ROK VOL. 15 Nr No. 3

Wrzesień September 2017

Zarządzanie i Finanse

Journal of Management and Finance

ISSN 2084-5189

http://jmf.wzr.pl

Rada Naukowa
prof. dr hab. Jerzy Bieliński — przewodniczący — Uniwersytet Gdański;
prof. Wojciech Charemza, University of Leicester;
prof. Halina Frydman, Leonard N. Stern School of Business;
prof. dr hab. Krzysztof Jajuga, Uniwersytet Ekonomiczny we Wrocławiu;
prof. dr hab. Stanisław Kasiewicz, Szkoła Główna Handlowa;
Agnieszka Kwapisz, Ph.D., College of Business, Montana State University;
Wojciech Piotrowicz, Ph.D., Saïd Business School;
Sofia Brito Ramos, Ph.D., ISCTE Business School;
prof. Gopichand C. Tikkiwal, Jai Narain Vyas University

Redaktor Naczelny
prof. UG dr hab. Krzysztof Najman
Z-ca Redaktora Naczelnego
prof. UG dr hab. Kamila Migdał-Najman
Redaktorzy językowi
prof. UG dr hab. Aneta Lewińska (język polski)
David Gagan, M.A. (język angielski)
Redaktor statystyczny
prof. UG dr hab. Anna Zamojska
Redaktorzy tematyczni
dr hab. Ewelina Sokołowska, dr Olga Martyniuk (Finanse), dr Jacek Maślankowski (Informatyka ekonomiczna), prof. UG dr hab. Kamila Migdał-Najman (Metody ilościowe), dr Aleksandra Wiercińska (Rachunkowość), prof. UG dr hab. Paweł Antonowicz, prof. UG dr hab. Tomasz Kawka, dr inż. Ewa Malinowska, dr Marek Rutka, dr Krzysztof Szczepaniak (Zarządzanie)
Redaktor numeru
dr Ewa Majerowska

Copyright by Wydział Zarządzania Uniwersytetu Gdańskiego
81-824 Sopot, ul. Armii Krajowej 101

ISSN 2084-5189
Wersja drukowana czasopisma ma charakter referencyjny.
W latach 2003–2011 czasopismo ukazywało się pod tytułem
„Prace i Materiały Wydziału Zarządzania Uniwersytetu Gdańskiego".

Korekta
Jerzy Toczek
Skład i łamanie
Mariusz Szewczyk

Druk
Zakład Poligrafii Uniwersytetu Gdańskiego, Sopot, ul. Armii Krajowej 119/121
tel. 523-13-75, 523-14-49, e-mail: poligraf@gnu.univ.gda.pl

Spis treści

Krzysztof Kompa, Dorota Witkowska, Czy obecność kobiet w kierownictwie wpływa na rentowność kapitału własnego spółek notowanych na GPW? 5

Joanna Olbryś, Michał Mursztyn, Elastyczność rynku jako jeden z wymiarów płynności Giełdy Papierów Wartościowych w Warszawie S.A. 21

Magdalena Gostkowska-Drzewicka, Ewa Majerowska, Oddziaływanie wybranych czynników kształtujących strukturę kapitału spółek notowanych na GPW w Warszawie 37

Błażej Kochański, 20+ years of Polish residential real estate prices – a house price index proxy 51

Andrzej Karpio, Dorota Żebrowska-Suchodolska, Analiza porównawcza aktywnych i pasywnych strategii inwestycyjnych otwartych funduszy emerytalnych i otwartych funduszy zrównoważonych w latach 2014–2017 63

Dorota Żebrowska-Suchodolska, Podobieństwo miar opartych na momentach cząstkowych wyznaczonych dla funduszy akcyjnych w latach 2004–2015 75

Stanisław Wieteska, Małgorzata Jeziorska, Piotr Majewski, Identyfikacja pola ubezpieczeniowego i ocena ryzyka eksploatacji samochodowych pojazdów elektrycznych w ubezpieczeniu autocasco 87

Agnieszka Pobłocka, Tworzenie rezerwy IBNR metodami deterministycznymi na potrzeby wypłacalności w zakładach ubezpieczeń majątkowych 101

Ewa Wycinka, Tablice wieloprzyczynowe w ocenie ryzyka niewypłacalności. Podejście kohortowe 121

Tomasz Jastrzębski, Zastosowanie symulacji Monte Carlo z korektą Imana-Conovera w pomiarze ryzyka systemowego 139

Adam Marszk, Structure of the US investment company industry over the period 2000 to 2017: substitution analysis 157

Barbara Batóg, Jacek Batóg, Analiza czynników wpływających na zmiany eksportu krajów Unii Europejskiej 175

Hanna Godlewska-Majkowska, Dorota Perło, Zastosowanie metody wagowo-korelacyjnej i modelowania miękkiego do analizy atrakcyjności inwestycyjnej powiatów w Polsce 187

Aleksandra Łuczak, Modelowanie rozwoju lokalnego z wykorzystaniem opinii wybranych grup ekspertów 215

Maciej Nowak, Tadeusz Trzaskalik, Zarządzanie zdolnością produkcyjną z wykorzystaniem interaktywnego programowania dynamicznego 231

Beata Jackowska, Metody analizy przeżycia w badaniu zróżnicowania rozkładów czasu istnienia przedsiębiorstw ze względu na rodzaj działalności 247

Iwona Markowicz, Modele trwania firm w województwie zachodniopomorskim według grup sekcji PKD 265

Tomasz Pisula, Zastosowanie ensemble klasyfikatorów do oceny ryzyka upadłości przedsiębiorstw na przykładzie firm sektora produkcyjnego działających na Podkarpaciu 279

Kamila Migdał-Najman, Sylwia Badowska, Wykorzystanie samouczących się sieci neuronowych w analizie zachowań zakupowych i identyfikacji ich wzorców wśród konsumentów w wieku 60 lat i więcej 295

Radosław Murkowski, Płeć a zaawansowanie procesu starzenia się ludności państw europejskich 309

Beata Bieszk-Stolorz, Modele regresji w ocenie wpływu płci na formę wyjścia z bezrobocia rejestrowanego 325

Rumiana Górska, Piotr Staszkiewicz, Zastosowanie algorytmu lasów losowych do prognozowania modyfikacji opinii biegłego rewidenta 339

Roman Machuga, Modelowanie systemu informacyjnego rachunkowości zarządczej na przykładzie przedsiębiorstwa branży spożywczej 349

Krzysztof Kompa*
Dorota Witkowska**

Czy obecność kobiet w kierownictwie wpływa na rentowność kapitału własnego spółek notowanych na GPW?

