuacji inwestycji. Autorzy uchylając założenie o stałości stóp inwestycji wprowadzają funkcję inwestycji zależną sinusoidalnie od czasu. Przyjęcie sinusoidalnej funkcji inwestycji tłumaczyć można tym, iż inwestycje (podobnie jak produkcja) w dużej mierze zależne są od koniunktury gospodarczej, która podlega okresowym fluktuacjom. W rozważaniach teoretycznych autorzy wprowadzają ponadto pojęcia cyklicznej oraz gładkiej ścieżki czasowej kapitału oraz produktu na jednostkę efektywnej pracy oraz wyznaczają ich względne odchylenia. Za pomocą tych narzędzi identyfikują wpływ inwestycji zależnych sinusoidalnie od czasu na ścieżki wzrostu kapitału oraz produktu w analizowanym modelu.

W części empirycznej dokonano kalibracji parametrów otrzymanego modelu oraz symulacji numerycznej. Bazując na danych panelowych dla krajów grupy UE15 w latach 2000–2013 w pierwszej kolejności oszacowano parametr α (elastyczność produkcji względem kapitału) na poziomie 0,349, którego wartość przyjęto do dalszych analiz numerycznych. W drugim etapie dokonano symulacji numerycznych skalibrowanych odchyleń cyklicznych od gładkich ścieżek czasowych kapitału oraz produktu na jednostkę efektywnej pracy. W analizach numerycznych rozważano wpływ różnych stóp inwestycji (15%, 20%, 25%) oraz okresów wahań cyklicznych (4 lub 10 lat) na kształtowanie się owych odchyleń. Analiz numerycznych dla gospodarek grupy UE15 dokonano w 100 letnim horyzoncie czasowym.

Struktura artykułu przedstawia się następująco. W części drugiej zawarto przegląd literatury dotyczący inwestycji oraz różnych rozwinięć neoklasycznego modelu wzrostu Solowa. W części trzeciej przedstawiono główne założenia oraz wyznaczono

¹ Uniwersytetu Jagielloński, Wydział Zarządzania i Komunikacji Społecznej, Katedra Ekonomii Matematycznej, ul. Prof. Stanisława Łojasiewicza 4, 30-348 Kraków, Polska.

² Politechnika Rzeszowska, Wydział Zarządzania, Katedra Ekonomii, Al. Powstańców Warszawy 8, 35-959 Rzeszów, Polska, autor prowadzący korespondencję – e-mail: tmisiak@prz.edu.pl.

³ Autorzy pragną podziękować anonimowym Recenzentom za cenne merytoryczne uwagi zgłoszone do pierwotnej wersji niniejszego artykułu. Oczywiście, cała odpowiedzialność za ewentualne mankamenty spada wyłącznie na autorów.

ścieżki wzrostu zarówno kapitału jak i produktu na jednostkę efektywnej pracy a także wyznaczono względne odchylenia cyklicznych od gładkich ścieżek wzrostu w analizowanym modelu z sinusoidalnymi inwestycjami. W czwartej części artykułu zawarto kalibrację parametrów otrzymanego modelu oraz przeprowadzono symulacje numeryczne w zaproponowanych wariantach. W ostatniej części artykułu znajduje się podsumowanie prowadzonych rozważań oraz ważniejsze wnioski.

2. PRZEGLĄD LITERATURY

Wzrost gospodarczy jest procesem długookresowym, prowadzącym do powiększania potencjału produkcyjnego a szeroko rozumiane inwestycje stanowią jeden z kluczowych czynników tego procesu. Przyczynia się on również do zaspokojenia potrzeb społeczeństwa. W literaturze przedmiotu rozróżnia się podział modeli wzrostu gospodarczego między innymi na modele keynesowskie, neoklasyczne, modele wzrostu endogenicznego oraz modele realnego cyklu koniunkturalnego (por. Tokarski, 2005). Można również, ze względu na sposób ustalania długookresowej stopy wzrostu, wyróżnić modele wzrostu: egzogenicznego, endogenicznego oraz heterodoksyjne (Malaga, 2015). Jednak zdaniem Malagi współczesna teoria wzrostu gospodarczego w niewielkim stopniu korzysta z dorobku keynesowskiej teorii wzrostu a u jej podstaw nadal pozostają modele neoklasyczne (Malaga, 2009). Koncepcja zaproponowana przez Solowa (1956) dała niejako przyczynek do rozwoju neoklasycznych modeli wzrostu gospodarczego. Solow zakładał m.in. stałe przychody ze skali produkcji oraz malejącą krańcową produktywność kapitału (Solow, 1956). Model Solowa nazywany również modelem Solowa-Swana posiada wiele zalet jak i wad. Do podstawowych zalet najczęściej zalicza się: łatwość wykorzystania matematycznych funkcji produkcji, możliwość badania udziału czynników produkcji we wzroście – zwłaszcza postępu technicznego oraz produktywności pracy, a także jego użyteczność do analiz procesów konwergencji realnej. Główne słabości modelu to przede wszystkim egzogeniczne traktowanie całości postępu technicznego, pominięcie takich czynników jak kapitał ludzki oraz kwestii rosnących korzyści skali co może powodować, że w przypadku niektórych inwestycji bodźce do inwestowania nie wygasają (szerzej na temat wad oraz krytyki modelu Solowa patrz np. Milo i inni, 2006 lub McCombie, 2000, 2001).

Mimo wielu słów krytyki w literaturze można znaleźć wiele opracowań dotyczących samego modelu Solowa (1956) jak i jego uogólnień. Mankiw i inni (1992) uogólnili model Solowa. Zasadnicza różnica między oryginalnym modelem Solowa, a jego rozszerzeniem zaproponowanym przez Mankiwa, Romera, Weila polega na tym, iż autorzy uwzględniają dodatkowy zasób kapitału, a mianowicie kapitał ludzki (Mankiw i inni, 1992; Romer, 2000). W roku 1996 Nonneman oraz Vanhoudt dokonali kolejnego uogólnienia modelu Solowa. W swojej pracy autorzy rozważają dowolną, skończoną liczbę dóbr kapitałowych, będących czynnikami produkcji (Nonneman, Vanhoudt, 1996). Kolejne rozszerzenia znaleźć można w artykule Dykasa, Sulimy i Tokarskiego, autorzy dowodzą istnienie stabilnego punktu stacjonarnego, w którym

model ten znajduje się w stanie długookresowej równowagi (Dykas i inni, 2008). Ponadto w opracowaniu Dykasa, Edigariana oraz Tokarskiego szczególny przypadek funkcji produkcji, jaką jest funkcja produkcji Cobba-Douglasa, zastąpiony został uogólnioną, neoklasyczną funkcją produkcji (Dykas i inni, 2011).

Inne rozszerzenie modelu Solowa można znaleźć w pracy Mroczek i inni (2014). Autorzy rozpatrują model wzrostu gospodarczego uwzględniając efekt grawitacyjny, będący nawiązaniem do prawa grawitacji Newtona. Zakładają, że regiony wzajemnie oddziałują na siebie gospodarczo oraz, iż ten efekt jest proporcjonalny do potencjału (np. PKB *per capita*, kapitału rzeczowego *per capita* itp.) danego regionu i odwrotnie proporcjonalny do kwadratu odległości między nimi (rozumianej jako fizyczna odległość między stolicami regionów). W opracowaniu wykazano również, że model ten ma stabilny (w sensie Lapunowa) punkt stacjonarny jest on zatem punktem długookresowej równowagi modelu (Mroczek i inni, 2014).

Większość dotychczasowych uogólnień modelu Solowa skoncentrowanych było na wprowadzeniu różnych rodzajów kapitału bądź analizie istnienia i jednoznaczności stanów stacjonarnych. W niniejszym artykule zaproponowano inny kierunek uogólnienia modelu Solowa, będący próbą uwzględnienia fluktuacji inwestycji. W teorii ekonomii zauważyć można bowiem utrzymujący się stan dychotomii między teorią wzrostu gospodarczego oraz teorią fluktuacji gospodarczych, co stanowi jeden z kluczowych dylematów wzrostu i rozwoju (Małaga, 2015). Z analizy danych takich wielkości agregatowych jak PKB, konsumpcji, zatrudniania oraz inwestycji wynika, iż ich trajektorie wzrostu nie są obrazem równomiernych ścieżek wzrostu lecz charakteryzują się okresowymi wahaniami. Przy czym okazuje się, iż to właśnie inwestycje są bardzo wrażliwe na zmiany sytuacji koniunkturalnej. Rzecz jasna: przyczyny owych wahań, amplitudy wahań, przebieg cykli, sprzężenia zwrotne, efekty mnożnikowe oraz akceleracji są przedmiotem badań szeroko rozumianej teorii wahań koniunkturalnych, której pogłębiona analiza jest zdecydowanie bogatsza niż przyjęte w niniejszym artykule założenie o cyklicznych waniach stóp inwestycji. (por. np. Maussner, 1994; Romer, 2000). Jednak kwestie te, ze względu na zakres ww. zagadnień oraz przyjęty cel artykułu nie będą szerzej analizowane.

3. MODEL

W rozważanym modelu przyjmuje się następujące założenia:

- 1) Proces produkcyjny opisuje funkcja produkcji Cobba-Douglasa (1928)⁴ dana wzorem (por. też Żółtowska, 1997 lub Tokarski, 2009; 2011)⁵:

⁴ Warto wspomnieć, iż analizowana funkcja produkcji została sformułowana przez Knuta Wicksella (Wicksell, 1893), a empirycznie przetestowana przez Paula H. Douglasa oraz Charlesa W. Cobba (Cobb, Douglas, 1928).

⁵ O wszystkich występujących dalej zmiennych makroekonomicznych zakłada się, iż są różniczkowalnymi funkcjami czasu $t \geq 0$. Zapis $x(t)$ oznaczał będzie dalej wartość zmiennej x w momencie t ,

$$Y(t) = (K(t))^\alpha (E(t))^{1-\alpha}, \quad (1)$$

gdzie $\alpha \in (0;1)$. Y oznacza strumień produktu, K , E – nakłady kapitału i tzw. jednostki efektywnej pracy, zaś α – elastyczność wytworzonego w gospodarce produktu względem nakładów kapitału⁶.

- 2) Przyrost zasobu kapitału, podobnie jak w modelu Solowa-Swana (1956), opisuje równanie różniczkowe postaci:

$$\dot{K}(t) = s(t)Y(t) - \delta K(t), \quad (2)$$

gdzie $\delta \in (0;1)$. $s(t) \in (0;1)$ oznacza stopę inwestycji w momencie t , δ zaś – stopę deprecjacji kapitału. O ile jednak w modelu Solowa zakłada się, że s i δ są stałe w czasie, o tyle w prezentowanym modelu wzrostu gospodarczego przyjmuje się, iż stopa inwestycji podlega cyklicznym, sinusoidalnym fluktuacjom opisanym przez równanie (3).

- 3) Stopę inwestycji w każdym momencie t określa sinusoida o stałej amplitudzie i okresie wahań dana wzorem:

$$s(t) = \bar{s} + \theta \sin\left(\frac{2\pi}{\omega}t\right), \quad (3)$$

gdzie $\bar{s}$, $(\bar{s} \pm \theta) \in (0;1)$ oraz θ , $\omega > 0$. $\bar{s}$ w równaniu (3) oznacza przeciętną stopę inwestycji w cyklu inwestycyjnym, θ – amplitudę wahań cyklicznych stopy inwestycji, natomiast ω – okres owych wahań.

- 4) Jednostki efektywnej pracy E rosną według stopy wzrostu $g + n$, gdzie $g > 0$ jest stopą egzogenicznego postępu technicznego w sensie Harroda, zaś $n > 0$ stopą wzrostu liczby pracujących. Dlatego też⁷:

$$E(t) = E_0 e^{(g+n)t} \Rightarrow \dot{E}(t)/E(t) = g + n, \quad (4)$$

gdzie $E_0 > 0$ oznacza liczbę jednostek efektywnej pracy (rozumianej jako zasób wiedzy naukowo-technicznej ważony liczbą pracujących) w momencie $t = 0$.

Niech $y_E = Y/E$ oraz $k_E = K/E$ oznaczają (odpowiednio) strumień produktu i zasób kapitału na jednostkę efektywnej pracy, zaś $y = Y/L$ oraz $k = K/L$ – wydajność pracy i techniczne uzbrojenie pracy. Wówczas:

$$y(t) = A_0 e^{gt} y_E(t), \quad (5)$$

zaś $\dot{x}(t) = dx/dt$ – pochodną zmiennej x po czasie t , czyli (ekonomicznie rzecz biorąc) przyrost wartości owej zmiennej w momencie t .

⁶ A zatem $1 - \alpha$ jest wówczas udziałem nakładów efektywnej pracy E w produkcji Y .

⁷ Ponieważ jednostki efektywnej pracy E w modelu Solowa są iloczynem egzogenicznej technologii A i liczby pracujących L , zatem zależność (4) można zapisać również następująco: $E(t) = A_0 e^{gt} L_0 e^{nt}$, gdzie $A_0 > 0$ i $L_0 > 0$ oznacza (odpowiednio) zasób technologii i liczbę pracujących w momencie $t = 0$.

$$k(t) = A_0 e^{gt} k_E(t), \quad (6)$$

$$Y(t) = L_0 e^{nt} y(t) \quad (7)$$

i:

$$K(t) = L_0 e^{nt} k(t). \quad (8)$$

Dzieląc funkcję produkcji (1) przez jednostki efektywnej pracy $E > 0$ dochodzi się do funkcji produkcji w postaci intensywnej określonej przez równanie:

$$y_E(t) = (K(t) / E(t))^\alpha = (k_E(t))^\alpha. \quad (9)$$

Funkcja (9) opisuje relacje pomiędzy nakładami kapitału na jednostkę efektywnej pracy k_E a wielkością produkcji na jednostkę owej pracy, czyli y_E .

Ponieważ kapitał na jednostkę efektywnej pracy można zapisać następująco: $k_E = K/E$, zatem (po zróżniczkowaniu powyższej zależności względem czasu t) mamy:

$$\dot{k}_E(t) = \frac{\dot{K}(t)E(t) - K(t)\dot{E}(t)}{(E(t))^2} = \frac{\dot{K}(t)}{E(t)} - \frac{\dot{E}(t)}{E(t)} k_E(t),$$

co wraz ze związkami (2–4) daje:

$$\dot{k}_E(t) = \left(\bar{s} + \theta \sin\left(\frac{2\pi}{\omega} t\right) \right) y_E(t) - \mu k_E(t), \quad (10)$$

przy czym $\mu = \delta + g + n > 0$ oznacza stopę ubytku kapitału na jednostkę efektywnej pracy. Równanie różniczkowe (10) jest odpowiednikiem równania Solowa w rozważanym w pracy modelu wzrostu gospodarczego.

Po wstawieniu funkcji produkcji w postaci intensywnej (9) do zależności (10) dochodzi się do następującego równania różniczkowego Bernoulliego:

$$\dot{k}_E(t) = \left(\bar{s} + \theta \sin\left(\frac{2\pi}{\omega} t\right) \right) (k_E(t))^\alpha - \mu k_E(t). \quad (11)$$

Równanie różniczkowe (11) posiada całkę trywialną $k_E(t) = 0$ dla każdego $t \geq 0$ oraz pewną rodzinę całek nietrywialnych, którą niebawem znajdziemy⁸.

Z równania (11) wynika, że dla $k_E > 0$ zachodzi związek:

$$(k_E(t))^{-\alpha} \dot{k}_E(t) = \bar{s} + \theta \sin\left(\frac{2\pi}{\omega} t\right) - \mu (k_E(t))^{1-\alpha}. \quad (12)$$

⁸ Całka trywialna (jako nieciekawa zarówno z matematycznego, jak i ekonomicznego punktu widzenia) będzie dalej pomijana. Nietrywialna zaś całka owego równania wyznaczała będzie ścieżkę czasową (lub ścieżkę wzrostu) kapitału na jednostkę efektywnej pracy.

Niech dane będzie podstawienie Bernoulliego postaci:

$$q(t) = (k_E(t))^{1-\alpha}, \quad (13)$$

co powoduje, że:

$$\frac{\dot{q}(t)}{1-\alpha} = (k_E(t))^{-\alpha} \dot{k}_E(t). \quad (14)$$

Wstawiając związki (13)–(14) do równania różniczkowego (12) sprowadza się je do równania różniczkowego liniowego niejednorodnego (względem q) danego wzorem:

$$\frac{\dot{q}(t)}{1-\alpha} = \bar{s} + \theta \sin\left(\frac{2\pi}{\omega} t\right) - \mu q(t),$$

które można przekształcić do zależności:

$$\dot{q}(t) = (1-\alpha)\bar{s} + (1-\alpha)\theta \sin\left(\frac{2\pi}{\omega} t\right) - (1-\alpha)\mu q(t). \quad (15)$$

Zapiszmy równanie jednorodne powstałe z równania (15) jako:

$$\dot{q}(t) = -(1-\alpha)\mu q(t), \quad (16)$$

Rozdzielając zmienne otrzymujemy

$$\ln(q(t)) = (1-\alpha)\mu t + C, \quad C \in R,$$

co daje:

$$q(t) = C \cdot e^{-(1-\alpha)\mu t}, \quad (17)$$

a stąd po uzmiennieniu stałej oraz zróżniczkowaniu względem czasu t mamy:

$$\dot{q}(t) = \dot{C}(t)e^{-(1-\alpha)\mu t} - C(t)(1-\alpha)\mu e^{-(1-\alpha)\mu t}. \quad (18)$$

Wstawiając zależności (17) oraz (18) do związku (15) otrzymuje się:

$$C(t) = \int e^{(1-\alpha)\mu t} (1-\alpha)\bar{s} dt + \int e^{(1-\alpha)\mu t} (1-\alpha)\theta \sin\left(\frac{2\pi}{\omega} t\right) dt. \quad (19)$$

Stąd oraz z równania (17) otrzymujemy

$$q(t) = \frac{\bar{s}}{\mu} + (1-\alpha)\theta \frac{(1-\alpha)\mu \sin\left(\frac{2\pi}{\omega} t\right) - \frac{2\pi}{\omega} \cos\left(\frac{2\pi}{\omega} t\right)}{(1-\alpha)^2 \mu^2 + \frac{4\pi^2}{\omega^2}} + F e^{-(1-\alpha)\mu t}. \quad (20)$$

Natomiast z równań (13) i (20) płynie wniosek, iż całkę równania różniczkowego (11) określa związek:

$$k_E(t) = \left(\frac{\bar{s}}{\mu} + (1-\alpha)\theta \frac{(1-\alpha)\mu \sin\left(\frac{2\pi}{\omega}t\right) - \frac{2\pi}{\omega} \cos\left(\frac{2\pi}{\omega}t\right)}{(1-\alpha)^2\mu^2 + \frac{4\pi^2}{\omega^2}} + Fe^{-(1-\alpha)\mu t} \right)^{\frac{1}{1-\alpha}}. \quad (21)$$

Stała całkowania F w równaniu (21) należy dobrać tak, by spełniony był warunek $k_E(0) = k_{E0} > 0$, gdzie k_{E0} jest zasobem kapitału na jednostkę efektywnej pracy w momencie $t = 0$. Dlatego też:

$$F = k_{E0}^{1-\alpha} + \frac{2\pi(1-\alpha)\theta}{(1-\alpha)^2\mu^2\omega + \frac{4\pi^2}{\omega}} - \frac{\bar{s}}{\mu}. \quad (22)$$

Równania (21)–(22) wyznaczają ścieżkę wzrostu kapitału na jednostkę efektywnej pracy w modelu Solowa z sinusoidalnymi stopami inwestycji.

Wstawiając ścieżkę czasową technicznego uzbrojenia pracy z równania (21) do funkcji produkcji w postaci intensywnej (9) mamy:

$$y_E(t) = \left(\frac{\bar{s}}{\mu} + (1-\alpha)\theta \frac{(1-\alpha)\mu \sin\left(\frac{2\pi}{\omega}t\right) - \frac{2\pi}{\omega} \cos\left(\frac{2\pi}{\omega}t\right)}{(1-\alpha)^2\mu^2 + \frac{4\pi^2}{\omega^2}} + Fe^{-(1-\alpha)\mu t} \right)^{\frac{\alpha}{1-\alpha}}. \quad (23)$$

Równanie (23) określa zaś ścieżkę czasową produktu na jednostkę efektywnej pracy w rozważanym tu modelu wzrostu gospodarczego.