Wstęp

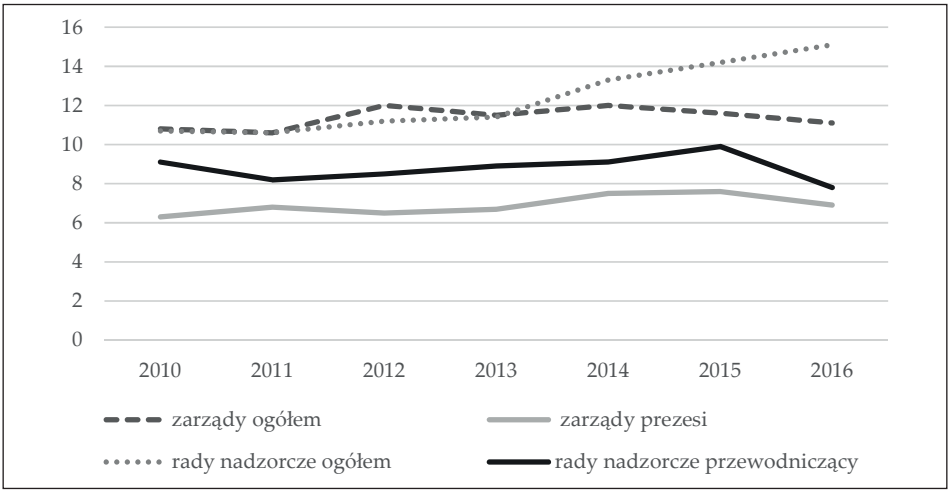
W Polsce niemal jedna trzecia pracujących na własny rachunek to kobiety, są one również założycielami lub zarządzającymi co trzeciej firmy w Polsce [Lewiatan, 2012, s. 18]. Trudno jest jednak znaleźć kobiety wśród menedżerów dużych firm, w tym spółek notowanych na GPW, a oczekuje się przyjęcia przez Komisję Europejską dyrektywy o 40% parytecie w radach nadzorczych spółek publicznych. W 2012 r. zaledwie 142 kobiety pełniły funkcje w zarządach i 299 w radach nadzorczych 396 polskich spółek giełdowych, a tylko 24 pełniły funkcje prezesa zarządu [tamże, s. 18]. W kolejnych latach sytuacja nie uległa zasadniczym zmianom, na co wskazują analizy przeprowadzane na koniec czerwca każdego roku.

W ciągu siedmiu lat 2010–2016 udział kobiet w zarządach (rys. 1) zwiększył się jedynie o niecałe 3% (a w porównaniu z 2012 rokiem spadł o 7,5%), chociaż w radach nadzorczych nastąpił istotny wzrost o ponad 41% (a 35% w relacji do roku 2012), co jednak w dalszym ciągu nie jest wystarczającym osiągnięciem w świetle zaleceń Komisji Europejskiej.

Nasuwa się zatem pytanie, czy istotny udział kobiet w kierownictwie dużych spółek jest potrzebny i jakie są argumenty za taką właśnie strukturą organów statutowych. Oprócz argumentów, że obecność kobiet poprawia praktyki korporacyjne spółek, m.in. lepsze zarządzanie ryzykiem i rozwiązywanie konfliktów interesów, pojawiają się również argumenty natury ekonomicznej, z których wynika, że obecność kobiet w organach statutowych przyczynia się do poprawy wyników finansowych spółek [Brown i inni, 2002; Radwan, 2010; Carter, Wagner, 2011; Curtis i inni, 2012; Bohdanowicz 2011a].

* Dr inż., Katedra Ekonometrii i Statystyki, Wydział Zastosowań Informatyki i Matematyki SGGW w Warszawie, ul. Nowoursynowska 166, 02-787 Warszawa, krzysztof_kompa@sggw.pl
** Prof. dr hab., Katedra Finansów i Strategii Przedsiębiorstwa, Wydział Zarządzania UŁ, ul. Jana Matejki 22/26, 90-237 Łódź, dorota.witkowska@uni.lodz.pl

Rysunek 1. Procentowy udział kobiet w organach statutowych spółek notowanych na rynku głównym GPW w Warszawie



Źródło: Opracowanie własne na podstawie danych Notoria Serwis.

W związku z tym celem prowadzonych analiz jest zbadanie występowania zależności między udziałem kobiet w zarządach i radach nadzorczych spółek Skarbu Państwa i należących do indeksu giełdowego WIG30 a rentownością kapitału własnego tych spółek. Badania przeprowadzono na podstawie danych pochodzących z Notoria Serwis dotyczących wyróżnionych 32 spółek publicznych. Określono strukturę organów kierowniczych i rentowność kapitału własnego (ROE) tych spółek w latach 2010–2015. Badania występowania współzależności między strukturą organów kierowniczych i rentownością spółek przeprowadzono za pomocą analizy korelacji i regresji.

Artykuł powstał w wyniku realizacji projektu badawczego „Zmiana pozycji kobiet na rynku pracy. Analiza sytuacji w Polsce na tle Unii Europejskiej w latach 2002–2014" Nr 2015/17/B/HS4/00930 finansowanego ze środków Narodowego Centrum Nauki.

1. Przegląd literatury

Argumenty, które w literaturze przedmiotu wymienia się w kontekście zwiększonego udziału kobiet w organach kierowniczych, dzieli się na dwie grupy: społeczne i ekonomiczne [Campbell, Minguez-Vera, 2010; Bohdanowicz, 2010]. Do tych pierwszych zalicza się twierdzenie, że nie można dyskryminować członków organów statutowych ze względu na płeć. Stąd idea ustanawiania parytetów płci, co ma już miejsce m.in. w Norwegii, Szwajcarii, Belgii, Niemczech i Francji oraz jest w trakcie uchwalania przez Komisję Europejską. Warto w tym

miejscu wspomnieć, że idea nie jest powszechnie popierana w Europie (np. nie została uchwalona w Szwecji [Radwan, 2010]), a sam pomysł jest poddawany krytyce, m.in. wskazuje się, że narzucanie parytetu i zwiększanie reprezentacji kobiet we władzach spółek może być sprzeczne z interesami właścicieli, a także może stanowić przejaw dyskryminacji mężczyzn [Oplustil, 2011, s. 10–11].

Pojawiają się również przykłady poddające w wątpliwość sens dywersyfikacji organów kierowniczych, wskazujące jednocześnie wątpliwe etycznie działania, kiedy to – wobec braku kandydatek posiadających stosowne uprawnienia – te same kobiety (zwane „złotymi spódniczkami") obejmowały funkcje w organach statutowych wielu norweskich spółek, tzw. zjawisko *overboarding* [Walat, 2010], podobne zjawisko zaobserwowano również we Francji. Innym argumentem przemawiającym przeciwko restrykcyjnym regulacjom prawnym w tym zakresie może być fakt, że narzucenie parytetu płci w Norwegii wpłynęło na zmniejszenie się liczby spółek publicznych o jedną czwartą, tj. z 600 na 450 [Oplustil, 2011, s. 10], z powodu dolegliwych represji grożących spółkom publicznym za nieprzestrzeganie wspomnianych regulacji prawnych.

W rozważaniach dotyczących udziału kobiet w menedżmencie często zwraca się uwagę na to, że grupy zdywersyfikowane charakteryzują się szerszym wykształceniem i większą kreatywnością, co pozwala na wypracowanie większej liczby rozwiązań poprzez zróżnicowanie poglądów członków organów statutowych, chociaż przyczynia się to jednocześnie do wydłużonego czasu podejmowania decyzji. W związku z tym postuluje się dywersyfikację rad nadzorczych, natomiast w przypadku zarządów zaleca się jednorodność wśród decydentów z uwagi na większą kooperację i mniej powodów do konfliktów, a zatem szybsze dopracowanie wspólnych decyzji [Bohdanowicz, 2010; 2011a; 2011b]. Wskazuje się również, że zwiększona liczba kobiet w kierownictwie spółek jest dobrze postrzegana przez inwestorów, chociaż zdarzają się sytuacje, w których powołuje się kobiety menedżerów w obliczu słabnących, czy wręcz złych wyników przedsiębiorstw, aby obarczyć je winą za niepowodzenie – jest to tzw. zjawisko szklanego klifu (*glass cliff*), opisane w pracy [Ryan, Haslam, 2007].

Spośród argumentów ekonomicznych postulujących dywersyfikację organów statutowych wymienia się poprawę wyników finansowych spółek wraz ze zwiększeniem się reprezentacji kobiet w organach kierowniczych spółek, aczkolwiek część badań empirycznych przeczy tej tezie lub wskazuje na brak istotnych relacji między zmiennymi opisującymi oba zjawiska. Dodatnią współzależność zwiększonego udziału kobiet w kierownictwie do wyników spółek zaobserwowano m.in. w odniesieniu do takich wskaźników, jak: rentowność kapitału własnego (ROE), rentowność sprzedaży

(ROS), zwrot z zainwestowanego kapitału (ROIC) oraz całkowita stopa zwrotu dla akcjonariuszy (TSR), a także wartość rynkowa [Carter i inni, 2003; 2010; Catalyst, 2004; Campbell, Minguez-Vera, 2008; Carter, Wagner, 2011; Bear i inni, 2010; Desvauxi i inni, 2007; Devillard i inni, 2012; Curtis i inni, 2012; Lenard i inni, 2014; McKinsey, 2007; Smith i inni, 2006], a dla Polski [Lisowska i inni, 2014, s. 4, 21–29; Bohdanowicz, 2011b].

Istnieją także publikacje, które wskazują na negatywną zależność pomiędzy zróżnicowaniem kierownictwa spółek pod względem płci i wynikami finansowymi [Zahra, Stanton, 1988; Shrader i inni, 1997; Adams, Ferreira, 2009; Lee, James, 2007; Adams i inni, 2009; Ahern, Dittmar, 2012], a także raportujące brak związku [Rose, 2007; Farrel, Hersch, 2005; Wang, Clift, 2009], a dla spółek z GPW [Kompa i inni, 2016].

2. Opis danych i metod analizy

Ocena wpływu obecności kobiet w kierownictwie spółek na wyniki finansowe tych spółek jest w literaturze opisywana dość enigmatycznie, bez wskazania, w jaki sposób dokonywano pomiaru udziału kobiet w organach statutowych. A przecież obserwacje tego typu tworzą szeregi czasowe momentów, zatem niezbędne jest określenie odpowiednich punktów pomiarowych. Osobną kwestią jest sam pomiar struktury organów kierowniczych. Oprócz tego należy ustalić lub założyć długość okresu, w jakim można mówić o rzeczywistym wpływie kobiet na wyniki finansowe spółek. Innymi słowy, wymagane jest ustalenie zależności przyczynowej między pojawieniem się określonej struktury menedżmentu a pomiarem wyniku finansowego, bowiem nie można mówić o natychmiastowym efekcie zmiany sposobu zarządzania spółką, a z taką (przynajmniej jakościową) zmianą mamy do czynienia wraz z dywersyfikacją pod względem płci zarządów i rad nadzorczych (w szczególności dotyczy to tych ostatnich). W prezentowanych badaniach przyjęto założenie, że minimalny okres, w jakim można obserwować efekty decyzji wynikających z wprowadzenia kobiet do organów statutowych, wynosi przynajmniej pół roku. W związku z tym dane dotyczące składów zarządów i rad nadzorczych wzięto z 30 czerwca każdego roku, a rentowność kapitału własnego spółek (ROE) na koniec roku.

Badania zrealizowano dla spółek wchodzących w listopadzie 2016 r. w skład indeksu WIG30 oraz będących spółkami Skarbu Państwa. Ostatecznie, po usunięciu spółek Energia S.A. i Polski Holding Nieruchomości S.A. z powodu braków danych, w analizach uwzględniono 32 spółki, których pełna lista została zamieszczona w tablicy 1. Spółki podzielono na dwie grupy spółek: finansowe – 9 spółek – oraz niefinansowe – 23 spółki. Analizy przeprowadzono za okres 2010–2015.

Tablica 1. Lista spółek wytypowanych do badania i ich oznaczenia

Spółki niefinansowe – nazwa	Symbol	Spółki finansowe – nazwa	Symbol
Asseco Poland S.A.	N1	GPW S.A.	F1
Grupa Azoty S.A.	N2	Alior Bank S.A.	F2
CCC S.A.	N3	Bank Zachodni WBK S.A.	F3
CD Projekt S.A.	N4	ING Bank Śląski S.A.	F4
Cyfrowy Polsat S.A.	N5	mBank S.A.	F5
Enea S.A.	N6	Bank Millennium S.A.	F6
Eurocash S.A.	N7	Bank PKO S.A.	F7
Globe Trade Centre S.A.	N8	PKO Bank Polski S.A.	F8
JSW S.A.	N9	PZU S.A.	F9
Kernel Holding S.A.	N10		
KGHM Polska Miedź S.A.	N11		
LPP S.A.	N12		
Grupa Lotos S.A.	N13		
LW Bogdanka S.A.	N14		
Orange Polska S.A.	N15		
PGE S.A.	N16		
PGNiG S.A.	N17		
PKN Orlen S.A.	N18		
PKP Cargo S.A.	N19		
Synthos S.A.	N20		
Tauron Polska Energia S.A.	N21		
ZCh Police S.A.	N22		
EC Będzin S.A.	N23		

Źródło: Opracowanie własne.

W badaniach dynamiki analizowanych zjawisk wykorzystano: (1) w przypadku porównań międzyokresowych – przyrosty łańcuchowe oraz (2) w analizach dynamiki struktury zatrudnienia – przyrosty jednopodstawowe. Strukturę zatrudnienia badano wykorzystując wskaźniki struktury oraz tzw. indeks heterogeniczności Blaua (1977)[1]:

$$BI = 1 - \Sigma_{i=1}^{k} p_i^2 \tag{1}$$

gdzie: p_i – prawdopodobieństwo wystąpienia i-tego ($i = 1, 2, \ldots, k$) wariantu analizowanej cechy, mierzone wskaźnikiem struktury, wyznaczonym dla konkretnej i-tej kategorii:

[1] Indeks ten wykorzystali w swoich badaniach m.in. [Campbell, Minguez-Vera, 2008; Bohdanowicz, 2011b].

$$p_i = \frac{n_i}{n} \tag{2}$$

gdzie: n_i – liczba obserwacji charakteryzujących się i-tym wariantem cechy, n – liczba wszystkich obserwacji. W prowadzonym badaniu $k = 2$ i oznacza udziały kobiet i mężczyzn w organach kierowniczych.

Występowanie współzależności między strukturą organów kierowniczych i rentownością spółek, mierzoną wskaźnikiem ROE, analizowano, wykorzystując współczynnik korelacji liniowej Pearsona oraz test istotności tego współczynnika ze statystyką testową postaci:

$$t = \frac{r_{xy}}{\sqrt{\frac{1-r_{xy}^2}{n-2}}} \tag{3}$$

gdzie: x, y odzwierciedlają badanie zjawiska, tj. strukturę organów kierowniczych i ROE.

Badania korelacyjne uzupełniono analizą regresji liniowej, przeprowadzoną na podstawie prób przekrojowo-czasowych. Analizy obejmowały zarówno próby zawierające wszystkie spółki, jak i wyróżnione ich grupy, np. tylko spółki niefinansowe lub tylko finansowe, albo tylko spółki o określonej strukturze menedżmentu. W efekcie oszacowano MNK modele postaci:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} \tag{4a}$$

$$y_{it} = \alpha + \beta x_{it-1} + \varepsilon_{it} \tag{4b}$$

$$y_{it} = \alpha + \beta x_{it-2} + \varepsilon_{it} \tag{4c}$$

$$\Delta y_{it} = \alpha + \beta \Delta x_{it} + \varepsilon_{it} \tag{5a}$$

$$\Delta y_{it} = \alpha + \beta \Delta x_{it-1} + \varepsilon_{it} \tag{5b}$$

gdzie: y_{it}, Δy_{it} – zmienne zależne wyrażające poziom lub przyrost z okresu na okres wskaźnika finansowego (ROE), x_{it}, Δx_{it} – zmienne niezależne wyrażające poziom lub przyrost z okresu na okres udziału kobiet w organach kierowniczych spółek (w modelach (4a) udział mierzony wskaźnikami struktury lub indeksem Blaua), ε_{it} – składnik losowy, α, β – parametry równania regresji, $i = 1, 2, \ldots, n$ – oznacza spółkę ($n = 9$ dla spółek finansowych, $n = 23$ dla spółek niefinansowych i $n = 32$ dla wszystkich spółek itd.), $t = 1, 2, \ldots, T$ – oznacza rok analizy (maksymalnie $T = 6$), $N = nT$ – wielkość próby estymacyjnej.