Wprowadźmy teraz dwa następujące pojęcia. Przez cykliczną ścieżkę czasową (lub ścieżkę wzrostu) pewnej zmiennej makroekonomicznej rozumiana będzie taka ścieżka wzrostu zmiennej x , która kształtuje się przy amplitudzie θ i okresie ω wahań stóp inwestycji większych od zera. Cykliczna ścieżka czasowa zmiennej x oznaczana będzie dalej przez $x(t)$. Natomiast przez gładką (acykliczną) ścieżkę wzrostu rozumiana będzie ścieżka wzrostu przy amplitudzie $\theta = 0$. Gładką ścieżkę wzrostu zmiennej x oznaczali będziemy dalej przez $\bar{x}(t)$.

Gładką ścieżkę wzrostu kapitału na jednostkę efektywnej pracy – zgodnie z zależnościami (21)–(22) – opisuje równanie:

$$\bar{k}_E(t) = \left(\frac{\bar{s}}{\mu} + \bar{F}e^{-(1-\alpha)\mu t} \right)^{\frac{1}{1-\alpha}}, \quad (24)$$

gdzie stałą całkowania $\bar{F}$ opisuje:

$$\bar{F} = k_{E0}^{1-\alpha} - \frac{\bar{s}}{\mu}. \quad (25)$$

Natomiast z równań (22) i (23) płynie wniosek, że gładka ścieżka wzrostu produktu na jednostkę efektywnej pracy dana jest wzorem:

$$\bar{y}_E(t) = \left(\frac{\bar{s}}{\mu} + \bar{F} e^{-(1-\alpha)\mu t} \right)^{\frac{\alpha}{1-\alpha}}, \quad (26)$$

przy czym stałą $\bar{F}$ określa zależność (25).

Oznaczmy też przez $\kappa(t) = \frac{k_E(t) - \bar{k}_E(t)}{\bar{k}_E(t)}$ względne odchylenie cyklicznej ścieżki czasowej kapitału na jednostkę efektywnej pracy k_E od jej gładkiej ścieżki wzrostu $\bar{k}_E$, zaś przez $\gamma(t) = \frac{y_E(t) - \bar{y}_E(t)}{\bar{y}_E(t)}$ – względne odchylenie y_E od $\bar{y}_E$. Wówczas, po uwzględnieniu zależności (5–8), mamy:

$$\kappa(t) = \frac{k_E(t)}{\bar{k}_E(t)} - 1 = \frac{k(t)}{\bar{k}(t)} - 1 = \frac{K(t)}{\bar{K}(t)} - 1 \quad (27)$$

oraz:

$$\gamma(t) = \frac{y_E(t)}{\bar{y}_E(t)} - 1 = \frac{y(t)}{\bar{y}(t)} - 1 = \frac{Y(t)}{\bar{Y}(t)} - 1. \quad (28)$$

Z równań (27) i (28) płynie wniosek, że ścieżki κ i γ można nazywać względnymi odchyleniami (odpowiednio) kapitałowych i produkcyjnych cyklicznych ścieżek wzrostu od ich gładkich ścieżek wzrostu. Wstawiając zaś równania (21) i (24) do związku (27) oraz (23) i (26) do (28) mamy:

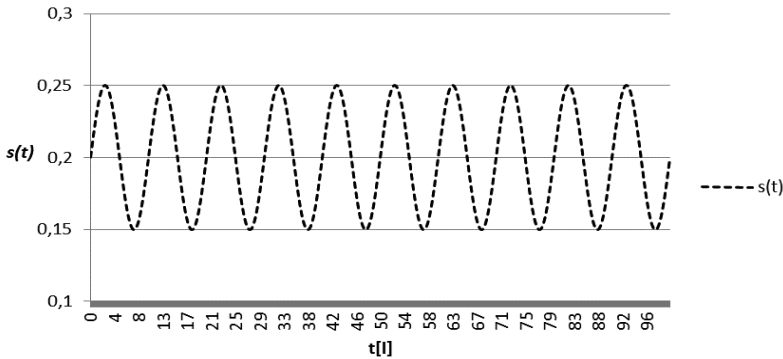
$$\kappa(t) = \left(\frac{\frac{\bar{s}}{\mu} + (1-\alpha)\theta \frac{(1-\alpha)\mu \sin\left(\frac{2\pi}{\omega}t\right) - \frac{2\pi}{\omega} \cos\left(\frac{2\pi}{\omega}t\right)}{(1-\alpha)^2\mu^2 + \frac{4\pi^2}{\omega^2}} + F e^{-(1-\alpha)\mu t}}{\frac{\bar{s}}{\mu} + \bar{F} e^{-(1-\alpha)\mu t}} \right)^{\frac{1}{1-\alpha}} - 1 \quad (29)$$

i:

$$\gamma(t) = \left(\frac{\frac{\bar{s}}{\mu} + (1-\alpha)\theta \frac{(1-\alpha)\mu \sin\left(\frac{2\pi}{\omega}t\right) - \frac{2\pi}{\omega} \cos\left(\frac{2\pi}{\omega}t\right)}{(1-\alpha)^2\mu^2 + \frac{4\pi^2}{\omega^2}} + Fe^{-(1-\alpha)\mu t}}{\frac{\bar{s}}{\mu} + \bar{F}e^{-(1-\alpha)\mu t}} \right)^{\frac{\alpha}{1-\alpha}} - 1. \quad (30)$$

4. KALIBRACJA PARAMETRÓW MODELU I SYMULACJE NUMERYCZNE

Uchylając założenie o stałości stóp inwestycji w czasie dopuszcza się po pierwsze, odchylenie od średniego poziomu stopy inwestycji (zarówno in plus jak i in minus w zależności od koniunktury gospodarczej). Po drugie zakłada się, zgodnie z teorią cykli koniunkturalnych, że owe odchylenia mają charakter cykliczny. Zatem uwzględniając powyższe założenia oraz przyjmując, że średnia stopa inwestycji wynosi np. 20% ($\bar{s} = 0,2$), odchylenia stopy inwestycji od jej średniego poziomu wynoszą ± 5 pkt. procentowych ($\theta = 0,05$) i okres wahań stóp inwestycji na poziomie 10 lat ($\omega = 10$) trajektoria stopy inwestycji będzie kształtowała się jak na rysunku 1.



Rysunek 1. Sinusoidalna funkcja inwestycji

Źródło: opracowanie własne.

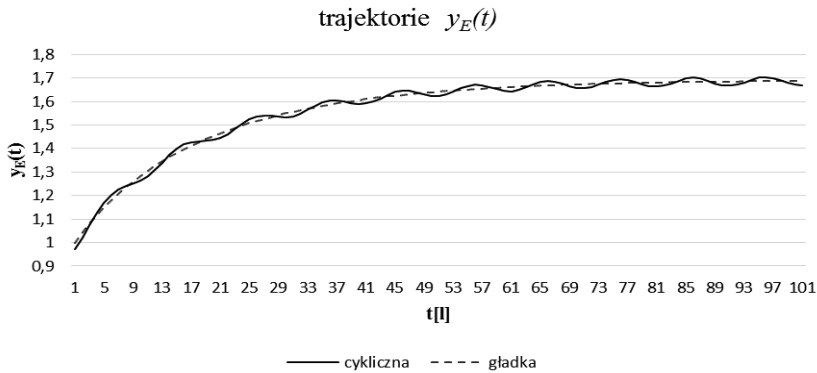
Przyjęcie powyższych założeń pozwoliło na przeprowadzenie symulacji numerycznych modelu teoretycznego (z punktu trzeciego), które prowadzone były w dwóch etapach. W pierwszym etapie dokonano kalibracji wartości parametrów otrzymanego modelu wzrostu gospodarczego, którą rozpoczęto od wyznaczenia wartości elastycz-

ności produkcji względem kapitału α , gdzie $\alpha \in (0;1)$. Bazując na danych panelowych dla krajów grupy UE15 w latach 2000–2013 oszacowano parametr α na poziomie 0,349. Wybór grupy krajów UE15 wynikał z faktu posiadania kompletnych danych statystycznych oraz z przeświadczenia, iż wśród krajów Unii Europejskiej są to gospodarki o ugruntowanych systemach rynkowych⁹. Zaś parametr α oszacowano z wykorzystanie dwóch metod: modelu SUR (ang. *seemingly unrelated regression*) oraz GLS (ang. *generalised least squares*). Przy czym w estymacjach SUR z grupy krajów UE15 wyeliminowano Luksemburg, co wynikało z ograniczeń metodologicznych modelu SUR (por. np. Madalla, 2006, s. 650). Uzyskane wartości parametrów α były na zbliżonym poziomie zatem do dalszych analiz numerycznych przyjęto $\alpha = 0,349$. Ponadto oszacowana wartość α jest zbliżona do szacunków wynikających z dekompozycji modelu Solowa $\alpha \approx 1/3$ (Solow, 1957). W drugim etapie dokonano symulacji numerycznych otrzymanego modelu. Do analiz numerycznych przyjęto trzy warianty średnich stóp inwestycji 15%, 20% oraz 25% ze stałymi odchyleniami na poziomie ± 5 pkt. procentowych ($\theta = 0,05$), parametr $\mu = 0,075$ oraz $k_{E0} = 1$ ¹⁰. Przyjęcie ww. poziomów parametrów modelu wynikało z analizy danych empirycznych w krajach UE 15 za lata 2000–2013 oraz wiedzy o długookresowych charakterystykach gospodarek. Ponadto założono dwa okresy cykli stóp inwestycji 4 oraz 10 lat. Przyjęte okresy cyklicznych zmian stóp inwestycji wynikają bezpośrednio z teorii wahań koniunkturalnych i dotyczą cykli krótkich Kitchina oraz cykli średnich Juglara. Analiz dokonano w 100 letnim horyzoncie czasowym, a ich wyniki zestawiono w tabeli 1 oraz zobrazowano na rysunkach 2–7. Ze względu na trudność uchwycenia 100 letniej perspektywy czasowej w tabeli przedstawiono jedynie wartości minimalne, maksymalne oraz średnie odchyleni cyklicznych ścieżek wzrostu kapitału kappa ($\kappa(t)$) oraz produktu gamma ($\gamma(t)$) od ich gładkich odpowiedników.

Analizując trajektorie $k_E(t)$ oraz $y_E(t)$, przedstawione na rysunkach 2–3, można zaobserwować, iż cykliczne ścieżki wynikające z prezentowanego modelu istotnie odchylają się zarówno in plus jak i in minus od gładkich trajektorii. Przy czym, odchylenia kapitału na jednostkę efektywnej pracy były wyższe niż odchylenia produktu na jednostkę efektywnej pracy. Wynika to zarówno z porównania rysunków 2 i 3 oraz analizy danych zestawionych w tabeli 1.

⁹ Rzecz jasna, tego typu symulacje można prowadzić zarówno dla poszczególnych gospodarek lub innych grup krajów (wykorzystując dane panelowe) np. dla gospodarek OECD.

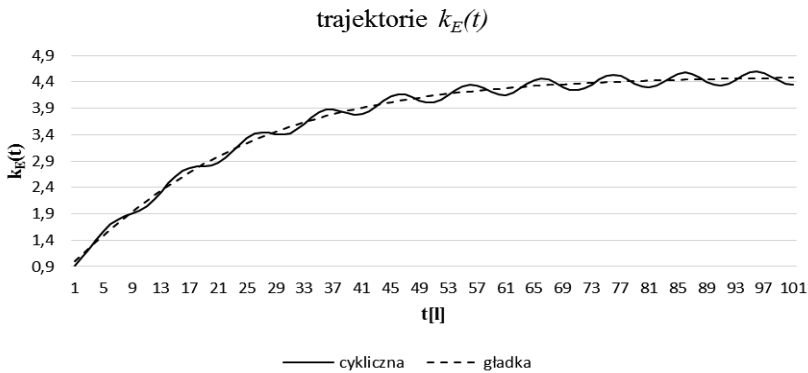
¹⁰ k_{E0} znormalizowano dla $t = 0$ do jedności, aby łatwo wyznaczyć relacje $\lim_{t \rightarrow +\infty} \frac{k_E(t)}{k_E(0)}$.



Rysunek 2. Symulacje ścieżek wzrostu produktu $y_E(t)$ przy $\omega = 10$ i $\bar{S} = 0,20$

Źródło: obliczenia własne na podstawie danych na stronie

<http://w3.unece.org/pxweb/database/STAT/20-ME/3-MELF/?lang=1>.



Rysunek 3. Symulacje ścieżek wzrostu kapitału $k_E(t)$ przy $\omega = 10$ i $\bar{S} = 0,20$

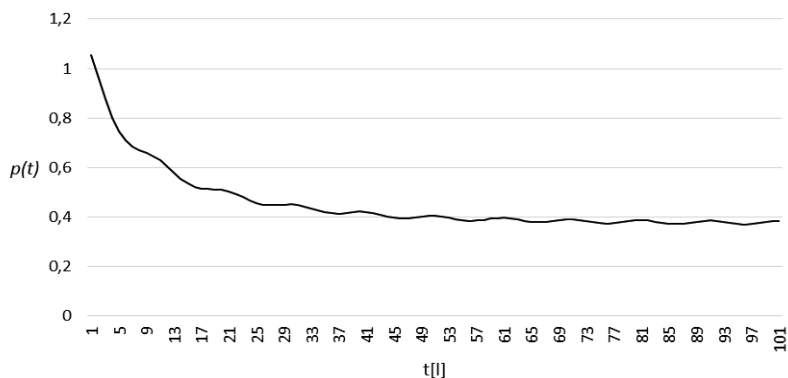
Źródło: obliczenia własne na podstawie danych na stronie

<http://w3.unece.org/pxweb/database/STAT/20-ME/3-MELF/?lang=1>.

Z przeprowadzonych symulacji wynika, iż przyrosty kapitału na jednostkę efektywnej pracy są wyższe niż przyrosty produktu na jednostkę efektywnej pracy to produktywność kapitału na jednostkę efektywnej pracy maleje w czasie. Ponieważ produktywność kapitału oznacza stosunek produktu do kapitału dany wzorem:

$$p(t) = \frac{Y(t)}{K(t)} = \frac{y_E(t)}{k_E(t)},$$

to można wyznaczyć wartość produktywności kapitału w czasie, której trajektorię przedstawiono na rysunku 4.

Rysunek 4. Czasowa ścieżka produktywności kapitału $p(t)$ przy $\omega = 10$ i $\bar{S} = 0,20$

Źródło: opracowanie własne.

W modelu Solowa z sinusoidalnymi inwestycjami produktywność kapitału na jednostkę efektywnej pracy jest funkcją malejącą z widocznymi cyklicznymi odchyleniami zarówno in plus jak i in minus. Wyznaczona ścieżka produktywności kapitału jest zgodna z założeniem modelu Solowa o malejącej produktywności krańcowej kapitału.

Kolejnym etapem symulacji było wyznaczenie względnych odchyłeń kapitału $\kappa(t)$ oraz produktu $\gamma(t)$ na jednostkę efektywnej pracy, których wyniki zestawiono w tabeli 1 oraz przedstawiono na rysunkach 5–7.

Tabela 1.

Wyniki symulacji numerycznych względnych odchyłeń $\kappa(t)$ oraz $\gamma(t)$ dla grupy krajów UE 15

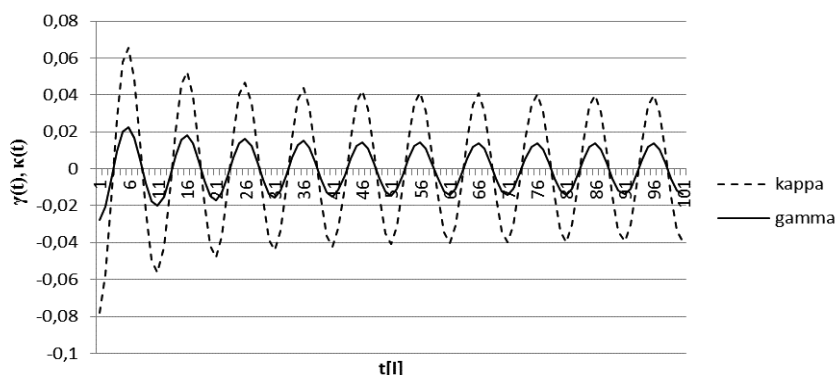
ω	$\bar{S}$	$\kappa(t)$			$\gamma(t)$		
		min	max	średnie	min	max	średnie
4	0,15	-0,0316	0,0292	-0,0002	-0,0112	0,0101	-0,0001
	0,20	-0,0316	0,0276	-0,0002	-0,0112	0,0096	-0,0001
	0,25	-0,0316	0,0262	-0,0002	-0,0112	0,0091	-0,0001
10	0,15	-0,0780	0,0657	-0,0004	-0,0279	0,0225	-0,0003
	0,20	-0,0780	0,0587	-0,0005	-0,0279	0,0201	-0,0003
	0,25	-0,0780	0,0530	-0,0006	-0,0279	0,0182	-0,0003

Źródło: obliczenia własne na podstawie danych na stronie <http://w3.unece.org/pxweb/database/STAT/20-ME/3-MELF/?lang=1>.

Z danych zawartych w tabeli 1 oraz z rysunków 5 i 7 wyciągnąć można następujące wnioski:

- Zakładając 4 letnie okresy cykliczności inwestycji zauważono, że zarówno dla kapitału jak i produktu najwyższe (najniższe) wartości odchyłeń cyklicznych ścieżek od

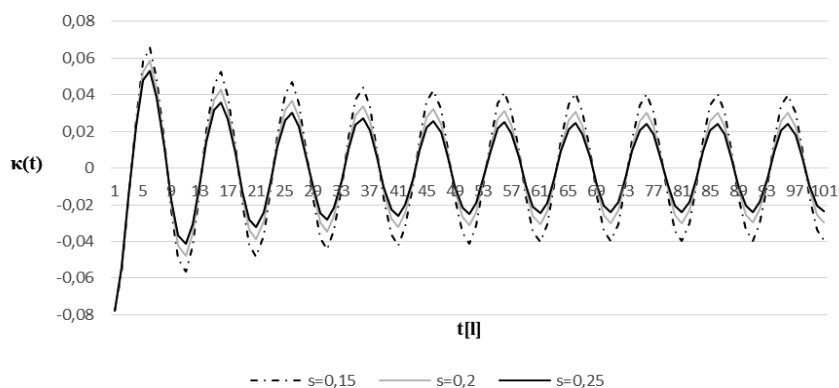
- ich gładkich odpowiedników uzyskano dla najniższej (najwyższej) średniej stopy inwestycji. Przy czym maksymalne wartości odchylen kapitału były kilkukrotnie wyższe niż odchylenia produktu. Maksymalna wartość odchylenia przy średniej stopie inwestycji na poziomie 15% wyniosła 0,03 w przypadku kapitału oraz 0,01 w przypadku produktu. Przyjmując średnią wartość stóp inwestycji na poziomie 25% okazuje się, że maksymalne odchylenia kapitału wyniosły 0,026 zaś produktu 0,009.
- W wariancie, w którym okres wahań inwestycji wynosi 10 lat okazuje się, że analogicznie jak w przypadku gdy $\omega = 4$, maksymalne wartości odchylen zarówno kapitału jak i produktu były najwyższe (najniższe) przy najniższych (najwyższych) średnich stopach inwestycji. Z tą różnicą, że w okresach 10 letnich cykli inwestycji otrzymane odchylenia były na ponad dwukrotnie wyższym poziomie niż w symulacjach 4 letnich cykli inwestycji (por. tab. 1).