Dodatkowo porównania średnich wyznaczonych dla różnie zdefiniowanych grup spółek przeprowadzono w oparciu o test równości średnich Cochrana-Coxa ze statystyką postaci:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \tag{6}$$

gdzie: $\bar{x}_1$, $\bar{x}_2$ – średnie, S_1^2, S_2^2 – wariancje, n_1, n_2 – liczebności porównywanych prób.

3. Analiza współzależności między wynikami finansowymi i strukturą organów kierowniczych

Badania korelacji i regresji liniowej przeprowadzono zarówno w odniesieniu do poziomów analizowanych mierników, jak i ich przyrostów, uwzględniając przy tym zależności współbieżne oraz opóźnione. U podstaw takiego podejścia leży spostrzeżenie, że jeśli obecność kobiet w organach kierowniczych i związane z tym decyzje zarządcze wpływają na wyniki spółek, to skutki te ujawniają się dopiero w kolejnych latach działalności. Zbadano również istotność wyznaczonych współczynników korelacji oraz parametrów modeli regresji, wskazując przypadki odrzucenia hipotezy zerowej.

W tablicy 2 zamieszczono wartości współczynników korelacji liniowej Pearsona wyznaczone w celu oceny siły i kierunku zależności między udziałem kobiet w organach zarządzających i wynikiem finansowym mierzonym wskaźnikiem ROE. Wnioskujemy, iż w całym 6-letnim okresie analizy zależności te są bardzo słabe i to niezależnie od charakteru spółek oraz od tego, czy były one mierzone współbieżnie (kolumna oznaczona jako 2010–2015), czy z przesunięciem w czasie, czyli czy struktura zatrudnienia była opóźniona o jeden rok w stosunku do pomiaru ROE (kolumna 2011–2015) lub o dwa lata (kolumna 2012–2015). W przypadku spółek niefinansowych w ogóle nie można mówić o występowaniu współzależności, ponieważ największa wartość współczynnika korelacji wynosi 0,3. Natomiast dla spółek finansowych odnotowano pojedyncze przypadki znacznej ujemnej korelacji między wartością wskaźnika ROE a udziałem kobiet w zarządach i radach nadzorczych – w latach 2012, 2013 i 2015 dla zależności mierzonych współbieżnie, a także w 2011 r. w zależności od udziału kobiet w organach zarządczych w 2010 r. oraz w 2015 r. w relacji do zatrudnienia z lat 2014 i 2013.

Wartości współczynników korelacji wyznaczone dla przyrostów bezwzględnych zestawiono w tablicy 3. Współczynniki Pearsona mają wartości niewielkie i są w większości ujemne. Tylko dla spółek finansowych przyrost udziału kobiet w organach kierowniczych w 2012 r. dodatnio wpływa, z maksymalną zaobserwowaną wartością 0,428, na wzrost wskaźnika rentowności kapitału własnego w następnym roku. Jednocześnie relacja

zaobserwowana w kolejnym roku jest przeciwna i generuje najmniejszą wartość współczynnika, tj. – 0,334.

Tablica 2. Wartości współczynników korelacji Pearsona wyznaczone dla poziomów zmiennych

Lata / Spółki	2010	2011	2012	2013	2014	2015	2010–2015
zależności bieżące							
niefinansowe	0,0443	–0,0487	–0,0342	–0,0993	–0,1262	–0,3024	–0,1285
finansowe	–0,0015	0,0469	**–0,6573**	**–0,6994**	–0,3848	**–0,7792**	–0,3474
zależności opóźnione o jeden rok							2011–2015
niefinansowe		–0,0140	–0,0389	–0,0589	–0,0765	–0,0263	–0,0467
finansowe		**0,6495**	–0,5490	–0,5305	–0,4948	**–0,6323**	–0,2921
zależności opóźnione o dwa lata							2012–2015
niefinansowe			0,0396	–0,0812	0,0254	–0,0514	–0,0281
finansowe			0,1662	–0,3047	–0,4133	**–0,6423**	–0,2087

Czcionka pogrubiona oznacza odrzucenie hipotezy zerowej na poziomie $\alpha = 0{,}05$.
Źródło: Obliczenia własne.

Tablica 3. Wartości współczynników korelacji Pearsona wyznaczone dla przyrostów bezwzględnych zmiennych

Lata / Spółki	2011	2012	2013	2014	2015	
	zależności bieżące					2010–2015
niefinansowe	0,0249	–0,1146	0,2227	–0,2373	–0,2408	–0,11636
finansowe	0,1411	–0,2379	–0,1050	0,0112	0,2586	–0,02303
wszystkie	0,0308	–0,1286	0,2466	–0,1583	–0,2203	–0,09506
	zależności opóźnione o 1 rok					2011–2015
niefinansowe		–0,0435	0,0353	–0,2012	0,0963	–0,1164
finansowe		–0,2001	–0,0430	0,4280	–0,3342	–0,0230
wszystkie		–0,0678	–0,0418	–0,1241	0,0510	–0,0951

Źródło: Obliczenia własne.

Z analiz modeli (4) i (5), oszacowanych na podstawie prób zawierających wszystkie spółki oraz z podziałem na spółki finansowe i niefinansowe, wynika, że statycznie istotne zależności zaobserwowano jedynie dla spółek finansowych. W tablicy 4 zestawiono jedynie te wersje modeli, w których relacja między ROE a strukturą zatrudnienia (mierzona wskaźnikiem struktury) była statystycznie istotna. Dotyczy to jedynie spółek finansowych i relacji (4a), (4b). Wzrost udziału kobiet w organach władzy spółek finansowych zarówno w bieżącym, jak i poprzednim roku istotnie wpływa na obniżenie się wartości współczynnika rentowności kapitału własnego.

Należy przy tym zauważyć, że współczynniki determinacji we wszystkich modelach są bardzo niskie (od 0,037 do 0,371), co potwierdza, że struktura zatrudnienia ma jedynie bardzo niewielki wpływ na kształtowanie się ROE. Najsilniejsza relacja między obiema zmiennymi zachodzi dla okresu czteroletniego 2012–2015 (współczynnik Pearsona wynosi – 0,609).

Uwzględniając zaobserwowaną prawidłowość zwiększonej korelacji między wskaźnikiem ROE a udziałem kobiet we władzach spółek finansowych dla skróconego okresu estymacji, oszacowano modele regresji na próbach czasowych o zróżnicowanej długości. Jak widać (tab. 4), siła korelacji między zmiennymi zwiększa się wraz ze skracaniem próby dla stałego jej końca w roku 2015 i maleje przy stałym początku próby w roku 2010.

Tablica 4. Parametry modeli regresji (4a) i (4b), oszacowanych dla spółek finansowych

Model	(4a)							
Próba	2010–2015		2011–2015		2012–2015		2013–2015	
Parametr	beta	alfa	beta	alfa	beta	alfa	beta	alfa
wartość	**–0,188**	**0,166**	**–0,238**	**0,177**	**–0,263**	**0,176**	**–0,227**	**0,170**
statystyki *t*	–2,672	13,255	–4,021	16,808	–4,482	16,291	–3,718	14,393
R^2	0,121	$N = 54$	0,273	$N = 45$	0,371	$N = 36$	0,356	$N = 27$
Model	(4b)		(4a)					
Próba	2011–2015		2010–2014		2010–2013		2010–2012	
Wartości	beta	alfa	beta	alfa	beta	alfa	beta	alfa
parametrów	**–0,133**	**0,160**	**–0,168**	**0,166**	**–0,181**	**0,168**	–0,134	**0,160**
statystyki *t*	–2,002	14,005	–1,985	11,382	–1,690	9,594	–0,985	7,354
R^2	0,0853	$N = 45$	0,0840	$N = 45$	0,0775	$N = 36$	0,0374	$N = 27$

N oznacza liczebność próby, a pogrubienie – odrzucenie hipotezy zerowej przy $\alpha = 0{,}05$.
Źródło: Obliczenia własne.

Przyjęto założenie, że wobec niewielkiej liczby kobiet pełniących w spółkach funkcje prezesa zarządu lub przewodniczącego rady nadzorczej decyzje wpływające na funkcjonowanie spółek podejmowane są kolegialnie, a udział kobiet w ich kreowaniu jest ściśle związany ze strukturą organów kierowniczych. Konsekwencją tego założenia jest stwierdzenie, że nie można oczekiwać wpływu kobiet na decyzje kierownicze, jeśli nie są one dostatecznie reprezentowane w organach władzy. W kolejnych analizach usunięto zatem z próby wszystkie te spółki, w których udział kobiet łącznie w radach nadzorczych i zarządach przynajmniej w jednym roku nie przekroczył 15%. Tak zbudowana próba zawiera 17 spółek (łącznie finansowych i niefinansowych).

Oceny estymatorów parametrów modeli (4) szacowanych na poziomach zmiennych zawiera tablica 5. Stwierdzono istotne statystycznie ujemne wartości parametru beta we wszystkich modelach (4a). Modele (5) szacowane na podstawie przyrostów nie wykazywały istotnej współzależności między badanymi zmiennymi.

Tablica 5. Parametry modeli regresji (4a) i (4b), oszacowanych dla spółek, w których przynajmniej w jednym roku został przekroczony parytet 15% kobiet w organach władzy

Modele	(4a)							
Próba	2010–2015		2011–2015		2012–2015		2013–2015	
Wartości	beta	alfa	beta	alfa	beta	alfa	beta	alfa
parametrów	**–0,2131**	**0,1439**	**–0,3046**	**0,1553**	**–0,2159**	**0,1196**	**–0,2207**	**0,1071**
statystyki *t*	–1,9240	6,5886	–2,4605	6,2982	–2,0347	5,4639	–1,8187	4,1997
R^2	0,0357	*N*=102	0,0680	*N*=85	0,0590	*N*=68	0,0632	*N*=51
Modele	(4b)		(4a)					
Próba	2011–2015		2010–2014		2010–2013		2010–2012	
Wartości	beta	alfa	beta	alfa	beta	alfa	beta	alfa
parametrów	–0,2012	**0,1352**	**–0,1610**	**0,1484**	**–0,0939**	**0,1480**	**–0,0782**	**0,1601**
statystyki *t*	–1,5752	5,5264	–1,3091	6,2966	–0,6325	5,3244	–0,4420	4,9389
R^2	0,0290	*N*=85	0,0202	*N*=85	0,0060	*N*=68	0,0040	*N*=51

N oznacza liczebność próby, a pogrubienie – odrzucenie hipotezy zerowej przy $\alpha = 0{,}05$.
Źródło: Obliczenia własne.

W badaniach wpływu kobiet na wyniki spółek przewija się pojęcie tzw. masy krytycznej, tj. granicznego parytetu kobiet w gremiach zarządczych, po osiągnięciu którego ich wpływ na standing przedsiębiorstwa jest istotny. W analizowanym zbiorze spółek wyszukano zatem takie przypadki, w których udział kobiet w organach kierowniczych wynosi przynajmniej 25%. Z oszacowań modeli (4a), (4b) dla takiej próby wynika (tab. 6), że wpływ kobiet na poziom ROE jest nieistotny statystycznie, przy korelacji 0,117 (model 4a) i 0,325 (model 4b).

Ponadto sprawdzono (zgodnie z metodologią Catalyst 2011) średnią wartość współczynnika ROA dla wyróżnionych 27 przypadków, w których udział kobiet we władzach spółek przekraczał 25%. Następnie porównano tę uśrednioną wartość ze średnią ROE, wyznaczoną dla przypadków, w których udział kobiet w gremiach kierowniczych był zerowy (37 przypadków). W tablicy 7 zestawiono wartości średnie wyznaczone dla wszystkich przypadków, a także wyznaczone z pominięciem dwóch spółek, w których w całym okresie badawczym nie zatrudniano żadnej kobiety ani w zarządzie, ani w radzie nadzorczej. Jak można zauważyć,

we wszystkich analizowanych przypadkach spółek z przynajmniej 25% reprezentacją kobiet średnia wartość ROE jest niższa niż dla spółek bez kobiet w organach kierowniczych. Różnice w średnich są przy tym statystycznie istotne tylko jeśli porównać ze sobą wszystkie przypadki.

Tablica 6. Parametry modeli regresji (4a) i (4b), oszacowanych dla przypadków zachowania 25% parytetu

Modele	(4a)		(4b)	
Próba	2010–2015		2011–2015	
Wartości	beta	alfa	beta	alfa
parametrów	0,2626	0,0074	0,6023	–0,0875
statystyki *t*	0,5884	0,0570	1,4584	–0,7452
R^2	0,0137	*N*=27	0,1057	*N*=20

N oznacza liczebność próby.
Źródło: Obliczenia własne.

Tablica 7. Uśredniona wartość ROE dla przypadków zachowania 25% parytetu oraz zerowego udziału kobiet we władzach spółek

Udział kobiet we władzach	25% kobiet		0% kobiet		Statystyki Cochrana–Coxa
Opis próby / Struktura próby	ROE	*N*	ROE	*N*	
Wszystkie przypadki	8,22%	27	15,50%	33	**–2,0950**
z pominięciem LPP i Synthos	8,22%	27	9,99%	21	–0,4138
Tylko spółki niefinansowe	7,63%	17	16,50%	28	–1,5628
z pominięciem LPP i Synthos	7,63%	17	6,68%	16	0,1753

N oznacza liczebność próby, a pogrubienie – odrzucenie hipotezy zerowej przy $\alpha = 0{,}05$.
Źródło: Obliczenia własne.

Podobne wnioski należy wyciągnąć z analiz porównawczych przedstawionych w tablicy 8, gdzie zestawiono średnie wartości wskaźnika ROE wyznaczone dla sześciu lat i wszystkich analizowanych spółek z podziałem na dwie grupy, biorąc pod uwagę udział kobiet w organach kierowniczych i przyjmując progowe wartości na poziomie 15% i 25%. Jak widać, uśredniona rentowność kapitału własnego jest mniejsza w spółkach o niższym udziale kobiet w menedżmencie, a statystycznie istotna różnica występuje, jeśli porównuje się spółki z 25% udziałem kobiet w ich kierownictwie.

Wszystkie powyższe spostrzeżenia potwierdziło wykorzystanie indeksu Blaua do pomiaru stopnia dywersyfikacji organów kierowniczych badanych spółek – w większości przypadków relacje między reprezentacją kobiet i wynikami finansowymi były ujemne. Analizując siłę korelacji, zauważamy wysokie wartości bezwzględne współczynnika Pearsona (tab. 9) dla spółek finansowych w latach 2012, 2013 i 2015, kiedy mierzone

są zależności współbieżne, a dla lat 2011 i 2015 dla zależności opóźnionych o jeden rok, które w jedynym 2011 r. są dodatnie.