Rysunek 5. Trajektorie $\kappa(t)$ oraz $\gamma(t)$ przy $\omega = 10$ i $\bar{S} = 0,15$

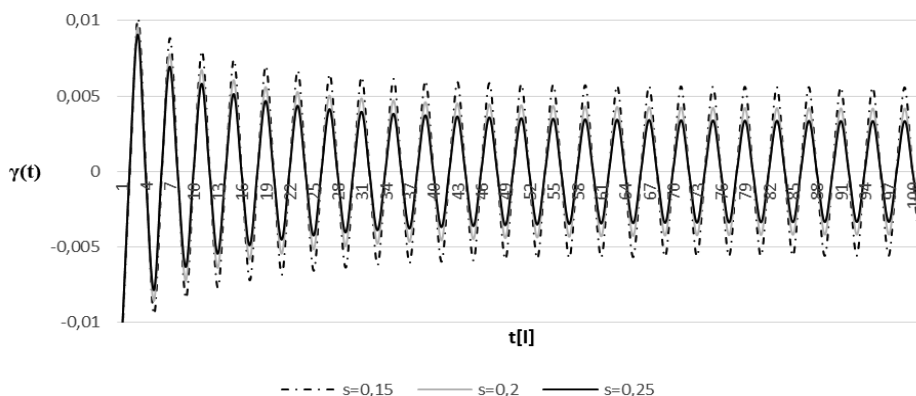
Źródło: opracowanie własne.

- Trajektorie $\kappa(t)$ oraz $\gamma(t)$ przebiegają wyraźnie cyklicznie z naprzemiennymi odchyleniami in plus oraz in minus zgodnie z założonymi okresami wahań 4 i 10 lat. Zaś średnie wartości tych odchylen zbliżone są do zera (tab. 1).
- Z symulacji numerycznych wynika ogólny wniosek, że niskie stopy inwestycji generują najwyższe odchylenia cyklicznych ścieżek od ich gładkich odpowiedników. Im wyższe średnie stopy inwestycji tym odchylenia $\kappa(t)$ oraz $\gamma(t)$ są niższe (patrz rys. 6–7).
- Najniższe odchylenia cyklicznych ścieżek kapitału $\kappa(t)$, biorąc pod uwagę analizowane w pracy warianty, zaobserwowano przy średnich stopach inwestycji na poziomie 25% oraz 4 letnich cyklach inwestycji. Najwyższe odchylenia kappa uzyskano przy 15% średnich stopach inwestycji oraz 10 letnich cyklach inwestycji.



Rysunek 6. Odchylenia cyklicznych ścieżek kapitału (κ) a stopa inwestycji przy $\omega = 10$

Źródło: opracowanie własne.



Rysunek 7. Odchylenia cyklicznych ścieżek produktu (γ) a stopa inwestycji przy $\omega = 4$

Źródło: opracowanie własne.

- Najniższe odchylenia cyklicznych ścieżek produktu $\gamma(t)$ od ich gładkich odpowiedników, w rozważanych wariantach, odnotowano przy 25% średnich stopach inwestycji oraz 4 letnich cyklach inwestycji. Najwyższe odchylenia $\gamma(t)$ cechowały gospodarki w scenariuszu, w których założono średnią stopę inwestycji na poziomie 15% oraz 10 letnie cykle inwestycji.
- Ponadto można zauważyć, iż amplitudy wahań $\kappa(t)$ oraz $\gamma(t)$ są najwyższe w początkowym okresie symulacji i cechują się tendencją malejącą do połowy okresu symulacji a następnie stabilizują się. Takie tendencje odchylen $\kappa(t)$ oraz $\gamma(t)$ mogą wynikać z trajektorii kapitału oraz produktu na jednostkę efektywnej pracy (rys. 2 i 3) które podlegają prawu malejącej produktywności krańcowej.

5. PODSUMOWANIE

Prowadzone w artykule rozważania można podsumować następująco:

1. Zaprezentowana w pracy modyfikacja modelu z sinusoidalnymi inwestycjami bazuje na neoklasycznym modelu Solowa-Swana (1956). Modyfikacja ta uchyla (zdaniem autorów) dość restrykcyjne założenie o stałości stóp inwestycji w czasie. Autorzy przyjmują założenie, że inwestycje odchylają się od swojego średniego poziomu, czyli podlegają cyklicznym wahaniom. Dlatego do opisu inwestycji wprowadzają funkcję zależną sinusoidalnie od czasu.
2. Bazując na wprowadzonych modyfikacjach wyprowadzono równania wyznaczające ścieżkę wzrostu kapitału oraz produktu na jednostkę efektywnej pracy w modelu Solowa z sinusoidalnymi stopami inwestycji. Pewnym ograniczeniem prezentowanej modyfikacji modelu wzrostu jest przyjęcie stałej amplitudy oraz okresu wahań cyklicznych. Przyjęcie stałych parametrów sinusoidy pozwoliło na domknięcie prezentowanego modelu.
3. Wyprowadzono ponadto równania na względne odchylenia cyklicznych ścieżek wzrostu kapitału oraz produktu od ich gładkich ścieżek wzrostu. Pozwoliło to na weryfikację oraz ocenę wpływu otrzymanych ścieżek wzrostu kapitału i produktu w modelu Solowa-Swana z sinusoidalnymi inwestycjami w stosunku do tych ścieżek w modelu Solowa ze stałymi w czasie stopami inwestycji.
4. Bazując na danych empirycznych dla grupy krajów UE15, dokonano kalibracji wartości parametrów modelu oraz analiz numerycznych. Wynika z nich, że wprowadzenie stóp inwestycji zależnych sinusoidalnie od czasu istotnie wpływa na cykliczność odchylenia kapitału oraz produktu na jednostkę efektywnej pracy.
5. Dla ustalonych wartości parametrów modelu przyjęcie sinusoidalnej funkcji stóp inwestycji silniej wpływa na ścieżkę wzrostu kapitału niż produktu na jednostkę efektywnej pracy, co wynika z kilkukrotnie wyższych wartości względnych odchylenia $\kappa(t)$ niż $\gamma(t)$.
6. Najniższe wartości odchylenia κ oraz γ zaobserwowano, gdy założono wysokie (w analizowanych wariantach) stopy inwestycji oraz 4 letnie okresy cykli inwestycji. Najwyższe odchylenia $\kappa(t)$ oraz $\gamma(t)$ charakterystyczne były przy niskich stopach inwestycji oraz 10-letnim okresie wahań inwestycji.
7. Z przeprowadzonych symulacji numerycznych wynika, iż na wielkość odchylenia $\kappa(t)$ oraz $\gamma(t)$ istotny wpływ ma zarówno średni poziom stóp inwestycji jak i okres wahań inwestycji.

LITERATURA

- Cobb C. W., Douglas P. H., (1928), A Theory of Production, *American Economic Review*, 18 (1), 139–165.
Dykas P., Sulima A., Tokarski T. (2008), Złote reguły akumulacji kapitału a N-kapitałowym modelu wzrostu gospodarczego, *Gospodarka Narodowa*, nr 11-12/2008, 47–75.

- Dykas P., Edigarian A., Tokarski T., (2011), Uogólnienie N-kapitałowego modelu wzrostu gospodarczego Nonnemana-Vanhoudta, w: Panek E., (red.), *Wzrost Gospodarczy Teoria. Rzeczywistość*, Zeszyty Naukowe nr 176, Poznań, 161–177.
- Madalla G. S., (2006), *Ekonometria*, Wydawnictwo Naukowe PWN, Warszawa.
- Malaga K., (2015), Jednolita teoria wzrostu gospodarczego – stan obecny a wyzwania współczesności, w: Fiedor B., (red.), *Ekonomia dla przyszłości. Stylizowane fakty a wyzwania współczesności*, PTE, Warszawa, 150–169.
- Malaga K., (2009), *O niektórych dylematach teorii wzrostu gospodarczego i ekonomii*, ZKP TE, Warszawa.
- Mankiw N. G., Romer D., Weil N. D., (1992), A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics*, 107 (2), 407–437.
- Maussner A., (1994), *Koniunkturtheorie*, Springer, Berlin.
- McCombie J. S. L., (2000), The Solow Residual, Technical Change and Aggregate Production Functions, *Journal of Post Keynesian Economics*, 23 (2), 267–297.
- McCombie J. S. L., (2001), What Does the Aggregate Production Function Tell Us? Second Thoughts on Solow's, Second Thoughts on Growth Theory, *Journal of Post Keynesian Economics*, 23 (4), 589–615.
- Mroczek K., Tokarski T., Trojak M., (2014), Grawitacyjny model zróżnicowania rozwoju ekonomicznego województw, *Gospodarka Narodowa*, nr 3, 5–34.
- Nonneman W., Vanhoudt P., (1996), A Further Augmentation of the Solow Model and the Empirics of Economics Growth for the OECD Countries, *Quarterly Journal of Economics*, 111 (3), 943–953.
- Solow R. M., (1956), A Contribution to the Theory of Economic Growth, *Quarterly Journal of Economics*, 70 (1), 65–94.
- Romer D., (2000), *Makroekonomia dla zaawansowanych*, Wydawnictwo Naukowe PWN, Warszawa.
- Tokarski T., (2005), *Statystyczna analiza regionalnego zróżnicowania wydajności pracy, zatrudnienia i bezrobocia w Polsce*, Wydawnictwo PTE, Warszawa.
- Tokarski T., (2009), *Matematyczne modele wzrostu gospodarczego (ujęcie neoklasyczne)*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków.
- Tokarski T., (2011), *Ekonomia matematyczna. Modele makroekonomiczne*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Wicksell K., (1893), *Über Wert, Kapital und Rente nach den Neueren Nationalökonomischen Theorien*, Jena.
- Żółtowska E., (1997), *Funkcja produkcji. Teoria, estymacja, zastosowania*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

NEOKLASYCZNY MODEL WZROSTU GOSPODARCZEGO Z SINUSOIDALNYMI INWESTYCJAMI

Streszczenie

Celem prezentowanego opracowania jest próba rozszerzenia neoklasycznego modelu wzrostu gospodarczego Solowa przez uwzględnienie fluktuacji inwestycji. Autorzy uchylając założenie o stałości stóp inwestycji wprowadzają funkcję inwestycji zależną sinusoidalnie od czasu. Przyjęcie sinusoidalnej funkcji inwestycji tłumaczyć można tym, iż inwestycje (podobnie jak produkcja) w dużej mierze zależne są od koniunktury gospodarczej, która podlega okresowym fluktuacjom. W rozważaniach teoretycznych autorzy wprowadzają ponadto pojęcia cyklicznej oraz gładkiej ścieżki czasowej technicznego uzbrojenia pracy oraz wydajności pracy i za pomocą tych narzędzi identyfikują, poprzez obliczenie względnych odchyłeń tych ścieżek, wpływ inwestycji zależnych sinusoidalnie od czasu.

W analizie empirycznej autorzy dokonali kalibracji parametrów otrzymanego modelu. Bazując na danych panelowych dla krajów grupy UE15 w latach 2000–2013 najpierw oszacowano parametr (elastyczność produkcji względem kapitału) na poziomie 0,349, którego wartość przyjęto do dalszych

analiz numerycznych. W drugim etapie dokonano symulacji numerycznych skalibrowanych odchyleń cyklicznej ścieżki czasowej technicznego uzbrojenia pracy (wydajności pracy) od gładkiej ścieżki wzrostu technicznego uzbrojenia pracy (wydajności pracy). Przy analizach numerycznych rozważano wpływ różnych stóp inwestycji (15%, 20%, 25%) oraz okresów wahań cyklicznych (4 lub 10 lat) na kształtowanie się owych odchyleń. Analiz numerycznych dla gospodarek grupy UE15 dokonano w 100 letnim horyzoncie czasowym.

Słowa kluczowe: cykliczne inwestycje, model wzrostu Solowa, ścieżka wzrostu produktu oraz kapitału na pracującego

THE NEOCLASSICAL GROWTH MODEL WITH SINUSOIDAL INVESTMENTS

Abstract

The aim of the present study is an attempt to extend the neoclassical model of economic growth of Solow by repealing the assumption of fixed investment and introducing an investment function dependent sinusoidally on the time. The adoption of the sinusoidal function of investment is substantiated by the fact that investments (like manufacturing) are largely depended on the economic situation, which is subject to periodic fluctuations. The authors introduce the theoretical considerations on the notion of cyclical and smooth path of temporal capital-labour ratio and labour productivity. When using these tools the authors identified, by calculating the relative deviations of the mentioned path, the impact of investment function dependent sinusoidally on the time.

In the empirical analysis the authors conducted the calibration of parameters used by the research model. Based on panel data for the EU15 between the years 2000–2013 the α parameter (production flexibility in relation to capital) was estimated first at 0,349. That value was adopted to further numerical analysis. In the second stage the simulations of numerical, calibrated deviations of temporal cyclical path of capital-labour ratio (labor productivity) from the smooth path of capital-labour ratio (labor productivity) was performed. When conducting numerical analysis the impact of different investment rates (15%, 20%, 25%) and periods of cyclical fluctuations (4 or 10 years) have been considered in relation to the formation of these deviations. Numerical analysis for the economies of the EU15 group was made for one hundred time series.

Keywords: cyclical investments, Solow growth model, the path of product growth and the capital-labour ratio

EWA SZLACHTOWSKA¹, DANIEL KOSIOROWSKI², DOMINIK MIELCZAREK³OCENA JAKOŚCI APLIKACYJNEJ ODPORNEGO ALGORYTMU
ANALIZY SKUPIEŃ TCLUST NA PRZYKŁADZIE ZBIORU DANYCH
DOTYCZĄCYCH JAKOŚCI POWIETRZA W KRAKOWIE^{4 5}

1. WPROWADZENIE

Metody wielowymiarowej analizy skupisk należą do procedur statystycznych najczęściej wykorzystywanych w praktyce gospodarczej. Mamy tutaj na uwadze m.in. eksploracyjną analizę danych – poszukiwanie modelu generującego dane ekonomiczne, wyodrębnianie typów klientów centrum handlowego bądź ostatnio, w kontekście badań prowadzonych w oparciu o tzw. wielkie zbiory danych – redukcję wymiaru zagadnienia statystycznego za pomocą tzw. mikroskupisk i danych funkcjonalnych (por. np. Jajuga, 1993; Krzyśko i inni, 2008; Walesiak, Gatnar, 2009; Kosiorowski i inni, 2015). Analiza skupień pozwala ogarnąć wielkie ilości danych będących w dyspozycji ekonomisty i spojrzeć na dane z właściwej perspektywy. Warto zwrócić uwagę, że jakość grupowania z wykorzystaniem tzw. klasycznych metod analizy skupień w krytycznym stopniu zależy od spełnienia założeń leżących u podłoża metod statystycznych stosowanych w ich obrębie (np. gdy zakładamy, że dane generuje mieszanina rozkładów normalnych). W praktyce badań ekonomicznych bardzo często mamy do czynienia z odstępstwem od przyjmowanych założeń.

Podejście odporne w modelowaniu statystycznym i analizie danych ma na celu zaproponowanie procedur statystycznych dających wiarygodne oszacowania, stanowiących użyteczne testy nie tylko w sytuacji, gdy dane generowane są przez zakładany

¹ Akademia Górniczo-Hutnicza im. S. Staszica w Krakowie, Wydział Matematyki Stosowanej, Katedra Analizy Matematycznej, Matematyki Obliczeniowej i Metod Probabilistycznych, al. Mickiewicza 30, 30-059 Kraków, Polska, autor prowadzący korespondencję – e-mail: szlachto@agh.edu.pl.

² Uniwersytet Ekonomiczny w Krakowie, Wydział Zarządzania, Katedra Statystyki, ul. Rakowicka 27, 31-510 Kraków, Polska.

³ Akademia Górniczo-Hutnicza im. S. Staszica w Krakowie, Wydział Matematyki Stosowanej, Katedra Analizy Matematycznej, Matematyki Obliczeniowej i Metod Probabilistycznych, al. Mickiewicza 30, 30-059 Kraków, Polska.

⁴ Autorzy uprzejmie dziękują za wsparcie finansowe ze strony UEK w Krakowie oraz AGH w Krakowie w postaci środków na utrzymanie potencjału badawczego 2015 i 2016.

⁵ Niniejszy artykuł bezpośrednio nawiązuje do przewodu doktorskiego Ewy Szlachtowskiej pt. *Odporna analiza skupisk w badaniach nowej gospodarki*, otwartego na Uniwersytecie Ekonomicznym w Krakowie, Wydział Zarządzania, Katedra Statystyki.

przez procedurę statystyczną rozkład, ale także w sytuacji, gdy rozkład generujący dane nieco odbiega od zakładanego rozkładu (znajduje się w pewnym sąsiedztwie zakładanego modelu). Proponowana procedura statystyczna powinna posiadać dobre własności zarówno, gdy w próbie nie ma elementów odbiegających od zasadniczej części chmury danych (tzw. obserwacji odstających, ang. *outliers*), ale także w sytuacji, gdy takie elementy występują (por. Maronna i inni, 2006). W ostatnich latach w literaturze proponuje się także odporne metody analizy skupień (por. np. Rocke, Wodurf, 2002; Kosiorowski, 2008). W związku z szeregiem trudności koncepcyjnych związanych z samym rozumieniem odporności metody analizy skupień – tematyka obfituje w szereg nierozwiązanych jak dotąd otwartych problemów (np. jak rozumieć, że procedura łamie się, w oparciu o którą miarę jakości grupowania danych definiować miarę wpływu jednostek odstających na wynik grupowania, czy odporność procedury wiązać z liczbą bądź kształtem skupień powstających w wyniku jej działania itd.)

W niniejszym artykule skupiono się na niehierarchicznym odpornym algorytmie grupowania o nazwie TCLUST, który jest jednym z najlepszych zaproponowanych jak dotąd w literaturze. Jego odporność rozumiemy w ramach jednolitego podejścia zaproponowanego w Genton, Lucas (2003).

2. ALGORYTM TCLUST

Odporny algorytm TCLUST mający swoją darmową implementację w postaci pakietu środowiska R (R Development Core Team 2010) o tej samej nazwie dobrze radzi sobie z niedoskonałościami danych ekonomicznych. Twórcami algorytmu TCLUST są H. Fritz, L. A. García-Escudero, A. Mayo-Iscar (por. Fritz i inni, 2011; Fritz i inni, 2012). Pakiet `tlust` dostępny jest na stronie <http://CRAN.R-project.org/package=tlust>.

Założenia modelu są następujące. Rozważamy próbę losową $\{x_1, \dots, x_n\} \subset \mathbb{R}^p$, gdzie x_i ($i = 1, \dots, n$) są zmiennymi losowymi o p -wymiarowym rozkładzie normalnym z funkcją gęstości $\phi(\cdot, \mu, \Sigma)$, wektorem wartości oczekiwanych $\mu = [\mu_1, \dots, \mu_p]^T$, i macierzą kowariancji Σ . Szukamy podziału $\{R_0, R_1, \dots, R_k\}$ zbioru indeksów $\{1, \dots, n\}$, gdzie $\#R_0 = \lceil n\alpha \rceil$, środków (centroidów) $m_1, \dots, m_k$, symetrycznych dodatnich pół-określonych macierzy rozproszenia $S_1, \dots, S_k$ oraz wag $p_1, \dots, p_k$, gdzie $p_j \in [0, 1]$

i $\sum_{j=1}^k p_j = 1$ maksymalizujących funkcję celu:

$$\sum_{j=1}^k \sum_{i \in R_j} \log(p_j \phi(x_i, m_j, S_j)). \quad (1)$$

Rozważamy „niejednorodny” problem grupowania obserwacji, tzn. dopuszczamy skupienia o eliptycznym kształcie oraz dopuszczamy istnienie małej α części obserwacji, które mogłyby wpłynąć negatywnie na jakość podziału danych na skupienia.