Tablica 8. Średnie wartości ROE wyznaczone na podstawie wszystkich spółek i lat 2010–2015

Opis próby / Spółki, w których udział kobiet w kierownictwie:	ROE	*N*	Statystyki Cochrana-Coxa
choć w jednym roku wynosił przynajmniej 15%	10,91%	19	
we wszystkich latach był poniżej 15%	13,63%	13	–1,0415
choć w jednym roku wynosił przynajmniej 25%	9,63%	12	
we wszystkich latach był poniżej 25%	13,36%	20	**–1,8886**

N oznacza liczebność próby, a pogrubienie – odrzucenie hipotezy zerowej przy $\alpha = 0{,}05$.
Źródło: Obliczenia własne.

Tablica 9. Wartości współczynników korelacji Pearsona wyznaczone dla indeksu Blaua i poziomów ROE

	zależności bieżące						
Spółki	2010	2011	2012	2013	2014	2015	2010–2015
niefinansowe	–0,0166	–0,0534	–0,0111	–0,1254	–0,1206	–0,3081	–0,1315
finansowe	–0,0663	0,0627	**–0,6563**	**–0,7099**	–0,4224	**–0,8756**	–0,3809
wszystkie	–0,0486	–0,0459	–0,1224	–0,1353	–0,0501	–0,2641	–0,1316
	zależności opóźnione o 1 rok						
Spółki		2011	2012	2013	2014	2015	2011–2015
niefinansowe		–0,0572	–0,0387	–0,0359	–0,0878	0,0113	–0,0407
finansowe		**0,5407**	–0,5333	–0,5563	–0,4847	**–0,7072**	–0,3354
wszystkie		–0,0101	–0,1107	–0,0906	–0,0632	0,0177	–0,0424

N oznacza liczebność próby, a pogrubienie – odrzucenie hipotezy zerowej przy $\alpha = 0{,}05$.
Źródło: Obliczenia własne.

Oszacowane funkcje regresji (tab. 10) wskazują występowanie statystycznie istotnych ujemnych zależności między poziomem ROA i strukturą organów zarządczych jedynie dla spółek finansowych, a także w przypadku wszystkich spółek w modelu (4a). W pozostałych modelach nie stwierdzono występowania wpływu reprezentacji kobiet na rentowność kapitału własnego.

Tablica 10. Parametry modeli regresji (4a) i (4b), oszacowanych dla indeksu Blaua, mierzącego strukturę organów kierowniczych

Modele: próba		(4a): 2010–2015		(4b): 2011–2015	
Parametry		beta	alfa	beta	alfa
Spółki niefinansowe					
Wartości	parametrów	–0,1868	**0,1474**	–0,0602	**0,1127**
	statystyki *t*	–1,5464	5,4462	–0,4326	3,6712
	R^2	0, 1073	$N = 138$	0,0017	$N = 115$
Spółki finansowe					
Wartości	parametrów	**–0,1547**	**0,1738**	**–0,1129**	**0,1667**
	statystyki *t*	–2,9703	12,5010	–2,3346	13,2257
	R^2	0,1451	$N = 54$	0,1125	$N = 45$
Wszystkie spółki					
Wartości	parametrów	**–0,1583**	**0,1513**	–0,0528	**0,1231**
	statystyki *t*	–1,8293	7,3788	–0,5334	5,3453
	R^2	0,0173	$N = 192$	0,0018	$N = 160$

N oznacza liczebność próby, a pogrubienie – odrzucenie hipotezy zerowej przy $\alpha = 0{,}05$.
Źródło: Obliczenia własne.

Zakończenie

Z przedstawionych analiz wynika, że niezależnie od udziału kobiet zajmujących kierownicze pozycje w spółkach ich wpływ na wyniki finansowe mierzone wskaźnikiem ROE jest nieistotny albo wręcz negatywny. Co więcej, w przeciwieństwie do badań raportowanych przez Catalyst, a realizowanych w krajach wysoko rozwiniętych, spółki o zwiększonym udziale kobiet we władzach charakteryzują się niższą rentownością kapitału własnego niż te, w których kobiet brakuje (tab. 6 i 7), chociaż nie wszystkie różnice są statystycznie istotne.

Trudno jest dokonać jednoznacznej oceny uzyskanych wyników, a tym bardziej stwierdzić, dlaczego przeczą one obserwacjom poczynionym przez uznawane ośrodki monitorujące to zagadnienie za granicą. Należy jednak przypomnieć, że rezultaty podobne do zaprezentowanych były już opisywane przez innych badaczy. W tym kontekście zasadne wydaje się postrzeganie wpływu kobiet w radach nadzorczych i zarządach także przez pryzmat sposobu doboru członków tych gremiów w Polsce, zwłaszcza w spółkach Skarbu Państwa, w którym kompetencje kandydatów ustępują regułom klucza politycznego.

Literatura

Adams R.B., Ferreira D. (2009), *Women in the boardroom and their impact on governance and performance*, „Journal of Financial Economics", Vol. 94, No. 2.

Adams S.M., Gupta A., Leeth J.D. (2009), *Are female executives over-represented in precarious leadership positions?* „British Journal of Management", Vol. 20, No. 1.

Ahern K.R., Dittmar A.K. (2012), *The change of the boards: The impact on Firm valuation of mandated female board representation*, „The Quarterly Journal of Economics", Vol. 127, No. 1.

Bear S., Rahman N., Post C. (2010), *The impact of board diversity and gender composition on corporate social responsibility and firm reputation*, „Journal of Business Ethics", Vol. 97, No. 2.

Blau P.M. (1977), *Inequality and Heterogeneity*, Free Press, New York.

Bohdanowicz L. (2010), *Kobiety w radach nadzorczych spółek publicznych – pokonywanie barier*, „Przegląd Corporate Governance", nr 3 (23).

Bohdanowicz L. (2011a), *Kobiety w radach nadzorczych i zarządach spółek: Polskie i światowe tendencje oraz wyzwania*, „Organizacja i Kierowanie", nr 3.

Bohdanowicz L. (2011b), *Zróżnicowanie organów statutowych pod wzg-lędem płci a wyniki finansowe polskich spółek publicznych*, „Acta Universitatis Nicolai Copernici", Ekonomia XLII, nr 403.

Brown D.A.H., Brown D.L., Anastasopoulos V. (2002), *Women on Boards: Not Just the Right Thing...But the "Bright" Thing*, The Conference Board of Canada, Report, http://utsc.utoronto.ca/~phanira/WebResearchMethods/women-bod&fp--conference%20board.pdf.

Campbell K., Minguez-Vera A. (2008), *Gender Diversity in the Boardroom and Firm Financial Performance*, „Journal of Business Ethics", Vol. 83, No. 3.

Carter D.A., D'Souza F., Simkins B.J., Simpson W.G. (2010), *The Gender and Ethnic Diversity of US Boards and Board Committees and Firm Financial Performance*, „Corporate Governance: An International Review", Vol. 18, No. 5.

Carter D.A, Simkins B.J., Simpson W.A. (2003), *Corporate Governance, Board Diversity, and Firm Value*, „The Financial Review", Vol. 38, No. 1.

Carter N.M., Wagner H.M. (2011), *The bottom line: Corporate performance and women's representation on boards (2004-2008)*, Catalyst, 2011, http://www.catalyst.org/system/files/the_bottom_line_corporate_performance_and_women%27s_representation_on_boards_%282004-2008%29.pdf, dostęp: 24.03.2017.

Catalyst (2004), *The Bottom Line: Connecting Corporate Performance and Gender Diversity*, Catalyst, New York.

Curtis M., Schmid Ch., Struber M. (2012), *Gender diversity and corporate performance*, Credit Suisse AG Research Institute, Zurich, http://www.calstrs.com/sites/main/files/fileattachments/csri_gender_diversity_and_corporate_performance.pdf, dostęp: 10.07.2016.

Desvaux G., Devillard-Hollinger S., Baumgarten P. (2007), *Women matter: Gender diversity a corporate performance driver*, McKinsey& Company, http://www.raeng.org.uk/publications/other/women-matter-oct-2007, dostęp: 13.07.2016.

Devillard S., Graven W., Lawson E., Paradise R., Sancier-Sultan S. (2012), *Women Matter 2012, Making the breakthrough*, McKinsey&Company, http://www.mckinsey.com/ features/women_matter.

Farrel K.A., Hersch P.L. (2005), *Additions to corporate boards: The effect of gender*, „Journal of Corporate Finance", Vol. 11, No. 1–2.

Julizaerma M.K., Sori Z.M. (2012), *Gender diversity in the boardroom and firm performance of Malaysian Public Listed Companies*, „Procedia Social and Behavioral Sciences", Vol. 65.

Kompa K., Mentel G., Witkowska D. (2016), *Czy obecność kobiet we władzach spółek giełdowych wpływa na poprawę sytuacji finansowej tych spółek*, „Metody Ilościowe w Badaniach Ekonomicznych", nr 17(3), http://qme.sggw.pl/wp-content/uploads/MIBE_T17_z3.pdf.

Lee P.M., James E.H.J. (2007), *She'-e-os: gender effects and investor reactions to the announcements of top executive appointments*, „Strategic Management Journal", Vol. 28, No. 3.

Lenard M.J., Yu B., York E.A., Wu S. (2014), *Impact of board gender diversity on firm risk*, „Managerial Finance", Vol. 40, No. 8.

Lewiatan (2012), *Przedsiębiorcy w Polsce. Fakty, liczby, Przykłady*, Raport PKPP Lewiatan, Warszawa.

Lisowska E., Zachorowska M., Sznajder A., Grabowska M. (2014), *Więcej kobiet w zarządzaniu – to się opłaca*, Seria: Równość w biznesie, Raport wykonany na zlecenie MPiPS, Warszawa.

McKinsey (2007), *Women Matter: Gender diversity, a corporate performance driver*, McKinsey & Company, Inc.

Oplustil K. (2011), *Ekspertyza na temat zielonej księgi Komisji Europejskiej pt. „Unijne ramy ładu korporacyjnego"*, Opinie i Ekspertyzy OE-173, Biuro Analiz i Dokumentacji, Kancelaria Senatu, Warszawa, http://ww2.senat.pl/k7/dok/opinia/2011/oe-173.pdf.

Radwan A. (2010), *Corporate governance w spódnicy – parytet równości również na parkiecie?* „Przegląd Corporate Governance", nr 3(23).

Rose C. (2007), *Does female board representation in uence rm performance? The Danish evidence*, „Corporate Governance: An International Review", Vol. 15, No. 2.

Ryan M.K., Haslam S.A. (2007), *Glass cliffs: Exploring the dynamics surrounding the appointment of women to precarious leadership positions*, „Academy of Management Review", Vol. 32, No. 2.

Shrader C., Blackburn V., Iles P. (1997), *Women in management and rm value: An exploratory study*, „Journal of Managerial Issues", Vol. 9.

Smith N., Smith V., Verner M. (2006), *Do women in top management affect firm performance? A panel study of 2,500 Danish firms*, „International Journal of Productivity and Performance Management", Vol. 55, No. 7.

Walat T. (2010), *Skandynawskie parytety w biznesie. Złote spódniczki*, „Polityka" 25.02.2010, http://www.polityka.pl/tygodnikpolityka/rynek/1503524,1,skandynawskie-parytety-w-biznesie.read.

Wang Y., Clift B. (2009), *Is There A Business Case For Board Diversity?*, „Pacific Accounting Review", Vol. 21, No. 2.

Zahra S.A., Stanton W.W. (1988), *The Implications of Board of Directors' Composition for Corporate Strategy and Performance*, „International Journal of Management", Vol. 5, No. 2.

Streszczenie

Komisja Europejska postuluje wprowadzenie 40% parytetu udziału kobiet w zarządach i radach nadzorczych dużych spółek publicznych i spółek Skarbu Państwa. Tymczasem w Polsce udział kobiet w organach kierowniczych takich spółek znacząco odbiega od postulowanego poziomu parytetu i prowokuje dyskusję o potrzebie wprowadzania takich regulacji prawnych. Zwolennicy wprowadzenia parytetu wskazują m.in. na istotną poprawę wyników finansowych firm, w kierownictwie których zatrudniono kobiety. Kontrowersje wokół tego stwierdzenia legły u podstaw prezentowanej pracy.

Celem pracy jest zbadanie, czy w przypadku spółek notowanych na GPW w Warszawie zachodzi relacja pomiędzy strukturą zatrudnienia kobiet w organach zarządczych a sytuacją finansową przedsiębiorstwa. Badania przeprowadzono za lata 2010–2015 na podstawie danych z Notoria Serwis. Analizy objęły 32 spółki z indeksu WIG30 i spółki z udziałem Skarbu Państwa.

Artykuł zawiera przegląd literatury przedmiotu, opis zastosowanych metod badawczych i część empiryczną z wynikami badania współzależności między standingiem finansowym analizowanych spółek a udziałem kobiet w menedżmencie.

Słowa kluczowe

spółki publiczne, rentowność kapitału własnego, obecność kobiet w kierownictwie

Does Presence of Women in Management Influences Return on Equity of Companies Listed on WSE? (Summary)

In many states, the gender parity assuring 30–40% women in boards of public and state companies, being classified as big ones, has been already affected. In Poland, the presentence of women in boards is smaller than the level requested by European Commission, and there are wide discussions on the necessity of such regulation. The supporters present advantages, among which the main argument is that performance of companies with the increasing number of female managers improved.

Therefore, the aim of our research is to examine if we can find similar relations for the companies listed on Warsaw Stock Exchange (WSE) in the years 2010–2015. Investigation was provided for 32 public companies which belonged to the stock index WIG 30 or were the State Treasury companies, using data from Notoria Serwis.

The paper consists of five sections. The first section contains literature review. The second one describes methods and data. Two following sections discuss the measurement and analyze investigated phenomena, i.e. share of women in management and return on equity (ROE) in considered companies. In the last section, we present the results of our research on relationship between ROE and the share of women in boardrooms.

Keywords

public companies, return on equity, women presence in boardrooms

Joanna Olbryś*
Michał Mursztyn**

Elastyczność rynku jako jeden z wymiarów płynności Giełdy Papierów Wartościowych w Warszawie S.A.

Wstęp

Analiza empiryczna wymiarów płynności jako dodatkowych charakterystyk rynku giełdowego może stanowić uzupełnienie badań dotyczących płynności/niepłynności aktywów kapitałowych. Problemem jest jednak brak formalnych definicji wymiarów płynności, który utrudnia analizy empiryczne oraz sprawia, że w literaturze przedstawionych jest wiele propozycji, jak estymować te charakterystyki rynku giełdowego. Niektórzy autorzy zwracają też uwagę na fakt, że nieformalne definicje wymiarów płynności są oparte głównie na intuicjach oraz ogólnych, trudno definiowalnych pojęciach, takich jak normalne warunki (np. rynkowe) czy normalny poziom (np. cen) [Doman, 2011]. Poza tym należy podkreślić, że kluczowe znaczenie mają również struktura rynku kapitałowego i organizacja handlu, jako czynniki wywierające istotny wpływ na metody pomiaru płynności [Harris, 2003].