Algorytm TCLUST łączy w sobie możliwości metody grupowania k-średnich z możliwością odpornego oszacowania macierzy kowariancji, jaką daje algorytm wyznaczania estymatora minimalnego wyznacznika macierzy kowariancji (ang. *fast-MCD algorithm*, por. Rousseeuw, Van Driessen, 1999). Nowe środki oraz nowe macierze rozproszenia są wyznaczane poprzez wyliczenie empirycznego wektora średnich i empirycznej macierzy kowariancji obserwacji przypisanych do każdego skupienia. Niestety, takie połączenie obu algorytmów nie zapewnia rozsądnych wyników podziału danych. Spowodowane jest to tym, że duże skupienia mają tendencję do „pochłaniania” najmniejszych. Aby wyeliminować działanie niepożądane algorytmu analizy skupień, nakłada się pewne ograniczenia kontrolujące rozproszenie w skupieniu. Zauważmy również, że problem bezpośredniej maksymalizacji funkcji danej wzorem (1), bez nałożenia ograniczeń na macierze rozproszenia, nie jest dobrze określony. Przykładowo, gdy dla macierzy rozproszenia S_j zachodzi $\det(S_j) \rightarrow 0$, wówczas funkcja celu jest nieograniczona. W związku z tym, aby problem maksymalizacji funkcji celu danej wzorem (1) był dobrze określony, w pracy García-Escudero i inni (2008) został nałożony warunek na iloraz wartości własnych macierzy rozproszenia:

$$\frac{\max_{j,l} \lambda_{j,l}}{\min_{j,l} \lambda_{j,l}} \leq c, \quad (2)$$

gdzie $\lambda_{j,l}$ są wartościami własnymi ($l = 1, \dots, p$) macierzy rozproszenia S_j ($j = 1, \dots, k$), a $c \geq 1$ jest stałą, która kontroluje siłę ograniczenia (2). Zauważmy, że im większa zostanie wybrana stała c , tym „luźniejsze” jest ograniczenie macierzy rozproszenia, co pozwala na większą różnorodność wśród skupień. Przeciwnie, dzięki małej wartości stałej c (blisko jeden) otrzymamy bardziej jednorodnie „rozproszone” skupienia. Dla $c = 1$, otrzymujemy ważoną metodę przyciętych k-średnich oraz skupienia o kulistym kształcie.

Algorytm TCLUST, maksymalizujący funkcję celu (1) przy warunku (2), opiera się głównie na stosowaniu pewnych kroków szacowania (E-krok, ang. *estimation*) i maksymalizacji (M-krok, ang. *maximization*). Algorytm można podzielić na dwa etapy: inicjalizacji oraz zagęszczania (ang. *concentration step*). Na etapie inicjalizacji losowo wybieramy k początkowych środków m_j^0 , k początkowych macierzy rozproszenia S_j^0 oraz k początkowych wag p_j^0 . W rezultacie otrzymujemy wektor $\theta^0 = (p_1^0, \dots, p_k^0, m_1^0, \dots, m_k^0, S_1^0, \dots, S_k^0)$. Jeśli macierze S_j^0 nie spełniają ograniczenia na iloraz wartości własnych, możemy po prostu wziąć wszystkie S_j^0 równe macierzy jednostkowej i/lub zmodyfikować otrzymane macierze rozproszenia tak, aby spełniały ograniczenia nałożone na iloraz wartości własnych.

Etap zagęszczania składa się z przypisywania obserwacji do skupień, a następnie aktualizacji parametrów. Następujące kroki wykonujemy naprzemiennie, aż nie nastąpią żadne zmiany (czyli $\theta^l = \theta^{l+1}$) lub jeżeli zostanie wykonana z góry zadana liczba kroków. W E-kroku, w danej iteracji przypisanie do skupień odbywa się poprzez pomiar odległości obserwacji x_i od każdego środka m_j . Ponieważ dopuszczamy różne

wagi i macierze rozproszenia, funkcje mierzące odległość od środków definiujemy jako:

$$D_j(x_i, \theta^l) = p_j^l \phi(x_i, m_j^l, S_j^l), \quad (3)$$

gdzie $\theta^l = (p_1^l, \dots, p_k^l, m_1^l, \dots, m_k^l, S_1^l, \dots, S_k^l)$ to zbiór parametrów w bieżącej iteracji algorytmu, tj. zbiór środków $m_1^l, \dots, m_k^l$, symetrycznych dodatnich półokreślonych macierzy rozproszenia $S_1^l, \dots, S_k^l$ oraz wag $p_1^l, \dots, p_k^l$, gdzie $p_k^l \in [0, 1]$ i $\sum p_j^l = 1$. Funkcje D_j nazywamy funkcjami dyskryminującymi (ang. *discriminant functions*). Najmniejsze $D_j(x_i, \theta)$ oznacza, że obserwacja x_i jest najbardziej odległa od środka m_j . Ponadto będziemy rozważać miarę „odstawania obserwacji” (ang. *outlyingness measure*) jako $D(x_i, \theta) = \max \{D_1(x_i, \theta^l), \dots, D_k(x_i, \theta^l)\}$. Na podstawie wartości miary odstawania dla poszczególnych obserwacji, wybieramy zbiór H składający się z indeksów $1-\alpha$ części obserwacji o największych wartościach $D(x_i, \theta^l)$. Następnie dzielimy H na podzbiory $H_1, \dots, H_k$, gdzie $H_j = \{i \in H : D_j(x_i, \theta^l) = D(x_i, \theta^l)\}$. Może się zdarzyć, że otrzymamy puste skupienie albo obserwację, która może należeć do więcej niż jednego skupienia. Wtedy przyjmujemy zasadę, że taką obserwację przypisujemy do skupienia o najniższym indeksie. Warto podkreślić, że dla $k = 1$ odległość dana wzorem (3) jest odległością Mahalanobisa wykorzystywaną w algorytmie fast-MCD. Natomiast dla $p_1 = \dots = p_k$ oraz $S_1 = \dots = S_k = \sigma^2 I$ jest to odległość stosowana w metodzie k-średnich.

W M-kroku aktualizujemy parametry korzystając z informacji przypisania obserwacji do skupień. W tym kroku bardzo ważne jest nałożenie ograniczenia na iloraz wartości własnych macierzy rozproszenia i jego kontrola. Sposób nakładania ograniczeń jest jednym z najważniejszych etapów w algorytmie. Jednakże jest to „wąskie gardło” obliczeniowe algorytmu, ponieważ złożony problem optymalizacji musi być rozwiązywany w każdej iteracji zagęszczania. Aktualizacja parametrów przebiega w następujący sposób. Wagi są aktualizowane przez $p_j^{l+1} = n_j / [n(1 - \alpha)]$, gdzie $n_j = \#H_j$. Natomiast środki dane są jako empiryczne wektory średnich

$$m_j^{l+1} = \frac{1}{n_j} \sum_{i \in H_j} x_i. \quad (4)$$

Aktualizacja estymatorów rozproszenia nie jest taka prosta, ponieważ przy obliczaniu estymatorów macierzy kowariancji

$$T_j = \frac{1}{n_j} \sum_{i \in H_j} (x_i - m_j^{l+1})(x_i - m_j^{l+1})^T \quad (5)$$

może się zdarzyć, że macierze $T_1, \dots, T_k$ nie spełniają ograniczenia (2) nałożonego na iloraz wartości własnych. W takim przypadku, rozważamy rozkład macierzy $T_j = U_j^T D_j U_j$, gdzie U_j jest ortogonalną macierzą, a $D_j = \text{diag}(d_{j1}, d_{j2}, \dots, d_{jp})$ macierzą diagonalną. Zdefiniujmy odcięte wartości własne jako

$$d_{jl}^m = \begin{cases} d_{jl} & \text{dla } d_{jl} \in [m, cm] \\ m & \text{dla } d_{jl} < m \\ cm & \text{dla } d_{jl} > cm \end{cases}, \quad (6)$$

gdzie $m > 0$. Macierze rozproszenia aktualizujemy w następujący sposób: $S_j^{l+1} = U_j^T D_j^* U_j$, gdzie $D_j^* = \text{diag}(d_{j1}^{m_{opt}}, d_{j2}^{m_{opt}}, \dots, d_{jp}^{m_{opt}})$, a m_{opt} jest wartością m ($m > 0$) minimalizującą wyrażenie

$$m \mapsto \sum_{j=1}^k n_j \sum_{l=1}^p \left(\log(d_{jl}) + \frac{d_{ij}^m}{d_{ij}} \right). \quad (7)$$

Zauważmy, że w każdym kroku zagęszczania następuje wzrost wartości funkcji celu. Po przeprowadzeniu wszystkich kroków algorytmu wartość funkcji celu (1) jest wyliczana. Natomiast wynikiem końcowym algorytmu jest zbiór parametrów, które prowadzą do jak największej wartości funkcji celu.

3. WYBÓR LICZBY PARAMETRÓW POCZĄTKOWYCH

Prawdopodobnie jednym z najbardziej złożonych problemów przy stosowaniu analizy skupień jest wybór liczby skupień. Wybór liczby skupień k oraz wybór poziomu przycinania α są to powiązane problemy, które powinny być rozwiązywane jednocześnie. Ważne jest, aby zauważyć, że dany poziom przycinania pociąga za sobą określoną liczbę skupień i vice versa. Ta zależność związana jest między innymi z tym, że całe skupienia mogą zostać całkowicie przycięte przy zbyt dużym zwiększeniu α . Z drugiej strony, gdy wybrany poziom α jest zbyt niski, to grupy obserwacji odstających mogą tworzyć nowe fałszywe skupienia. W rezultacie wydaje się, że liczba skupień występująca w zbiorze danych jest wyższa. Ponadto równoczesny wybór k i α zależy od rodzaju skupień, których szukamy jak i dopuszczalnych różnic pomiędzy rozmiarami skupień. W pracy García-Escudero i inni (2011) zostało zaproponowane pewne narzędzie graficzne pomagające dokonać właściwego wyboru liczby skupień k oraz poziomu przycinania α .

Załóżmy najpierw, że stała c została wcześniej ustalona przez badacza, który stosuje metody odpornej analizy skupień. Tradycyjną metodą wyboru liczby skupień, gdy $\alpha = 0$, jest uważne kontrolowanie wartości maksymalnej funkcji celu. Jednakże zwiększanie liczby skupień k zawsze spowoduje wzrost maksymalnej wartości funkcji (1), co może prowadzić do „przeszacowania” liczby skupień. Do podejmowania rozsądnych wyborów parametrów α i k , w pracy Fritz i inni (2012) proponuje się monitorowanie funkcji największej wiarygodności $(\alpha, k) \rightarrow \mathcal{L}(\alpha, k)$, gdzie $\mathcal{L}(\alpha, k)$ jest maksymalną wartością osiągniętą przez funkcję (1) dla każdej kombinacji z danego

zbioru wartości dla k i α . W praktyce proponuje się, aby wybrać liczbę skupień jako najmniejszą wartość k taką, że

$$\mathcal{L}(\alpha, k+1) - \mathcal{L}(\alpha, k) \quad (8)$$

jest (prawie) 0 z wyjątkiem niewielkich wartości α . Gdy liczba skupień jest ustalona, jako poziom przycinania wybieramy pierwsze α_0 takie, że (8) jest bliskie 0 dla każdego $\alpha \geq \alpha_0$.

W pakiecie `tclust` jest dostępna funkcja `ctlcurves`, która przybliża funkcję największej wiarygodności poprzez wykonanie funkcji `tclust` dla sekwencji wartości k i α . Opisana powyżej procedura, dokonywania rozsądnych wyborów parametrów k i α , wymaga aktywnego udziału badacza. Wielkość ograniczenia na iloraz wartości własnych musi być z góry określona. W konsekwencji decyzja badacza, dotycząca wielkości nałożonych ograniczeń, determinuje właściwe ustalenie parametrów k i α . Na przykład, niektóre specyficzne zastosowania analizy skupień wymagają niemal kulistych skupień, które mogą być uzyskane przez ustalenie stałej c blisko 1. Domyślnie jest ustawiona wartość $c = 50$, jednakże można w razie potrzeby zmienić wartość stałej c .

Otrzymane w ten sposób „rozsądne” wartości k i α oraz związane z nimi skupienia należy dokładnie zbadać. Ponadto algorytm `TCLUST` wysyła ostrzeżenie, gdy uzyskane skupienia zostały „sztucznie ograniczone” algorytmem. Oznacza to, że wartości własne macierzy rozproszenia spełniają warunek

$$\frac{\max_{j,l} \lambda_{j,l}}{\min_{j,l} \lambda_{j,l}} = c, \quad (9)$$

ponieważ algorytm wymusił takie ograniczenie. W tej sytuacji, jeśli nie są wymagane żadne szczególne ograniczenia, stała c może być stopniowo zwiększana do czasu, aż ostrzeżenie zniknie.

Ponadto w pakiecie `tclust` dostępne jest dodatkowe narzędzie graficzne, które może być stosowane w celu oceny jakości podziału danych i podjętych decyzji przycinania. Oceny tej dokonuje się przez zastosowanie funkcji dyskryminujących.

Niech $R = \{R_0, R_1, \dots, R_k\}$, $\theta = (p_1, \dots, p_k, m_1, \dots, m_k, \dots, S_1, \dots, S_k)$ będą wartościami otrzymanymi z algorytmu poprzez maksymalizację funkcji (1). $D_j(x_i, \theta)$ jest miarą stopnia przynależności obserwacji x_i do j -tego skupienia. Wartości funkcji dyskryminujących, mierzących odległości obserwacji x_i od poszczególnych środków skupień, można posortować rosnąco

$$D_{(1)}(x_i, \theta) \leq \dots \leq D_{(k)}(x_i, \theta), \quad (10)$$

gdzie $D_{(k)}(x_i, \theta) = D(x_i, \theta)$. Jakość decyzji przyporządkowania nieprzyciętych obserwacji x_i można ocenić poprzez porównanie jej stopnia przynależności do najbliższego skupienia z drugim z możliwych najlepszych dopasowań, tj.

$$DF(i) = \log \left(\frac{D_{(k-1)}(x_i, \theta)}{D_{(k)}(x_i, \theta)} \right). \quad (11)$$

Ponadto jeżeli $x_{(1)}, \dots, x_{(n)}$ będą obserwacjami w próbie posortowanymi względem ich wartości $D_{(k)}(\cdot, \theta)$ ($D_{(k)}(x_{(1)}, \theta) \leq \dots \leq D_{(k)}(x_{(n)}, \theta)$) to $x_{(1)}, \dots, x_{(\lceil na \rceil)}$ są przycięte obserwacje, które nie są przypisane do żadnego skupienia. Niemniej jednak, możliwe jest obliczenie stopnia przynależności przyciętej obserwacji x_i do jej najbliższego skupienia. Tym samym jakość decyzji przycinania owej obserwacji można ocenić przez porównanie $D_{(k)}(x_i, \theta)$ z $D_{(k)}(x_{\lceil na \rceil+1}, \theta)$, tj.

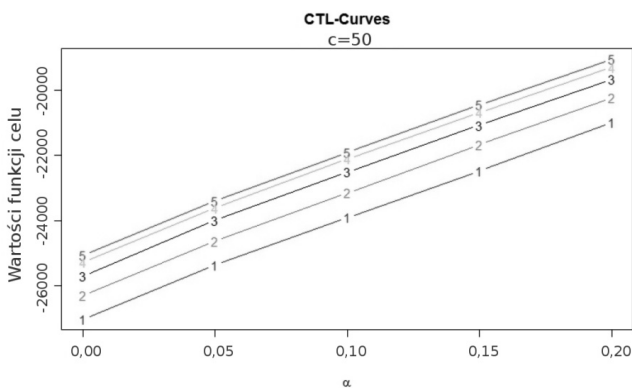
$$DF(i) = \log \left(\frac{D_{(k)}(x_i, \theta)}{D_{(k)}(x_{\lceil na \rceil+1}, \theta)} \right). \quad (12)$$

W rezultacie czynniki dyskryminujące $DF(i) \leq 0$ są uzyskiwane dla każdej obserwacji w zbiorze danych, przyciętej czy też nie. Do oceny uzyskanych czynników dyskryminujących służy funkcja DiscrFact. Wynikiem funkcji DiscrFact są trzy wykresy. Wykres „Klasyfikacja” (ang. *classification*) ilustruje przypisanie obserwacji do skupień oraz podjęte decyzje przycinania. Wykres zarysu (ang. *silhouette plot*, por. Rousseeuw, 1987) przedstawia wartości czynników dyskryminujących dla poszczególnych skupień. Czynniki dyskryminujące określają dobroć podjętej decyzji grupowania. Obiekty na wykresie zarysu z wieloma dużymi wartościami $DF(i)$ (tj. blisko 0) wskazują na istnienie niezbyt „dobrze dobranych” skupień. Najbardziej „wątpliwe” przypisania o czynniku $DF(i)$ większym niż ustalona wartość progowa ($\log(prog)$) są wyświetlane przez funkcję DiscrFact. Przykładowo, wybór $prog = 0,1$ oznacza, że decyzja dla danej obserwacji x_i jest uważana za wątpliwą, jeżeli wielkość czynnika dyskryminującego drugiej najlepszej z możliwych decyzji ($D_{(k-1)}(x_i, \theta)$ lub $D_{(k)}(x_{\lceil na \rceil+1}, \theta)$) jest większa niż jedna dziesiąta wielkości czynnika dyskryminującego rzeczywiście podjętej decyzji ($D_{(k)}(x_i, \theta)$). Natomiast im mniejsze wartości czynników dyskryminujących tym lepiej dopasowane skupienia. Wątpliwe decyzje są zaznaczone na wykresie „Obserwacje o wątpliwej przynależności” (ang. *doubtful assignments*).

4. ZASTOSOWANIE ALGORYTMU TCLUST DLA ZBIORU DANYCH DOTYCZĄCYCH JAKOŚCI POWIETRZA

Aby zobrazować wady i zalety algorytmu TCLUST wybrano dane dotyczące jakości powietrza ze stacji Aleja Krasińskiego w Krakowie za okres od 1 do 31 października 2015 r. Wyniki pomiarów siedmiu substancji są prezentowane co godzinę na stronie <http://monitoring.krakow.pios.gov.pl/>.

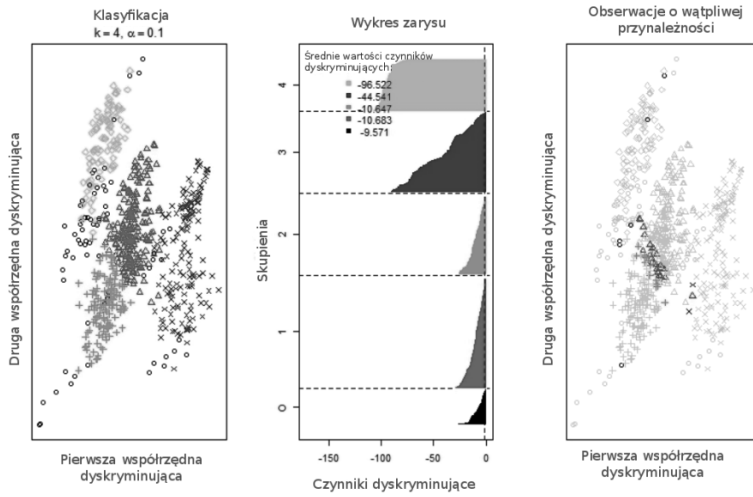
W związku z próbą pogrupowania otrzymanych danych ze względu na wielkość poziomu niektórych substancji w powietrzu w zależności od pory dnia zastosowano algorytm TCLUST. W celu ustalenia optymalnej liczby skupień zastosowano funkcję *ctlcurves* dostępną w pakiecie *tlust* dla sekwencji parametrów: k od 1 do 5 oraz α od 0 do 0,2.



Rysunek 1. Klasyfikacja funkcji największej wiarygodności
dla parametrów $k = 1, \dots, 5$, $\alpha = 0, 0,05, \dots, 0,2$

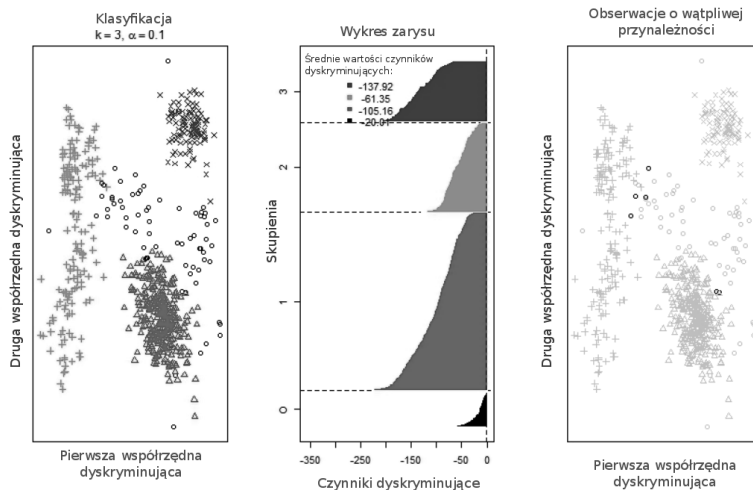
Źródło: opracowanie własne za pomocą programu R.

Na rysunku 1 widać wyraźną różnicę pomiędzy wartościami funkcji celu obliczonymi dla jednego skupienia oraz dwóch skupień. Również widoczny wzrost jest wartości funkcji celu przechodząc od dwóch skupień do trzech, jak również od trzech do czterech. Jednakże trudno określić czy rozważać cztery skupienia, czy konieczne jest rozważanie pięciu skupień. Podobnie, różnica w wartościach funkcji celu dla różnych poziomów przycinania, przechodząc od trzech do czterech skupień wydaje się praktycznie stała dla uwzględnionych wartości przycinania.

Rysunek 2. Wykres zarysu dla parametrów $k = 4$, $\alpha = 0,1$, $c = 50$

Źródło: opracowanie własne za pomocą programu R.

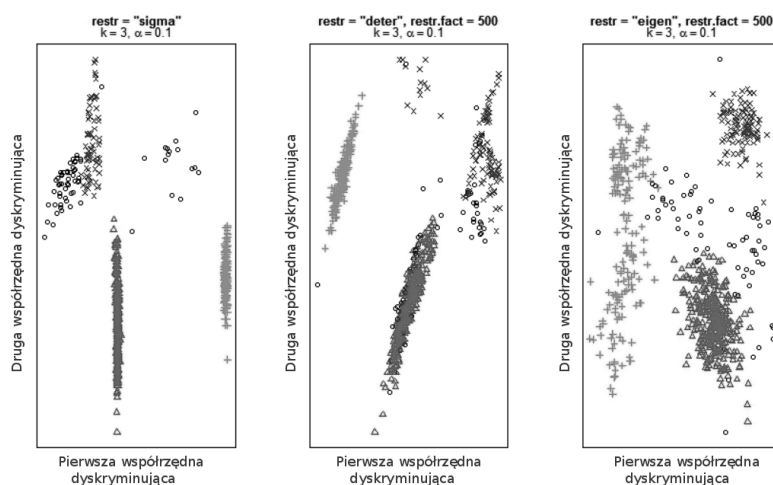
Rysunek 2 przedstawia wykres zarysu. Na wykresie „Klasyfikacja” można zaobserwować, że wątpliwe obserwacje występują głównie na styku dwóch skupień. Wykres zarysu przedstawia wartości czynników dyskryminujących dla poszczególnych skupień. Dla rozważanych parametrów $k = 4$, $\alpha = 0,1$ oraz $c = 50$ wartości czynników dyskryminujących tylko w dwóch skupieniach są oddalone od zera, co świadczy o nie-najlepszym wyborze parametrów. Dodatkowo otrzymaliśmy informację, iż algorytm wymusił ograniczenie na iloraz wartości własnych macierzy rozproszenia.

Rysunek 3. Wykres zarysu dla parametrów $k = 3$, $\alpha = 0,1$, $c = 500$

Źródło: opracowanie własne za pomocą programu R.

Na rysunku 3 został przedstawiony wykres zarysu dla parametrów $k = 3$, $\alpha = 0,1$, $c = 500$. Na wykresie „Obserwacje o wątpliwej przynależności” niewiele decyzji zostało zaznaczonych jako decyzje wątpliwe. Natomiast na wykresie zarysu wartości czynników dyskryminujących są oddalone od zera, co świadczy o dobrym doborze parametrów. Warto podkreślić, że mimo sugestii, iż dla zestawu parametrów $k = 4$ i $\alpha = 0,1$ wartość funkcji celu będzie znacząco wyższa, to jednak zestaw parametrów $k = 3$, $\alpha = 0,1$ wskazuje na lepsze decyzje grupowania.

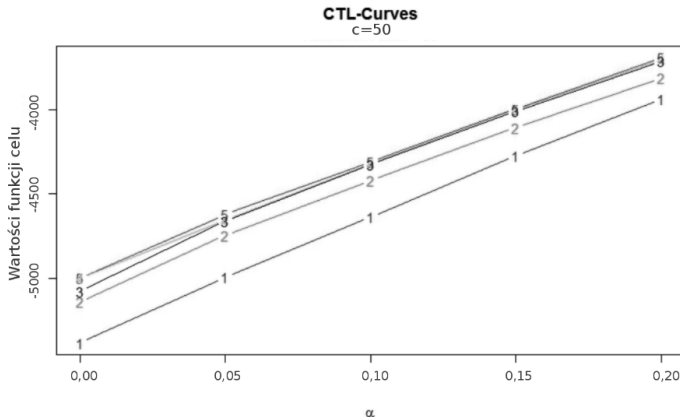
W pakiecie tclust istnieje również możliwość nałożenia innych ograniczeń na macierze rozproszenia. Oprócz ograniczenia na iloraz wartości własnych, które to decyduje o kształcie skupień, można zażądać, aby wszystkie macierze rozproszenia były takie same albo nałożyć ograniczenie na wyznaczniki macierzy rozproszenia. Tego typu ograniczenia wpływają na objętość skupień, a nie na ich kształt.



Rysunek 4. Wyniki grupowania przy różnych ograniczeniach na macierze rozproszenia dla parametrów $k = 3$, $\alpha = 0,1$, $c = 500$

Źródło: opracowanie własne za pomocą programu R.

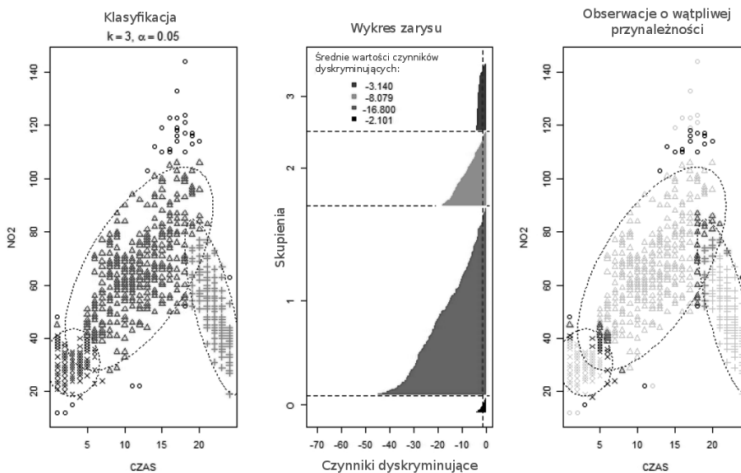
Po wstępnej analizie zbioru danych dotyczących jakości powietrza, trudno znaleźć zależność pomiędzy porą dnia a przypisaniem do skupień. Otrzymany podział wymaga dalszej i bardziej szczegółowej analizy dotyczącej jakości powietrza. W związku z tym w kolejnym etapie ograniczyliśmy się tylko do dwutlenku azotu.



Rysunek 5. Klasyfikacja funkcji największej wiarygodności dla parametrów $k = 1, \dots, 5$, $\alpha = 0, 0,05, \dots, 0,2$

Źródło: opracowanie własne za pomocą programu R.

Na rysunku 5 widać wyraźny wzrost wartości funkcji celu przechodząc od jednego skupienia do dwóch, jak również od dwóch do trzech. W związku z tym zastosowano algorytm TCLUST dla parametrów $k = 3$ i $\alpha = 0,05$.



Rysunek 6. Wykres zarysu dla parametrów $k = 3$, $\alpha = 0,05$, $c = 500$

Źródło: opracowanie własne za pomocą programu R.

Rysunek 6 przedstawia wykres zarysu. Na wykresie „Klasyfikacja” można zaobserwować, że wątpliwe obserwacje występują głównie na styku skupień. Ponadto na rysunku 6 można zaobserwować, że między godziną 6 a 19 występuje największe natężenie emisji dwutlenku azotu. Na szczególną uwagę zasługują również obserwacje odstające, które powinny zostać poddane dalszej szczegółowej analizie. W tabeli 1

podano środki poszczególnych skupień. Można zaobserwować, że koło południa występuje najwyższe stężenie dwutlenku azotu i utrzymuje się do późnych godzin wieczornych. Natomiast tuż nad ranem jest najniższe.

Tabela 1.

Środki poszczególnych skupień dla parametrów $k = 3$, $\alpha = 0,05$, $c = 500$

Środki skupień	C1	C2	C3
Czas	11,91	21,84	3,04
NO ₂	66,17	51,77	30,99

Źródło: opracowanie własne za pomocą programu R.

5. PODSUMOWANIE

Jak już wcześniej wspomniano, w analizie skupień bardzo istotna jest rola badacza. W pierwszym kroku należy wybrać odpowiedni algorytm analizy skupień. Tutaj istotne jest jaką informację o zbiorze danych posiada badacz. Pakiet tclust służy do analizy bardzo dużych zbiorów danych o różnej wymiarowości. Algorytm TCLUSST prezentuje podejście odporne w modelowaniu statystycznym. Ponadto zakłada się między innymi, że dane generowane są przez rozkład normalny lub zbiór danych jest mieszaniną rozkładów normalnych. W związku z tym badacz powinien mieć pewną informację na wejściu dotyczącą danych. Bardzo dobrze sprawdza się w sytuacji, gdy szukamy niejednorodnych skupień o eliptycznych kształtach. Dzięki możliwości nakładania pewnych ograniczeń na macierze rozproszenia, można wpływać na kształt, jak również na objętość otrzymanych skupień. Zauważmy, że $\lceil na \rceil$ obserwacji nie jest brane pod uwagę przy obliczaniu funkcji celu (1). W rezultacie algorytm dobrze się sprawdza w sytuacji, gdy rozkład generujący dane nieco odbiega od zakładanego rozkładu, jak również można uniknąć szkodliwego wpływu odstających obserwacji. Kolejnym krokiem jest dobór parametrów wejściowych: liczba skupień, poziom przycinania, wartość ograniczenia na iloraz wartości własnych. W pakiecie tclust zaproponowano kilka rozwiązań pomocnych przy wyborze parametrów wejściowych. Jest to funkcja ctlcurves, która służy do wyboru liczby skupień oraz poziomu przycinania, oraz funkcja DicsrFact, która służy do weryfikacji podjętych decyzji grupowania. Ponadto algorytm informuje badacza czy nałożone ograniczenie na iloraz wartości własnych macierzy rozproszenia zostało „wymuszone”. Ostatnim i najtrudniejszym etapem grupowania jest interpretacja wyników. Przykładowo, po wstępnej analizie zbioru danych dotyczących jakości powietrza, trudno znaleźć zależność pomiędzy porą dnia a przypisaniem do skupień. Otrzymany podział wymagał dalszej i bardziej szczegółowej analizy danego zbioru danych. Jednakże po ograniczeniu się do danych pomiarowych dotyczących zanieczyszczeń powietrza dwutlenkiem azotu udało się znaleźć zależność pomiędzy porą dnia a wielkością stężenia NO₂ w Krakowie.

LITERATURA

- Fritz H., García-Escudero L. A., Mayo-Iscar A., (2011), A Fast Algorithm for Robust Constrained Clustering, URL http://www.eio.uva.es/infor/personas/tclust_algorithm.pdf.
- Fritz H., García-Escudero L. A., Mayo-Iscar A., (2012), tclust: An R Package for a Trimming Approach to Cluster Analysis, *Journal of Statistical Software*, 47 (12), 1–26.
- García-Escudero L. A., Gordaliza A., Matrán C., Mayo-Iscar A., (2011), Exploring the Number of Groups in Robust Model-Based Clustering, *Statistics and Computing*, 21 (4), 585–599.
- García-Escudero L. A., Gordaliza A., Matrán C., Mayo-Iscar A., (2008), A General Trimming Approach to Robust Cluster Analysis, *The Annals of Statistics*, 36 (3), 1324–1345.
- Genton M. G., Lucas A., (2003), Comprehensive Definitions of Breakdown Points for Independent and Dependent Observations, *Journal of the Royal Statistical Society Series B*, 65, 81–84.
- Jajuga K., (1993), *Statystyczna analiza wielowymiarowa*, PWN, Warszawa.
- Kosiorowski D., Mielczarek D., Szlachetowska E., (2015), Clustering of Functional Objects in Energy Load Prediction Issues, w: Papież M., Śmiech S., (red.), *Proceedings from 9th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena*, Fundacja Uniwersytetu Ekonomicznego w Krakowie, 108–118.
- Kosiorowski D., Zawadzki Z., (2014), *DepthProc*: An R Package for Robust Exploration of Multidimensional Economic Phenomena, <http://arxiv.org/abs/1408.4542>.
- Kosiorowski D., (2008), Robust Classification and Clustering Based on the Projection Depth Function, w: Brito P., (red.), *COMPSTAT 2008, Proceedings in Computational Statistics*, Physica-Verlag, Heidelberg, 209–216.
- Krzyśko M., Wołyński W., Górecki T., Skorzybut M., (2008), *Systemy uczące się*, WNT.
- Maronna R. A., Martin R. D., Yohai V. J., (2006), *Robust Statistics – Theory and Methods*, John Wiley & Sons, Chichester.
- Rocke D. M., Woodruff D. L., (2002), Computational Connections Between Robust Multivariate Analysis and Clustering, w: Härdle R. B., (red.), *COMPSTAT 2002 Proceedings in Computational Statistics*, 255–260.
- Rousseeuw, P. J., Van Driessen K., (1999), A Fast Algorithm for the Minimum Covariance Determinant Estimator, *Technometrics*, 41 (3), 212–223.
- Rousseeuw P. J., (1987), Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis, *Journal of Computational and Applied Mathematics*, 20 (1), 53–65.
- Walesiak M., Gatnar E., (red.), (2009), *Statystyczna analiza danych z wykorzystaniem programu R*, PWN, Warszawa.

OCENA JAKOŚCI APLIKACYJNEJ ODPORNEGO ALGORYTMU ANALIZY SKUPIEŃ
TCLUST NA PRZYKŁADZIE ZBIORU DANYCH DOTYCZĄCYCH JAKOŚCI POWIETRZA
W KRAKOWIE

Streszczenie

Pozyskiwanie i gromadzenie danych to obecnie bardzo dynamiczne procesy. Przy ogromnych ilościach danych proces przetwarzania danych w celu uzyskania na ich podstawie użytecznych informacji i wniosków nie jest zadaniem trywialnym. W tym pomocna jest analiza skupień, a wynik grupowania pozwala ogarnąć dostępną informację i spojrzeć na nią z innej perspektywy. W żadnym razie nie jesteśmy w stanie pokazać całego spektrum zagadnień związanych analizą skupień, dlatego też ograniczymy się do omówienia algorytmu TCULST, którego twórcami są H. Fritz, L. A. García-Escudero, A. Mayo-Iscar (por. Fritz i in., 2011, 2012). W pracy zostaną przedstawione wady i zalety odpornego algorytmu

analizy skupień oraz omówione podstawowe funkcje dostępne w pakiecie tclust. Następnie zostanie dokonana ocena jakości aplikacyjnej algorytmu TCLUST na przykładzie zbioru danych dotyczących jakości powietrza w Krakowie.

Słowa kluczowe: odporna analiza skupień, algorytm tclust, badanie jakości powietrza

EVALUATION OF THE QUALITY OF ROBUST CLUSTERING ALGORITHM TCLUST ON THE EXAMPLE OF DATASET OF AIR POLLUTANTS EMISSION IN KRAKOW

Abstract

Acquisition and data collection is currently a very dynamic processes. In order to obtain from data useful information, when huge quantities of data, the processing of the data is not a trivial task. Cluster analysis is very helpful in this and the result of grouping the result of grouping allows us to comprehend the available information and look at it from a different perspective. In any case, we are not able to show the entire spectrum of issues related to data analysis. Therefore we limit our discussion to the analysis of clusters, then we describe the TCLUST algorithm. The authors of the algorithm are H. Fritz, L. A. García-Escudero, A. Mayo-Isar (see Fritz et al. 2011, 2012). In the paper we present the pros and cons robust clustering algorithm, and we discuss the available functions in the package tclust. Then on the example of dataset of air pollutants emission in Krakow we try to evaluate the quality of robust clustering algorithm.

Keywords: robust cluster analysis, tclust algorithm, air quality testing

MIROSLAW KRZYŚKO^{1,3}, AGNIESZKA MAJKA², WALDEMAR WOŁYŃSKI³

OCENA ZRÓŻNICOWANIA POZIOMU ŻYCIA MIESZKAŃCÓW WOJEWÓDZTW W LATACH 2003–2013 ZA POMOCĄ SKŁADOWYCH GŁÓWNYCH DLA WIELOWYMIAROWYCH DANYCH FUNKCJONALNYCH ORAZ ANALIZY SKUPIEŃ

1. WSTĘP

Badanie poziomu życia i jego zróżnicowania nabiera szczególnego znaczenia w kontekście analizy stopnia przemian gospodarczych, porównania rozwoju wybranych obszarów czy wskazania dysproporcji życia społeczeństwa zamieszkującego dany region. Dzięki takim ocenom można wskazać dystans dzielący poszczególne regiony, wyodrębnić grupy o zbliżonym poziomie życia, uchwycić podobieństwa i różnice występujące pomiędzy poziomem życia w poszczególnych jednostkach administracyjnych czy określić zagrożenia danego regionu.

Wstępując do Unii Europejskiej (UE) Polska włączyła się w realizację polityki spójności mającej na celu promowanie harmonijnego rozwoju całego terytorium UE poprzez działania prowadzące do zmniejszenia zróżnicowania w rozwoju jej regionów, a tym samym do wzmocnienia spójności gospodarczej, społecznej i terytorialnej Wspólnoty. Efektem tej polityki powinno być wyrównywanie dysproporcji w poziomie życia mieszkańców poszczególnych regionów UE. W ramach polityki spójności w latach 2007–2013 Polska otrzymała łącznie 67 mld euro, czyli 20% całego budżetu UE przeznaczonego na ten cel.

Biorąc pod uwagę, że wszystko co dzieje się w społeczeństwie i w gospodarce zmienia się w miarę upływu czasu, zasadne wydaje się spojrzenie na zmiany, jakie nastąpiły w zróżnicowaniu poziomu życia mieszkańców poszczególnych regionów Polski. Celem artykułu jest ocena zróżnicowania poziomu życia mieszkańców województw w latach 2003–2013.

¹ Państwowa Wyższa Szkoła Zawodowa im. Prezydenta Stanisława Wojciechowskiego, Wydział Zarządzania, ul. Nowy Świat 4, 62-800 Kalisz, Polska, autor prowadzący korespondencję, e-mail: mkrzysko@amu.edu.pl.

² Uniwersytet Rzeszowski, Wydział Ekonomii, Katedra Metod Ilościowych i Informatyki Gospodarczej, ul. Ćwiklińskiej 2, 35-601 Rzeszów, Polska.

³ Uniwersytet im. Adama Mickiewicza w Poznaniu, Wydział Matematyki i Informatyki, ul. Umultowska 87, 61-614 Poznań, Polska.

Nowością tej pracy jest fakt rozpatrywania lat 2003–2013 łącznie, a nie oddzielnie każdego roku. Można to było osiągnąć po przekształceniu danych oryginalnych w postaci szeregów czasowych dla każdej cechy oddzielnie na wektorowe funkcje ciągle określone na ustalonym przedziale czasowym zwane wielowymiarowymi danymi funkcjonalnymi (patrz Jacques, Preda, 2014 oraz Górecki i inni, 2014). Do oceny przestrzennego zróżnicowania poziomu życia zastosowano analizę składowych głównych dla wielowymiarowych danych funkcjonalnych (patrz sekcja 3) oraz dendrytową metodę analizy skupień (patrz sekcja 4). Metody te pozwoliły na wyodrębnienie grup województw o zbliżonym poziomie wartości rozpatrywanych cech dla całego rozpatrywanego okresu łącznie. Wszystkie obliczenia wykonane zostały przy użyciu programu R.

2. DOBÓR ZMIENNYCH ORAZ ICH UNITARYZACJA

Na podstawie przesłanek merytorycznych ustalono zestaw zmiennych diagnostycznych dotyczących wielu aspektów życia, takich jak: wynagrodzenia, rynek pracy, opieka zdrowotna i społeczna, komunikacja i infrastruktura gospodarcza, sytuacja mieszkaniowa, oświata i kultura, środowisko, bezpieczeństwo. Przy doborze zmiennych kierowano się koniecznością w miarę wszechstronnego opisu poziomu życia, dostępnością i kompletnością danych statystycznych. Wybrane zmienne miały charakter wskaźnikowy. Na liście zmiennych diagnostycznych znalazły się:

wynagrodzenia i rynek pracy:

- x_1 — przeciętne miesięczne wynagrodzenie brutto w zł (w podmiotach gospodarczych o liczbie pracujących powyżej 9 osób),
- x_2 — przeciętny miesięczny dochód rozporządzalny na 1 osobę,
- x_3 — stopa bezrobocia rejestrowanego w % (d),
- x_4 — liczba pracujących na 1000 mieszkańców,
- x_5 — odsetek pracujących w rolnictwie (d),
- x_6 — odsetek pracujących w usługach,
- x_7 — bezrobotni z wykształceniem wyższym w ogólnej liczbie ludności w wieku produkcyjnym (d),
- x_8 — bezrobotni pozostający bez pracy powyżej 24 miesięcy w ogólnej liczbie bezrobotnych (d),
- x_9 — ludność w wieku produkcyjnym na 1 tys. osób w wieku nieprodukcyjnym,
- x_{10} — ludność w wieku poprodukcyjnym na 1 tys. osób w wieku produkcyjnym (d);

opieka zdrowotna i społeczna:

- x_{11} — wydatki budżetowe w dziale ochrona zdrowia na 1 mieszkańca w zł,
- x_{12} — liczba lekarzy na 1 tys. mieszkańców,
- x_{13} — liczba mieszkańców przypadających na 1 aptekę ogólnodostępną (d),
- x_{14} — placówki stacjonarnej opieki społecznej na 1 tys. ludności;

infrastruktura gospodarcza:

- x_{15} — rozdzielcza sieć wodociągowa w km na 100 km²,
- x_{16} — rozdzielcza sieć kanalizacyjna w km na 100 km²,
- x_{17} — rozdzielcza sieć gazowa w km na 100 km²,
- x_{18} — gęstość dróg (drogi o twardej nawierzchni w km na 100 km²),
- x_{19} — liczba ludności przypadająca na 1 placówkę pocztową (d);

zasoby mieszkaniowe:

- x_{20} — wydatki budżetowe w dziale gospodarka mieszkaniowa na 1 mieszkańca w zł,
- x_{21} — przeciętna powierzchnia mieszkań w przeliczeniu na 1 osobę,
- x_{22} — liczba mieszkań na 1 tys. ludności,
- x_{23} — odsetek mieszkań wyposażonych w wodociąg,
- x_{24} — odsetek mieszkań wyposażonych w łazienkę,
- x_{25} — odsetek mieszkań wyposażonych w gaz sieciowy;

oświata, kultura i rekreacja:

- x_{26} — odsetek dzieci w wieku 3–6 lat objętych wychowaniem przedszkolnym,
- x_{27} — studenci szkół wyższych na 10 tys. ludności,
- x_{28} — liczba uczniów szkół podstawowych przypadających na 1 komputer z dostępem do Internetu (d),
- x_{29} — liczba uczniów szkół ponadgimnazjalnych przypadających na 1 komputer z dostępem do Internetu (d),
- x_{30} — wydatki budżetowe w dziale kultura i sport na 1 mieszkańca w zł,
- x_{31} — liczba klubów sportowych na 1 tys. ludności,
- x_{32} — księgozbiór bibliotek na 1 tys. ludności,
- x_{33} — domy i ośrodki kultury, kluby i świetlice na 10 tys. mieszkańców;

bezpieczeństwo i środowisko:

- x_{34} — przestępstwa stwierdzone w zakończonych postępowaniach przygotowawczych w przeliczeniu na 10 tys. mieszkańców (d),
- x_{35} — nakłady na środki trwale służące ochronie środowiska na 1 mieszkańca,
- x_{36} — odpady wytworzone na 1 km² (poza odpadami komunalnymi) (d),
- x_{37} — lesistość w %.

Literą (d) oznaczono destymulanty. Pozostałe cechy są stymulantami.

W celu ujednolicenia wartości rozpatrywanych cech, które są wyrażone w różnych jednostkach pomiarowych i mają różne przedziały zmienności, przeprowadzono ich unitaryzację zerowaną (por. np. Walesiak, 2014).

Niech X_{kij} będzie wartością cechy X_k zaobserwowaną w i -tym województwie oraz j -tym roku, gdzie $k = 1, \dots, 37$, $i = 1, \dots, 16$, $j = 1, \dots, 11$.

Wówczas zunitaryzowana wartość z_{kij} wartości X_{kij} ma postać:

$$z_{kij} = \begin{cases} \frac{x_{kij} - \min_i x_{kij}}{r_{kj}}, & \text{gdy cecha } X_k \text{ jest stymulantą,} \\ \frac{\max_i x_{kij} - x_{kij}}{r_{kj}}, & \text{gdy cecha } X_k \text{ jest destymulantą,} \end{cases}$$

gdzie

$$r_{kj} = \max_i x_{kij} - \min_i x_{kij}$$

jest rozstępem k -tej cechy w j -tym roku.

Równoważnie, zunitaryzowane wartości z_{kij} można zapisać w postaci:

$$z_{kij} = b_{kj}x_{kij} + a_{kj},$$

gdzie

$$b_{kj} = \frac{1}{r_{kj}}, \quad a_{kj} = -\frac{\min_i x_{kij}}{r_{kj}}, \quad (1)$$

w przypadku stymulant oraz

$$b_{kj} = \frac{1}{r_{kj}}, \quad a_{kj} = -\frac{\max_i x_{kij}}{r_{kj}}, \quad (2)$$

w przypadku destymulant.

Niech $\bar{x}_{kj}$, s_{kj}^2 oraz r_{kj} będą odpowiednio: wartością średnią, wariancją oraz rozstępem cechy X_k w j -tym roku. Wówczas zunitaryzowane wartości tych wielkości są równe:

$$\bar{z}_{kij} = \begin{cases} \frac{\bar{x}_{kij} - \min_i x_{kij}}{r_{kj}}, & \text{gdy cecha } X_k \text{ jest stymulantą,} \\ \frac{\max_i x_{kij} - \bar{x}_{kij}}{r_{kj}}, & \text{gdy cecha } X_k \text{ jest destymulantą,} \end{cases}$$

$$\tilde{s}_{kj}^2 = \frac{s_{kj}^2}{r_{kj}^2},$$

$$\tilde{r}_{kj} = 1,$$

dla $k = 1, \dots, 37, j = 1, \dots, 11$.

Jeśli s_{klj} jest kowariancją między cechami X_k oraz X_j w j -tym roku, gdzie $k \neq l$, to kowariancja danych zunitaryzowanych jest równa:

$$\bar{s}_{klj} = b_{kj}b_{lj}s_{klj},$$

gdzie b_{kj} oraz b_{lj} dane są wzorem (1) dla stymulant oraz wzorem (2) dla destymulant.

Widzimy, że po unitaryzacji zerowanej rozstęp wszystkich cech we wszystkich latach jest stały i równy 1, natomiast wariancje i kowariancje w danym roku są proporcjonalne do wariancji i kowariancji cech bez unitaryzacji zerowanej w tym roku.

3. WIELOWYMIAROWE FUNKCJONALNE SKŁADOWE GŁÓWNE

W rozdziale tym omówiona zostanie analiza składowych głównych dla wielowymiarowych danych funkcjonalnych (MFPCA) (por. Jacques, Preda, 2014; Górecki i inni, 2014). Klasycznymi pozycjami z zakresu metod statystycznych dla jednowymiarowych danych funkcjonalnych są monografie Ramsaya, Silvermana (2005) oraz Horvátha, Kokoszki (2012).

Załóżmy, że obserwujemy p -wymiarowy proces stochastyczny $X(t) = (X_1(t), X_2(t), \dots, X_p(t))'$ z ciągłym parametrem $t \in I$. Dalej założmy, że $E(X(t)) = \mathbf{0}$ i $X(t) \in L_2^p(I)$, gdzie $L_2(I)$ jest przestrzenią Hilberta funkcji całkowalnych z kwadratem na przedziale I z iloczynem skalarnym postaci:

$$\langle u(t), v(t) \rangle = \int_I u'(t)v(t)dt.$$

Ponadto założmy, że k -ta składowa procesu $X(t)$ może być reprezentowana przez skończoną liczbę ortonormalnych funkcji bazowych $\{\varphi_b\}$

$$X_k(t) = \sum_{b=0}^{B_k} c_{kb} \varphi_b(t), \quad t \in I, \quad k = 1, 2, \dots, p,$$

gdzie c_{kb} są zmiennymi losowymi takimi, że $E(c_{kb}) = 0$, $\text{Var}(c_{kb}) < \infty$, $k = 1, 2, \dots, p$, $b = 0, \dots, B_k$.

Niech

$$\mathbf{c} = (c_{10}, \dots, c_{1B_1}, \dots, c_{p0}, \dots, c_{pB_p})',$$

$$\Phi(t) = \begin{bmatrix} \varphi'_1(t) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \varphi'_2(t) & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \varphi'_p(t) \end{bmatrix}, \quad (3)$$

gdzie $\varphi_k(t) = (\varphi_0(t), \dots, \varphi_{B_k}(t))'$, $k = 1, 2, \dots, p$.

Używając notacji macierzowej proces $X(t)$ ma następującą reprezentację

$$X(t) = \Phi(t)c, \quad t \in I, \quad E(c) = 0, \quad \text{Var}(c) = \Sigma_c.$$

Poszukujemy zmiennej losowej

$$U = \langle u(t), X(t) \rangle = \int_I u'(t) X(t) dt$$

mającej maksymalną wariancję dla wszystkich $u(t) \in L_2^p(I)$ takich, że $\langle u(t), u(t) \rangle = 1$. Możemy założyć, że wektor funkcji wagowych $u(t)$ oraz proces $X(t)$ należą do tej samej przestrzeni, tzn. funkcja $u(t)$ może być przedstawiona w następującej postaci:

$$u(t) = \Phi(t)u,$$

gdzie $u \in \mathbb{R}^{K+p}$, $K = B_1 + \dots + B_p$. Wtedy

$$\langle u(t), X(t) \rangle = \langle \Phi(t)u, \Phi(t)c \rangle = u' \langle \Phi(t), \Phi(t) \rangle c = u'c$$

oraz

$$E(\langle u(t), X(t) \rangle) = u'E(c) = u'0 = 0,$$

$$\text{Var}(\langle u(t), X(t) \rangle) = u'E(cc')u = u'\Sigma_c u.$$

Niech

$$\lambda_1 = \sup_{u(t) \in L_2^p(I)} \text{Var}(\langle u(t), X(t) \rangle) = \text{Var}(\langle u_1(t), X(t) \rangle) = u_1' \Sigma_c u_1,$$

gdzie $\langle u_1(t), u_1(t) \rangle = u_1' u_1 = 1$.

Zmienną losową $U_1 = \langle u_1(t), X(t) \rangle = u_1' c$ nazywać będziemy pierwszą funkcjonalną składową główną, a funkcję wektorową $u_1(t)$ pierwszym wektorem funkcji wagowych. Następnie szukamy drugiej funkcjonalnej składowej głównej $U_2 = \langle u_2(t), X(t) \rangle = u_2' c$, maksymalizującej $\text{Var}(\langle u(t), X(t) \rangle) = u' \Sigma_c u$ takiej, że $\langle u_2(t), u_2(t) \rangle = u_2' u_2 = 1$ oraz nieskorelowanej z pierwszą funkcjonalną składową główną U_1 , tzn. spełniającą warunek $\langle u_1(t), u_2(t) \rangle = u_1' u_2 = 0$.

Ogólnie k -ta funkcjonalna składowa główna $U_k = \langle u_k(t), X(t) \rangle = u_k' c$ spełnia warunki:

$$\lambda_k = \sup_{u(t) \in L_2^p(I)} \text{Var}(\langle u(t), X(t) \rangle) = \text{Var}(\langle u_k(t), X(t) \rangle) = u_k' \Sigma_c u_k,$$

$$\langle u_{\kappa_1}(t), u_{\kappa_2}(t) \rangle = \delta_{\kappa_1 \kappa_2}; \quad \kappa_1, \kappa_2 = 1, \dots, k.$$

Parę $(\lambda_k, u_k(t))$ będziemy nazywać k -tym układem głównym procesu $X(t)$.

Rozważmy teraz zagadnienie składowych głównych dla wektora losowego $\mathbf{c}$. k -ta składowa główna $U_k^* = \langle \mathbf{u}_k, \mathbf{c} \rangle$ tego wektora spełnia warunki:

$$\gamma_k = \sup_{\mathbf{u} \in \mathbb{R}^{K+p}} \text{Var}(\langle \mathbf{u}, \mathbf{c} \rangle) = \sup_{\mathbf{u} \in \mathbb{R}^{K+p}} \mathbf{u}' \text{Var}(\mathbf{c}) \mathbf{u} = \sup_{\mathbf{u} \in \mathbb{R}^{K+p}} \mathbf{u}' \Sigma_c \mathbf{u} = \mathbf{u}_k' \Sigma_c \mathbf{u}_k,$$

$$\mathbf{u}_{k_1}' \mathbf{u}_{k_2} = \delta_{k_1 k_2},$$

gdzie $k_1, k_2 = 1, \dots, k$, $K = B_1 + \dots + B_p$. Parę $(\gamma_k, \mathbf{u}_k)$ nazywać będziemy k -tym układem głównym wektora $\mathbf{c}$.

Wyznaczenie k -tego układu głównego wektora $\mathbf{c}$ jest równoważne z wyznaczeniem wartości własnych oraz odpowiadających im wektorów własnych macierzy kowariancji Σ_c spełniających warunek $\mathbf{u}_{k_1}' \mathbf{u}_{k_2} = \delta_{k_1 k_2}$.

Z powyższych rozważań wynika następujące twierdzenie.

Twierdzenie k -ty układ główny $(\lambda_k, \mathbf{u}_k(t))$ procesu stochastycznego $\mathbf{X}(t)$ jest związany z k -tym układem głównym $(\gamma_k, \mathbf{u}_k)$ wektora losowego $\mathbf{c}$ następującymi zależnościami:

$$\lambda_k = \gamma_k, \quad \mathbf{u}_k(t) = \Phi(t) \mathbf{u}_k, \quad t \in I,$$

gdzie $k = 1, \dots, K + p$, $K = B_1 + B_2 + \dots + B_p$.

Analiza składowych głównych dla wektora losowego $\mathbf{c}$ bazuje na macierzy Σ_c . W praktyce macierz ta nie jest znana. Możemy ją oszacować na podstawie n niezależnych realizacji $\mathbf{x}_1(t)$, $\mathbf{x}_2(t)$, ..., $\mathbf{x}_n(t)$ procesu losowego $\mathbf{X}(t)$.

W typowych sytuacjach dane pochodzą z obserwacji zmiennych w dyskretnych momentach czasowych. Proces transformacji takich danych dyskretnych do danych funkcjonalnych wykonujemy osobno dla każdej zmiennej $X_1, X_2, \dots, X_p$.

Niech x_{kj} oznacza obserwowaną wartość zmiennej X_k , $k = 1, 2, \dots, p$ w j -tym momencie czasowym t_j , gdzie $j = 1, 2, \dots, J$. Zatem dane składają się z pJ par (t_j, x_{kj}) . Te dane dyskretnie wygładzamy za pomocą funkcji ciągłych $x_k(t)$, gdzie $t \in I$ (por. Ramsay, Silverman, 2005). Niech I będzie zbiorem zwartym takim, że $t_j \in I$, dla $j = 1, \dots, J$. Załóżmy, że funkcja $x_k(t)$ ma następującą reprezentację

$$x_k(t) = \sum_{b=0}^{B_k} c_{kb} \varphi_b(t), \quad t \in I, \quad k = 1, \dots, p, \quad (4)$$

gdzie $\{\varphi_b\}$ są ortonormalnymi funkcjami bazowymi, a $c_{k0}, c_{k1}, \dots, c_{kB_k}$ są współczynnikami.

Niech $\mathbf{x}_k = (x_{k1}, x_{k2}, \dots, x_{kJ})'$, $\mathbf{c}_k = (c_{k0}, c_{k1}, \dots, c_{kB_k})'$ oraz $\Phi(t)$ będzie macierzą wymiaru $J \times (B_k + 1)$ zawierającą wartości $\varphi_b(t_j)$, $b = 0, 1, \dots, B_k$, $j = 1, 2, \dots, J$, $k = 1, \dots, p$. Współczynnik c_k we wzorze (4) jest oszacowany metodą najmniejszych kwadratów tak, aby minimalizował funkcję:

$$S(c_k) = (\mathbf{x}_k - \Phi_k(t)c_k)' (\mathbf{x}_k - \Phi_k(t)c_k), \quad k = 1, \dots, p.$$

Różniczkując $S(c_k)$ względem wektora c_k , otrzymujemy estymator najmniejszych kwadratów postaci:

$$\hat{c}_k = (\Phi_k'(t)\Phi_k(t))^{-1}\Phi_k'(t)\mathbf{x}_k, \quad k = 1, \dots, p.$$

Stopień gładkości funkcji $x_k(t)$ zależy od wartości B_k (mała wartość B_k oznacza większe wygładzenie krzywej). Optymalną wartość B_k możemy ustalić przy pomocy bayesowskiego kryterium informacyjnego BIC (por. Schwarz, 1978; Shmueli, 2010).

Załóżmy, że dysponujemy n niezależnymi parami wartości (t_j, x_{kij}) , $k = 1, \dots, p$, $i = 1, \dots, n$, $j = 1, \dots, J$. Dane te wygładzamy za pomocą funkcji ciągłych postaci:

$$x_{ki}(t) = \sum_{b=0}^{B_{ki}} \hat{c}_{kib} \varphi_b(t), \quad k = 1, \dots, p, \quad i = 1, \dots, n, \quad t \in I.$$

Spośród wszystkich wartości $B_{k1}, B_{k2}, \dots, B_{kn}$ wybieramy jedną wspólną wartość B_k jako modę z wartości $B_{k1}, B_{k2}, \dots, B_{kn}$ oraz zakładamy, że funkcje $x_{ki}(t)$ mają postać

$$x_{ki}(t) = \sum_{b=0}^{B_k} \hat{c}_{kib} \varphi_b(t), \quad k = 1, \dots, p, \quad i = 1, \dots, n, \quad t \in I.$$

Dane postaci $\{x_{k1}(t), \dots, x_{kn}(t)\}$ noszą nazwę danych funkcjonalnych (por. Ramsay, Silverman, 2005).

Ogólnie założmy, że n niezależnych realizacji $\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_n(t)$ może być przedstawionych w postaci $\mathbf{x}_i(t) = \Phi(t)\hat{c}_i$ gdzie $\Phi(t)$ dane jest wzorem (3) oraz że wektory $\hat{c}_i = (\hat{c}_{i0}, \dots, \hat{c}_{iB_1}, \dots, \hat{c}_{ip0}, \dots, \hat{c}_{ipB_p})'$ są scentrowane, $i = 1, 2, \dots, n$.

Oznaczmy $\hat{C} = (\hat{c}_1, \hat{c}_2, \dots, \hat{c}_n)$. Wtedy

$$\hat{\Sigma}_c = \frac{1}{n} \hat{C} \hat{C}'.$$

Niech $\hat{\gamma}_1 \geq \hat{\gamma}_2 \geq \dots \geq \hat{\gamma}_s$ będą niezerowymi wartościami własnymi macierzy $\hat{\Sigma}_c$, a $\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \dots, \hat{\mathbf{u}}_s$ odpowiadającymi im wektorami własnymi, gdzie $s = \text{rank}(\hat{\Sigma}_c)$.

Ponadto k -ty układ główny procesu losowego $\mathbf{X}(t)$ wyznaczony na podstawie próby ma następującą postać:

$$(\hat{\lambda}_k = \hat{\gamma}_k, \hat{\mathbf{u}}_k(t) = \Phi(t)\hat{\mathbf{u}}_k), \quad k = 1, \dots, s.$$

Współrzędne rzutu i -tej realizacji $\mathbf{x}_i(t)$ procesu $\mathbf{X}(t)$ na kierunek wyznaczony przez k -tą funkcjonalną składową główną są równe:

$$\hat{U}_{ik} = \langle \hat{u}_k(t), x_i(t) \rangle = \langle \Phi(t) \hat{u}_k, \Phi(t) \hat{c}_i \rangle = \hat{u}'_k \langle \Phi(t), \Phi(t) \rangle \hat{c}_i = \hat{u}'_k \hat{c}_i,$$

dla $i = 1, 2, \dots, n$, $k = 1, 2, \dots, s$. Ogólnie współrzędne rzutu i -tej realizacji $x_i(t)$ procesu $X(t)$ na płaszczyznę wyznaczoną przez dwie pierwsze funkcjonalne składowe główne są równe:

$$(\hat{u}'_1 \hat{c}_i, \hat{u}'_2 \hat{c}_i), i = 1, 2, \dots, n.$$

4. DENDRYTOWA ANALIZA SKUPIEŃ

Omówimy teraz krótko metodę analizy skupień bazującą na dendrycie. Z punktu widzenia teorii grafów dendryt jest synonimem drzewa. Dendryt lub drzewo jest grafem spinającym (każde dwa wierzchołki łączy jakaś droga), który nie zawiera cykli. Minimalny dendryt, to taki dendryt, w którym suma wag przy krawędziach jest minimalna. Najczęściej wagami są odległości. Konstrukcja dendrytu podana została przez grupę matematyków wrocławskich w pracy Florek i inni (1951a). Metoda ta znana pod nazwą *taksonomia wrocławska* została spopularyzowana w pracach Florek i inni (1951b) oraz Perkal (1953). Istnieją dwie niezwykle proste procedury postępowania pozwalające skonstruować minimalny dendryt. Pierwszą z nich jest algorytm Kruskala:

1. Wybieramy krawędź o najmniejszej długości.
2. Z pozostałych krawędzi wybieramy tę o najmniejszej długości, która nie prowadzi do cyklu (z połączeń o jednakowych długościach wybieramy dowolne).
3. Powtarzamy poprzedni krok do zakończenia budowy minimalnego dendrytu.

Algorytm Prima rozpoczyna od drzewa składającego się z jednego wierzchołka, ciągle dodając najkrótszą krawędź drzewa.

Najkrótszy dendryt można wykorzystać w analizie skupień. Idea polega na usunięciu z minimalnego dendrytu wszystkich krawędzi, których długość jest większa od wspólnej wartości krytycznej d . Wyliczamy wartość średnią $\bar{x}$ oraz odchylenie standardowe s z długości wszystkich krawędzi najkrótszego dendrytu i przyjmujemy $d = \bar{x} + s$. Wierzchołki, które pozostają połączone w minimalnym dendrycie tworzą skupienie. Zaprezentowana tutaj metoda jest szerzej omówiona w sekcji 10.3 podręcznika Dziechciarza (2003).

5. WYNIKI BADAŃ EMPIRYCZNYCH

Analizą objęto 16 województw Polski ($n = 16$). Na prezentowanych dalej wykresach poszczególne województwa są oznaczone numerami przedstawionymi w tabeli 1.

Analizowane dane obejmują okres 11 lat, od 2003 do 2013 roku ($J = 11$). Każde województwo scharakteryzowano za pomocą 37 cech, zgrupowanych w 6 podzbiorach obrazujących różnorodne aspekty życia mieszkańców danego regionu, jak:

- wynagrodzenia i rynek pracy;

- opieka zdrowotna i społeczna;
- infrastruktura gospodarcza;
- zasoby mieszkaniowe;
- oświata, kultura i rekreacja oraz
- bezpieczeństwo i środowisko.

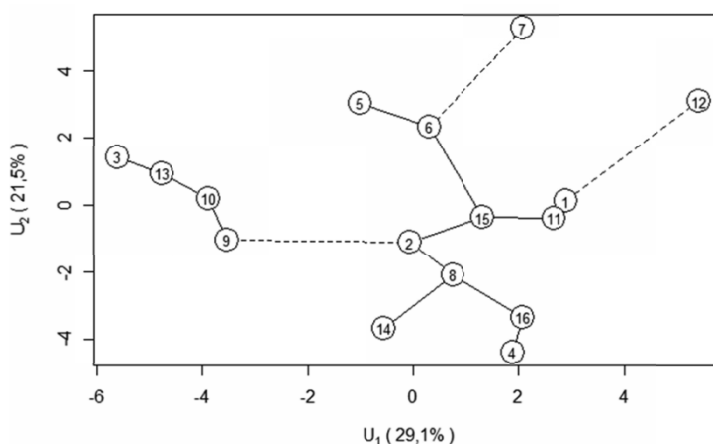
Tabela 1.

Numeryczne oznaczenia województw

Nr	Województwo	Nr	Województwo
1	dolnośląskie	9	podkarpackie
2	kujawsko-pomorskie	10	podlaskie
3	lubelskie	11	pomorskie
4	lubuskie	12	śląskie
5	łódzkie	13	świętokrzyskie
6	małopolskie	14	warmińsko-mazurskie
7	mazowieckie	15	wielkopolskie
8	opolskie	16	zachodniopomorskie

Źródło: opracowanie własne.

Dane pierwotne zostały poddane unitaryzacji zerowanej (patrz sekcja 2), a następnie przekształcone do wielowymiarowych danych funkcjonalnych. Posłużono się funkcjami bazowymi Fouriera. Przedział czasowy $I = [0, 11]$ został podzielony na momenty czasowe następująco: $t_1 = 0,5$ (2003), $t_2 = 1,5$ (2004), ..., $t_{11} = 10,5$ (2013). Następnie dla wszystkich 37 cech łącznie oraz oddzielnie dla każdej z 6 grup cech zostały skonstruowane funkcjonalne składowe główne. Każde z 16 województw zostało przedstawione jako punkt w układzie dwóch pierwszych funkcjonalnych składowych głównych. Kolejno dla wszystkich grup cech zostały zbudowane dendryty (por. Florek i inni, 1951) na podstawie tablicy wzajemnych odległości euklidesowych między województwami w przestrzeni wszystkich funkcjonalnych składowych głównych. Odległości te są równe odległościom euklidesowym w oryginalnej przestrzeni danych funkcjonalnych. Dendryty te rozpięto na punktach reprezentujących poszczególne województwa i wykorzystano do analizy skupień. W każdym z dendrytów policzono wartość średnią $\bar{x}$ długości krawędzi oraz odchylenie standardowe s tych długości. Krawędzie, których długość była większa od $d = \bar{x} + s$ zaznaczone są na rysunkach linią przerywaną. W ten sposób uzyskano podział województw na względnie jednorodne skupienia.

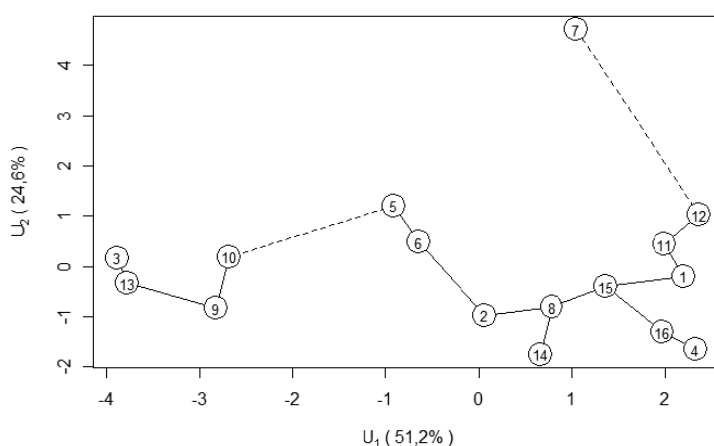


Rysunek 1. Województwa w układzie dwóch pierwszych funkcjonalnych składowych głównych – grupowanie na podstawie wszystkich 37 cech

Źródło: opracowanie własne.

Analiza przeprowadzona w oparciu o wszystkie 37 cech diagnostycznych pozwoliła wyodrębnić cztery grupy województw względnie jednorodnych pod względem poziomu życia mieszkańców. Dwie z nich to grupy jednoelementowe: woj. śląskie i woj. mazowieckie. Kolejną grupę utworzyły województwa: dolnośląskie, pomorskie, wielkopolskie, zachodniopomorskie, lubuskie, opolskie, małopolskie, kujawsko-pomorskie, łódzkie i warmińsko-mazurskie. Grupę czwartą utworzyły cztery województwa ściany wschodniej, często określane mianem „wschodniej ściany płaczu”: podkarpackie, podlaskie, świętokrzyskie i lubelskie. Województwa te są najuboższymi regionami Polski, a do momentu wejścia w struktury UE Bułgarii i Rumunii, były też uznawane za najuboższe w całej Wspólnocie. Niepokojącym wydaje się także fakt, iż w świetle badań prowadzonych przez OECD we współpracy z Ministerstwem Rozwoju Regionalnego a także danych z ostatniego Spisu Powszechnego, różnice w rozwoju gospodarczym i społecznym polskich regionów pogłębiają się. Z raportu „Przegląd Regionalny Polski 2012” opracowanego przez Ministerstwo Rozwoju Regionalnego wynika, iż mimo długoletnich działań oraz liczonej w miliardach euro pomocy z Unii Europejskiej przepaść między Polską wschodnią a resztą kraju niebezpiecznie rośnie. „Najwolniej dystans do średniej unijnej nadrabiały województwa o najniższym poziomie PKB per capita (województwa Polski Wschodniej i woj. zachodniopomorskie). Pięć województw Polski Wschodniej nadal znajdowało się w grupie 20 europejskich regionów NUTS 2 o najniższym poziomie PKB per capita, z poziomem tego wskaźnika w relacji do średniej unijnej od 42% (lubelskie) do 47% (świętokrzyskie)”.

Wykorzystując podzbiory cech diagnostycznych wyodrębniono grupy województw względnie jednorodnych pod względem poszczególnych aspektów kształtujących ogólny poziom życia mieszkańców. Uzyskane wyniki zaprezentowano na rysunkach 2–7.

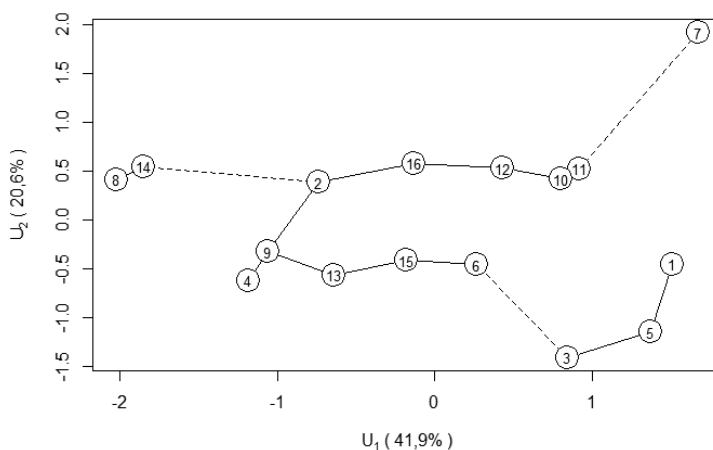


Rysunek 2. Województwa w układzie dwóch pierwszych funkcjonalnych składowych głównych – grupowanie na podstawie cech opisujących „wynagrodzenie i rynek pracy”

Źródło: opracowanie własne.

Cechy charakteryzujące wynagrodzenia i rynek pracy podzieliły województwa Polski na trzy grupy. Jedną z nich tworzą województwa: lubelskie, świętokrzyskie, podkarpackie i podlaskie. Województwa te, będące typowo rolniczym regionem naszego kraju, charakteryzuje relatywnie wysoki odsetek zatrudnionych w rolnictwie, mała konkurencyjność gospodarki, niski poziom bezpośrednich inwestycji zagranicznych w skali kraju – co niewątpliwie wpływa na bardzo niski poziom dochodów ludności oraz relatywnie wysokie bezrobocie. Drugą grupę, względnie jednorodną pod względem poziomu wynagrodzeń i sytuacji na rynku pracy, utworzyło 11 województw. Trzecią grupą stanowi województwo mazowieckie, w którym przeciętne miesięczne wynagrodzenia brutto w analizowanych 11 latach wahały się pomiędzy 129,8% (w roku 2003) a 123,1% (w roku 2013) średniej krajowej. Woj. mazowieckie jest jednym z dwóch województw Polski, w których przeciętne miesięczne wynagrodzenia przekraczają średnią krajową, przy czym dystans wynagrodzeń w drugim, tj. w woj. śląskim do średniej krajowej jest zdecydowanie mniejszy. W mazowieckim obserwuje się też relatywnie niski poziom bezrobocia.

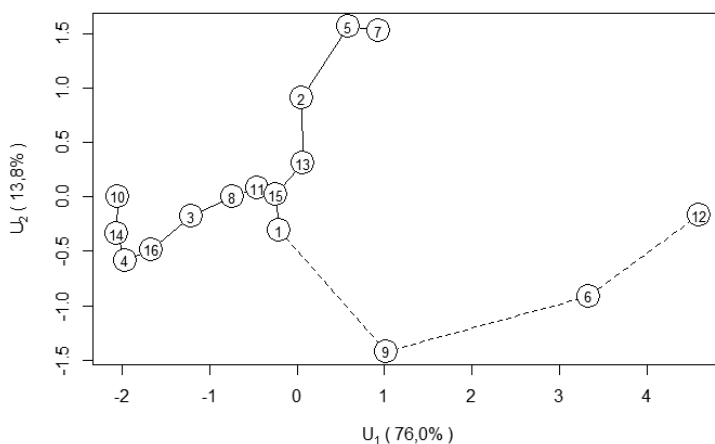
Pod względem „opieki zdrowotnej i społecznej” województwo mazowieckie – po raz kolejny – nie weszło w skład żadnej z grup względnie jednorodnych utworzonych z pozostałych województw. Wiele „przewag” województwa mazowieckiego wynika niewątpliwie z faktu, iż jest to region ze stolicą kraju, mamy tu zlokalizowaną dużą bazę akademicką z największą uczelnią medyczną w Polsce, dużą liczbę szpitali i lekarzy (przykładowo w roku 2010 na terenie tego województwa działało 14% ogółu szpitali publicznych i 12% ogółu szpitali niepublicznych, pracowało tu blisko 30 tys. (tj. 17%) spośród 172 tys. pracujących lekarzy w Polsce).



Rysunek 3. Województwa w układzie dwóch pierwszych funkcjonalnych składowych głównych – grupowanie na podstawie cech opisujących „opiekę zdrowotną i społeczną”

Źródło: opracowanie własne.

Grupy względnie jednorodne pod względem poziomu opieki zdrowotnej i społecznej utworzyły województwa: łódzkie z dolnośląskim i lubelskim, opolskie z warmińsko-mazurskim i grupę trzecią – dziesięć pozostałych województw.

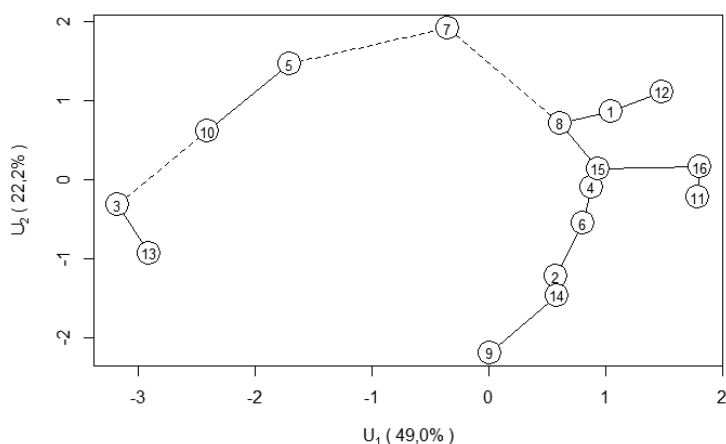


Rysunek 4. Województwa w układzie dwóch pierwszych funkcjonalnych składowych głównych – grupowanie na podstawie cech opisujących „infrastrukturę gospodarczą”

Źródło: opracowanie własne.

Aż 13. spośród szesnastu województw utworzyło grupę względnie jednorodną pod względem wyposażenia i stanu infrastruktury gospodarczej. Poza wspomnianą grupą znalazły się – postrzegane jako najzasobniejsze pod względem infrastruktury technicznej – województwa śląskie i małopolskie (usytuowane wzdłuż międzynarod-

wych szlaków komunikacyjnych i transportowych, z największym w kraju zagęszczeniem sieci drogowej i kolejowej oraz wodno-kanalizacyjnej) oraz podkarpackie, które cechuje słaba dostępność komunikacyjna, wynikająca z niedostatecznej i niskiej jakości sieci drogowej i kolejowej, niskie tempo rozwoju infrastruktury teleinformatycznej (zwłaszcza szerokopasmowego dostępu do Internetu) oraz wymagający rozbudowy i modernizacji stan infrastruktury komunalnej i energetycznej.

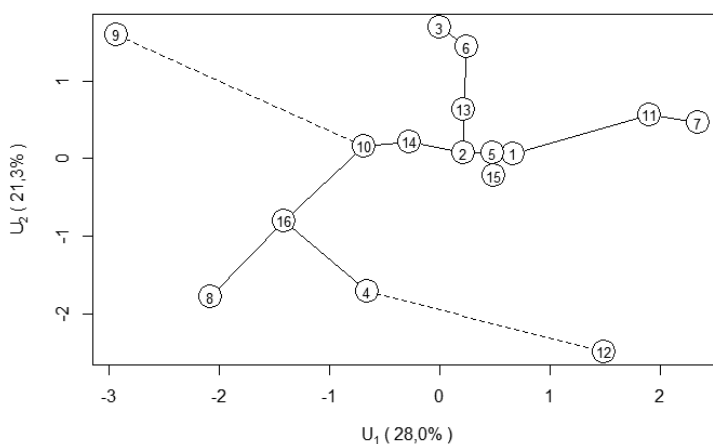


Rysunek 5. Województwa w układzie dwóch pierwszych funkcjonalnych składowych głównych – grupowanie na podstawie cech opisujących „zasoby mieszkaniowe”

Źródło: opracowanie własne.

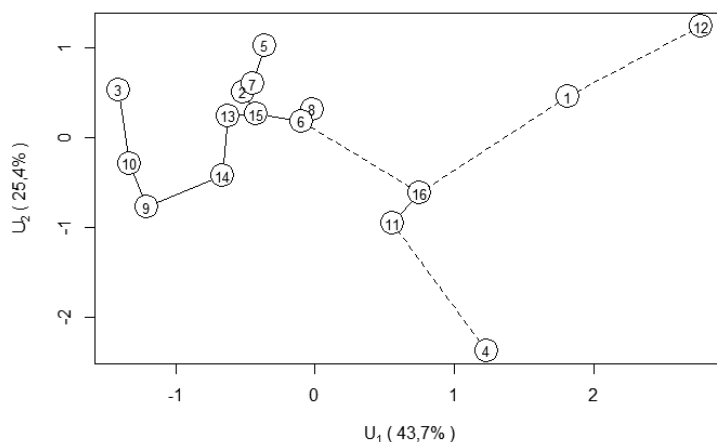
Analiza przeprowadzona w oparciu o cechy obrazujące stan zasobów mieszkaniowych w latach 2003–2013 pozwoliła na wyodrębnienie czterech grup województw o zbliżonym poziomie rozpatrywanych cech. Jedną z tych grup, tworzy (po raz kolejny „jednoosobowo”) województwo mazowieckie, w którym notuje się największą liczbę mieszkań na 1 tys. mieszkańców i największą powierzchnię użytkową mieszkania w przeliczeniu na 1 osobę.

W wyniku analizy przeprowadzonej w oparciu o cechy opisujące wyposażenie i stan „oświaty, kultury i rekreacji” w poszczególnych województwach wyodrębniono grupę względnie jednorodną w skład, której weszło aż 14 województw Polski. Poza nią znalazły się jedynie województwa śląskie i podkarpackie. „Oddalenie” województwa podkarpackiego od pozostałych wynika z faktu, iż mamy tu do czynienia z relatywnie największą liczbą klubów sportowych w przeliczeniu na 10 tys. mieszkańców ale też z jednym z najniższych – w skali kraju – odsetkiem dzieci objętych wychowaniem przedszkolnym. W śląskim, z kolei, mamy najslabsze wyposażenie szkół w komputery z dostępem do Internetu oraz relatywnie małą liczbę klubów sportowych oraz domów, klubów i ośrodków kultury w przeliczeniu na liczbę mieszkańców.



Rysunek 6. Województwa w układzie dwóch pierwszych funkcjonalnych składowych głównych – grupowanie na podstawie cech opisujących „oświatę, kulturę i rekreację”

Źródło: opracowanie własne.



Rysunek 7. Województwa w układzie dwóch pierwszych funkcjonalnych składowych głównych – grupowanie na podstawie cech opisujących „bezpieczeństwo i środowisko”

Źródło: opracowanie własne.

Grupowanie w oparciu o cechy opisujące „bezpieczeństwo i środowisko” doprowadziło do wyodrębnienia dwóch grup względnie jednorodnych, z których jedna liczyła jedenaście, a druga dwa województwa. Poziom rozpatrywanych cech w pozostałych województwach: śląskim, dolnośląskim i lubuskim, wyłączył je poza grupy względnie jednorodne. W każdym z tych trzech województw notuje się relatywnie wysoką przestępczość, w śląskim i dolnośląskim – najwyższe w skali kraju ilości wytwarzanych odpadów na 1 km² (lubuskie pod tym względem plasuje się na ostatnich pozycjach w rankingu województw). „Odmienność” lubuskiego jest także wynikiem

dużego zalesienia tego województwa – udział lasów w całkowitej powierzchni jest tu najwyższy w Polsce i przekracza 50% (podczas, gdy w kolejnym pod tym względem, woj. podkarpackim wynosi ok. 38%).

Podsumowując, warto zauważyć, że pozycję województwa śląskiego, w którym poziom życia mieszkańców w latach 2003–2013 był relatywnie najwyższy (rys. 1) determinowały przede wszystkim: wyposażenie w infrastrukturę gospodarczą, wysoki poziom wynagrodzeń, dobra sytuacja na rynku pracy oraz posiadane zasoby mieszkaniowe. O położeniu woj. mazowieckiego zdecydowały najwyższy w Polsce poziom wynagrodzeń, dobra sytuacja na rynku pracy, sytuacja w zakresie opieki zdrowotnej i społecznej oraz oświaty, kultury i rekreacji. Z kolei o położeniu czwartej, wyodrębnionej w oparciu o 37 cech diagnostycznych, grupy województw o relatywnie najniższym poziomie życia mieszkańców, tj. grupy woj. ściany wschodniej, zdecydowały: najniższe z obserwowanych w kraju wynagrodzenia, trudna sytuacja na lokalnych rynkach pracy oraz słabe zasoby mieszkaniowe.

LITERATURA

- Diechciarz J., (red.), (2003), *Ekonometria. Metody, przykłady, zadania*, Wydanie 2 poprawione, Wydawnictwo AE we Wrocławiu.
- Florek K., Łukaszewicz J., Perkal J., Steinhaus H., Zubrzycki S., (1951a), Sur la Liaison et la Division des Points d'un Ensemble Fini, *Colloquium Mathematicum*, 2, 282–285.
- Florek K., Łukaszewicz J., Perkal J., Steinhaus H., Zubrzycki S., (1951b), Taksonomia wrocławska, *Przegląd Antropologiczny*, 17, 193–211.
- Górecki T., Krzyśko M., Waszak Ł., Wołyński W., (2014), Methods of Reducing Dimension for Functional Data, *Statistics in Transition – New Series*, 15 (2), 231–242.
- Horváth L., Kokoszka P., (2012), *Inference for Functional Data with Applications*, Springer.
- Jacques J., Preda C., (2014), Model-Based Clustering for Multivariate Functional Data, *Computational Statistics & Data Analysis*, 71, 92–106.
- Kruskal J. B., (1956), On the Shortest Spanning Subtree of a Graph and the Travelling Salesman Problem, *Proceedings of the American Mathematical Society*, 7, 48–50.
- Perkal J., (1953), Taksonomia wrocławska, *Przegląd Antropologiczny*, 19, 209–221.
- Prim R. C., (1957), Shortest Connection Networks and Some Generalizations, *Bell System Technical Journal*, 36, 1389–1401.
- Ramsay J. O., Silverman B. W., (2005), *Functional Data Analysis*, Second Edition, Springer.
- Schwarz G., (1978), Estimating the Dimension of a Model, *Annals of Statistics*, 6, 461–464.
- Shmueli G., (2010), To Explain or to Predict? *Statistical Science*, 25 (3), 289–310.
- Walesiak M., (2014), Przegląd formuł normalizacji wartości zmiennych oraz ich własności w statystycznej analizie wielowymiarowej, *Przegląd Statystyczny*, 61 (4), 363–372.

OCENA ZRÓŻNICOWANIA POZIOMU ŻYCIA MIESZKAŃCÓW WOJEWÓDZTW
W LATACH 2003–2013 ZA POMOCĄ SKŁADOWYCH GŁÓWNYCH
DLA WIELOWYMIAROWYCH DANYCH FUNKCJONALNYCH ORAZ ANALIZY SKUPIEŃ

Streszczenie

W artykule przedstawiono ocenę zróżnicowania poziomu życia mieszkańców województw w latach 2003–2013. Do oceny zastosowano analizę składowych głównych dla wielowymiarowych danych funkcjonalnych oraz dendrytową analizę skupień. Metody te pozwoliły na wyodrębnienie względnie jednorodnych grup województw o zbliżonym poziomie rozpatrywanych cech dla całego rozpatrywanego okresu łącznie.

Słowa kluczowe: wielowymiarowe dane funkcjonalne, funkcjonalna analiza danych, analiza składowych głównych

ESTIMATION OF DIVERSITY OF LIVING STANDARDS IN POLISH VOIVODSHIPS
IN 2003–2013 USING PRINCIPAL COMPONENTS
FOR MULTIDIMENSIONAL FUNCTIONAL DATA AND CLUSTER ANALYSIS

Abstract

The paper presents an estimation of life standard diversity for residents of Polish voivodships in 2003–2013. The principal component analysis was applied for multidimensional functional data and the dendrite method was used for cluster analysis. These methods made it possible to isolate relatively homogeneous groups of voivodships that had similar values of characteristics under consideration, for the whole period at issue.

Keywords: multivariate functional data, functional data analysis, principal components analysis

SPRAWOZDANIA

ANNA JUREK¹, ELŻBIETA ZALEWSKA²

SPRAWOZDANIE Z XXXIV MIĘDZYNARODOWEJ KONFERENCJI NAUKOWEJ WIELOWYMIAROWA ANALIZA STATYSTYCZNA (WAS 2015)

W dniach od 16 do 18 listopada 2015 r. w Centrum Szkoleniowo-Konferencyjnym Uniwersytetu Łódzkiego odbyła się XXXIV międzynarodowa konferencja naukowa Wielowymiarowa Analiza Statystyczna WAS 2015 (*The 34th International Annual Conference on Multivariate Statistical Analysis MSA 2015*). Konferencja zorganizowana została przez Instytut Statystyki i Demografii UŁ, Katedrę Metod Statystycznych UŁ, Polskie Towarzystwo Statystyczne (PTS) oraz Komitet Statystyki i Ekonometrii Polskiej Akademii Nauk. Funkcję Przewodniczącego Komitetu Organizacyjnego pełnił prof. dr hab. Czesław Domański, a obowiązki sekretarzy naukowych konferencji sprawowały mgr Anna Jurek oraz mgr Elżbieta Zalewska.

XXXIV Konferencja Wielowymiarowa Analiza Statystyczna WAS 2015 objęta została honorowym patronatem Prezydenta Miasta Łodzi – Hanny Zdanowskiej. Jej organizacja została dofinansowana przez Narodowy Bank Polski w ramach projektu edukacyjnego wpisującego się w priorytetowe obszary edukacji ekonomicznej NBP, Polską Akademię Nauk oraz Urząd Miasta Łodzi – w ramach otwartego konkursu ofert w formie wsparcia realizacji zadania publicznego w zakresie przeprowadzenia działań zmierzających do przekształcenia Łodzi w centrum wydarzeń naukowych i edukacji wyższej o znaczeniu międzynarodowym oraz propagujących osiągnięcia naukowe Łodzi. Partnerem konferencji była także firma StatSoft Polska Sp. z o.o.

Głównymi celami XXXIV edycji Konferencji Wielowymiarowa Analiza Statystyczna były: prezentacja najnowszych osiągnięć z zakresu wielowymiarowej analizy statystycznej oraz wymiana doświadczeń, będących wynikiem jej stosowania. Szeroki zakres tematów statystyki matematycznej, metod statystyki wielowymiarowej oraz ich innowacyjnych aplikacji będących przedmiotem publikacji można ująć następująco: rozkłady wielowymiarowe, testy statystyczne, metody nieparametryczne, ana-

¹ Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Katedra Metod Statystycznych, ul. POW 3/5, 90-255 Łódź, Polska.

² Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Katedra Metod Statystycznych, ul. POW 3/5, 90-255 Łódź, Polska, autor prowadzący korespondencję – e-mail: zalewska.e@uni.lodz.pl.

liza czynnikowa, analiza skupień, analiza dyskryminacyjna, analiza wariancji i regresji, metody bayesowskie, analizy Monte Carlo, data mining, procedury odporne, analiza danych cenzurowanych, rozpoznawanie obrazów, analizy stochastyczne oraz zastosowanie metod statystycznych w naukach ekonomicznych, marketingu, finansach, ubezpieczeniach, rynku kapitałowym, zarządzaniu ryzykiem oraz zastosowaniu m.in. w medycynie i opiece zdrowotnej.

W Konferencji wzięły udział 93 osoby z różnych ośrodków akademickich w Polsce – Białystok, Gdańsk, Katowice, Kraków, Poznań, Szczecin, Toruń, Warszawa, Wrocław, przedstawiciele Głównego Urzędu Statystycznego, Urzędu Statystycznego w Łodzi, Narodowego Banku Polskiego, firmy Bosch i Siemens, a także goście z zagranicy oraz studenci UŁ. Po otwarciu konferencji przez Przewodniczącego Komitetu Organizacyjnego – Czesława Domańskiego, Rektora Uniwersytetu Łódzkiego – Włodzimierza Nykiela oraz Dziekana Wydziału Ekonomiczno-Socjologicznego UŁ – Pawła Starosty uczestnicy konferencji wysłuchali pierwszego wykładu zaproszonego: *Proces Statystyczny jako Społeczny Proces Kwantyfikacji* wygłoszonego przez Włodzimierza Okrasę z Uniwersytetu Kardynała Stefana Wyszyńskiego w Warszawie. Drugi wykład zaproszony przedstawili Krzysztof Najman, Kamila Migdał-Najman oraz Mirosław Szreder z Uniwersytetu Gdańskiego, na którym zaprezentowano temat: „*Big Data – nowe możliwości, nowe ograniczenia*”. Tradycją konferencji MSA jest sesja historyczna poświęcona wybitnym przedstawicielom polskiej myśli statystycznej. *Wspomnienie o Włodzimierzu Krysiem* wygłosił Tadeusz Gerstenkorn z Uniwersytetu Łódzkiego. Kolejny referat pt.: *Wątki statystyczne w pracach Jana Długosza* zaprezentował Tadeusz Kowaleski – Uniwersytet Łódzki. Jan Kordos (Wyższa Szkoła Menedżerska w Warszawie) zaprezentował dorobek naukowy Profesora Tadeusza Walczaka w referacie – *Moja współpraca z Profesorem Tadeuszem Walczakiem*. Ostatni referat w tej sesji przedstawił Czesław Domański, zaprezentował postać *Stanisława Staszica jako sympatyka Łodzi*. Podczas konferencji odbyło się posiedzenie członków Rady Głównej Polskiego Towarzystwa Statystycznego (PTS), któremu przewodniczył prof. dr hab. Czesław Domański – prezes PTS.

Oprócz 2 wykładów zaproszonych oraz 4 referatów w sesji jubileuszowo-historycznej uczestnicy konferencji wygłosili 59 referatów (w tym 11 referatów wygłoszonych w języku angielskim) – poniżej przedstawiono ich tytuły (w kolejności alfabetycznej według nazwisk autorów):

- Balcerzak Adam, Pietrzak Michał – *Jakość instytucji a poziom łącznej produktywności czynników produkcji w Unii Europejskiej*
Baszczyńska Aleksandra – *Redukcja efektu brzegowego w estymacji jądrowej dystrybuanty zmiennej losowej*
Bąk Tomasz – *Adaptacyjne losowanie przestrzenne oparte na krigingu*
Białek Jacek – *Aproksymacja indeksu Fishera za pomocą indeksu Lloyda–Moultona: badanie symulacyjne i empiryczne*

- Budnikas Germanas – *Discovery of aggression property from person behavior in social media*
- Budny Katarzyna – *Estymacja funkcji momentów wektora losowego opartych na definicji potęgi wektora*
- Ceranka Bronisław, Graczyk Małgorzata – *O pewnych własności układów doświadczalnych*
- Cheba Katarzyna – *Zastosowanie rachunku wektorowego do badania jednorodności czasowej rozwoju obiektów społeczno-gospodarczych*
- Dehnel Grażyna, Klimanek Tomasz – *Niepełnosprawność w Narodowych Spisach Powszechnych 2002 i 2011 – porównanie zakresu informacyjnego*
- Domański Czesław – *Znaczenie modeli i statystyk w procesie wnioskowania statystycznego*
- Dudek Andrzej – *Comparison of selected classical and modern algorithms of cluster analysis*
- Dziechciarz Józef – *Metody statystyki w identyfikacji brakujących kompetencji miękkich*
- Fałdziński Marcin – *The Multidimensional Value-at-Risk backtesting*
- Gołata Elżbieta – *Dylematy szacunku liczby ludności – implikacje dla rynku prac*
- Gruchociak Hanna – *Próba analizy porównawczej szeregów czasowych na przykładzie chodu pieszych*
- Grzenda Wioletta – *Estymacja bayesowska modeli predykcyjnych i jej wpływ na dokładność regresji i klasyfikacji*
- Jewczak Maciej – *Skłonność do płacenia za wybrane świadczenia zdrowotne – różnice i podobieństwa metod grupowania*
- Jędrzejczak Alina, Kubacki Jan – *Estimation of income for small areas using Rao-Yu model*
- Kończak Grzegorz, Rajda-Tasior Angelina – *O monitorowaniu złożonych wielowymiarowych procesów*
- Korzeniewski Jerzy – *Selekcja zmiennych w analizie skupień danych binarnych*
- Kozera Agnieszka, Głowicka-Wołoszyn Romana – *Zjawisko autokorelacji przestrzennej w ocenie samodzielności finansowej gmin województwa wielkopolskiego*
- Kręzolek Dominik – *Miara ryzyka GlueVaR a postawy wobec ryzyka – zastosowanie na rynku metali nieżelaznych*
- Krzyśko Mirosław, Górecki Tomasz, Wołyński Waldemar – *Analiza korelacji dla wielowymiarowych danych funkcjonalnych*
- Kubus Mariusz – *Lokalnie regularyzowana regresja liniowa w wycenie nieruchomości*
- Kurzawa Izabela, Stanisławska Joanna – *Zastosowanie regresji kwantylowej do badania zróżnicowania najpilniejszych wydatków gospodarstw domowych w Polsce w relacji miasto-wieś*
- Lucińska Anna – *Metody wyznaczania hedonicznego indeksu cen malarstwa na polskim rynku aukcyjnym*
- Łuczak Aleksandra – *Rozmyta i rozszerzona interwałowa metoda TOPSIS. Analiza porównawcza*

- Łukaszonek Wojciech – *Wielowymiarowa i zdynamizowana klasyfikacja województw na podstawie charakterystyk szkolnictwa wyższego obserwowanych w latach 2002–2013*
- Manikowski Arkadiusz – *Podjęcie wielomodelowe w prognozowaniu zapotrzebowania na pieniądź*
- Markowicz Iwona, Bieszk-Stolorz Beata – *Tablice trwania w bezrobociu*
- Mikulec Artur – *Charakterystyka populacji przedsiębiorstw zlikwidowanych w województwie łódzkim*
- Misztal Małgorzata – *O zastosowaniu wybranych technik ordynacyjnych do analizy wyników wyborów parlamentarnych*
- Parys Dariusz – *The modification of Hochberg-Hommel type step-up procedure*
- Pekasiewicz Dorota – *Semiparametryczna estymacja bootstrapowa kwantyli rozkładów zmiennych losowych*
- Pietrzyk Radosław, Rokita Paweł – *Residual Wealth at Risk – the concept and main properties*
- Piontek Krzysztof – *Analiza mocy wybranych testów wstecznych VaR bazujących na szeregach przekroczeń – podejście symulacyjne*
- Podogrodzka Małgorzata – *Przestrzenna konwergencja feminizacji starości demograficznej w Polsce*
- Ptak-Chmielewska Aneta – *Application of methods of multivariate statistical analysis in banking*
- Roszk Wojciech – *Statystyczna integracja badań częściowych – techniki i porównanie metod*
- Roszko-Wójtowicz Elżbieta, Białek Jacek – *Statystyczna analiza czynników determinujących innowacyjność krajów UE*
- Rozmus Dorota – *Porównanie stabilności czynnikowej metody odległości probabilistycznej oraz podejścia zagregowanego w taksonomii*
- Sitek Grzegorz – *Estymacja parametrów mieszanin regresji*
- Skródzka Iwona – *Gospodarka oparta na wiedzy w krajach Unii Europejskiej – analiza przestrzenna*
- Sokołowski Andrzej, Markowska Małgorzata – *Iterative method for linear ordering of multidimensional objects*
- Stapor Katarzyna, Smolarczyk Tomasz, Fabian Piotr – *Heteroscedastic discriminant analysis combined with feature selection for credit scoring*
- Szczepocki Piotr – *O numerycznych aproksymacjach funkcji przejścia dla procesów dyfuzji*
- Sztemberg-Lewandowska Mirosława – *Osiągnięcia uczniów na II-IV etapie edukacji z wykorzystaniem funkcjonalnej analizy głównych składowych*
- Szymańska Anna – *Zastosowanie modelu Bülmanna-Strauba do estymacji stawek składki netto zależnych od wieku ubezpieczonych w ubezpieczeniu odpowiedzialności cywilnej posiadaczy pojazdów mechanicznych*
- Szymczak Michał – *An influence of observable data on estimates of the modified Weibull distribution parameters*

- Trzpiot Grażyna – *Globalne starzenie się – natura ryzyka długowieczności*
- Trzpiot Grażyna, Orwat-Acedańska Agnieszka, Ojrzyńska Anna – *Przestrzenna regresja kwantylowa w analizie umieralności*
- Walesiak Marek – *Wybór formuły normalizacji wartości zmiennych w analizie skupień*
- Więtczak Michał – *Wykorzystanie skoringu marketingowego w zapobieganiu migracji klientów banków*
- Wiktorowicz Justyna – *Wielowymiarowa analiza aktywności ekonomicznej pokolenia 50+ – szanse i bariery*
- Wilak Kamil – *Warstwy i jednostki losowania pierwszego stopnia w Badaniu Aktywności Ekonomicznej Ludności*
- Wywiał Janusz – *Evaluation of necessary sample size asserting normality distribution of sample mean*
- Zaborski Artur – *Prezentacja zmian zachodzących w preferencjach z wykorzystaniem skalowania dynamicznego*
- Zieliński Wojciech – *Przedział ufności dla wskaźnika struktury w dwuwarstwowej populacji skończonej*
- Żądło Tomasz – *On prediction of future small areas' characteristics*

Kolejna, XXXV międzynarodowa konferencja naukowa **Wielowymiarowa Analiza Statystyczna, WAS 2016** (*The 35th International Annual Conference on Multivariate Statistical Analysis, MSA 2016*) odbędzie się w Łodzi w dniach od 7 do 9 listopada 2016 r. Przewodnictwem Komitetu Organizacyjnego przyszłorocznej edycji objął prof. dr hab. Czesław Domański, a sekretarzami naukowymi zostali: mgr Piotr Szczepocki oraz dr hab. inż. Jacek Białek.