Temat wymiarów płynności rynku (*dimensions of market liquidity*) prawdopodobnie pojawił się po raz pierwszy w fundamentalnej pracy Kyle'a z 1985 r., dotyczącej mikrostruktury rynku. Autor wymienił pewne charakterystyki wspomagające ocenę płynności, takie jak: naprężenie rynku (*market tightness*), głębokość rynku (*market depth*) oraz elastyczność rynku (*market resiliency*) [Kyle, 1985, s. 1316]. Analogiczny podział wymiarów płynności podają również Wong i Fung [2002] oraz Doman [2011]. Niektórzy autorzy wymieniają czas (*trading time*) jako dodatkowy wymiar płynności, np. [Ranaldo, 2001, s. 311–312; von Wyss, 2004, s. 5–6]. Należy jednak zauważyć, że czas w sensie natychmiastowości transakcji jest powiązany z pozostałymi trzema wymiarami, więc nie musi być rozpatrywany oddzielnie.

Wymiary płynności rynku są zwykle szacowane dla pojedynczych papierów wartościowych. Naprężenie rynku odnosi się do kosztów odwrócenia zajętej pozycji w krótkim czasie. Głębokość rynku informuje o zachowaniu

* Prof. PB, dr hab., Katedra Informatyki Teoretycznej, Wydział Informatyki, Politechnika Białostocka, Wiejska 45A, Białystok 15-351, j.olbrys@pb.edu.pl
** Student, Wydział Informatyki, Politechnika Białostocka, Wiejska 45A, Białystok 15-351, m.mursztyn@student.pb.edu.pl

ciągłości kupna i sprzedaży instrumentów finansowych. Natomiast pojęcie elastyczności rynku odnosi się do szybkości, z jaką cena waloru powraca do ceny równowagi, po losowym zaburzeniu. W literaturze występuje również pojęcie szerokości rynku, jednak nie ma ono jednoznacznego określenia, nawet w języku angielskim. Spotykane nazwy to: *market width* [Lin i inni, 2012] lub *market breadth*, np. [Bernstein, 1987]. Pomimo swojej nazwy szerokość rynku nie jest traktowana jako wymiar płynności, lecz informuje o przewadze kupujących nad sprzedającymi. Jest to wskaźnik liczony najczęściej jako stosunek liczby spółek, których ceny zwyżkowały, do liczby spółek, których ceny zniżkowały [Jankowski, Olbryś, 2015, s. 649].

Należy dodać, że wymiary płynności mogą być rozumiane też w inny sposób. Na przykład Foucault i inni w monografii z 2013 r. analizują trzy wymiary płynności (*dimensions of liquidity*): płynność rynku (*market liquidity*), płynność finansowania (*funding liquidity*) oraz płynność pieniężną (*monetary liquidity*) [Foucault i inni, 2013].

Większość badań empirycznych dotyczących wymiarów płynności rynków kapitałowych na świecie pochodzi z ostatnich kilkunastu lat. Temat ten jest jednak stosunkowo rzadko poruszany przez polskich autorów. Monografie Domana [2011] oraz Olbrysia [2014] przedstawiają ten temat od strony teoretycznej, natomiast pojedyncze prace zawierają również badania empiryczne wybranych wymiarów płynności na giełdzie warszawskiej. Prace dotyczące głębokości rynku to np. [Będowska-Sójka, 2014; Olbryś, 2017; Olbryś, Mursztyn, 2016; 2017b]. Z kolei temat naprężenia rynku jest analizowany empirycznie np. w artykułach [Olbryś, 2017; Olbryś, Mursztyn, 2017a; 2017b]. Nie są natomiast znane autorom żadne opracowania prezentujące wyniki badań elastyczności polskiego rynku giełdowego. Jest to najtrudniejszy do oszacowania wymiar płynności. W tym kontekście celem artykułu była analiza empiryczna elastyczności jako jednego z wymiarów płynności rynku na Giełdzie Papierów Wartościowych w Warszawie S.A. Badanie objęło 53 firmy z trzech grup spółek – dużych, średnich i małych – w okresie od stycznia 2005 r. do czerwca 2015 r. Podział na grupy umożliwił zaobserwowanie wpływu efektu wielkości spółki na wyniki badań.

Przeprowadzono również analizę otrzymanych wyników pod kątem wyboru okresu badania, z uwzględnieniem całej próby statystycznej oraz trzech jednakowo licznych podokresów: przed kryzysem, kryzys, po kryzysie. Okres ostatniego kryzysu finansowego na giełdzie warszawskiej został ustalony w sposób formalny jako przedział czasu od początku czerwca 2007 do końca lutego 2009 r., na podstawie wyników przedstawionych w pracy Olbryś i Majewskiej [2015].

Ponadto dokonano podstawowej analizy zależności korelacyjnych pomiędzy alternatywnymi estymatorami elastyczności rynku i zweryfikowano hipotezę badawczą zakładającą stabilność tych zależności w czasie. W tym celu testowano równość macierzy korelacji w podokresach, z wykorzystaniem modyfikacji testu Larntza–Perlmana [1985], zaproponowanej w artykule [Nowak, Olbryś, 2015]. Wyniki empiryczne wskazały konieczność odrzucenia hipotezy zerowej o stabilności zależności korelacyjnych pomiędzy estymatorami elastyczności w podokresach: (kryzys, przed kryzysem) oraz (kryzys, po kryzysie).

Praca ma następującą strukturę. W części pierwszej przedstawiono przegląd literatury dotyczącej elastyczności rynku na tle badań innych wymiarów płynności. Część druga prezentuje metody i wyniki analiz empirycznych na giełdzie warszawskiej oraz dyskusję. Artykuł zakończono podsumowaniem otrzymanych wyników.

Opracowanie powstało w ramach realizacji projektu badawczego 2016/21/B/HS4/02004 „Badania komparatywne wspólności w płynności na rynkach giełdowych Europy Środkowo-Wschodniej", finansowanego przez Narodowe Centrum Nauki.

1. Elastyczność jako jeden z wymiarów płynności rynku

Jak podkreślono we wstępie, nie ma jednomyślności w literaturze w kwestii metod estymacji wymiarów płynności rynku. W szczególności problem dotyczy pomiaru elastyczności. Wymiar ten jest miarą elastyczności ceny papieru wartościowego, zatem trudno zdefiniować odpowiedni indykator. Tym niemniej niezależnie od proponowanego przez różnych autorów podejścia wszystkie procedury wymagają zastosowania danych śródziennych o ultrawysokiej częstotliwości.

Według definicji podanej przez Kyle'a [1985, s. 1331] elastyczność rynku jest miarą prędkości, z jaką cena waloru powraca do ceny równowagi, po losowym, nieinformacyjnym zaburzeniu. Na rynku o wysokiej płynności cena powinna natychmiast powracać do poziomu efektywnego [Engle, Lange, 1997, s. 3]. Dlatego też niektórzy badacze argumentują, że naturalną i zgodną z intuicją aproksymację wymiaru elastyczności można uzyskać z wykorzystaniem estymatora szybkości powrotu ceny do poprzedniego poziomu (*price reversal proxy*) [Alan i inni, 2015].

Glosten [1987] wykazał, że można dokonać dekompozycji rozpiętości cen sprzedaż/kupno (*bid/ask spread*) na dwa komponenty: pierwszy związany z asymetrią informacji oraz drugi, uwzględniający pozostałe czynniki wpływające na wielkość rozpiętości cen. W nawiązaniu do wyników Glostena Hasbrouck i Seppi [2001] wyróżnili dwa komponenty spreadu: trwały (*permanent*) oraz tymczasowy (*transitory*). Huang i Stoll [1996, 1997]

zaproponowali procedurę dekompozycji efektywnej rozpiętości cen sprzedaż/kupno (*effective bid/ask spread*) na dwa komponenty: informacyjny (trwały) i tymczasowy, reprezentujący szybkość powrotu ceny papieru wartościowego do poprzedniego poziomu po losowym, nieinformacyjnym zaburzeniu. Autorzy zastosowali zrealizowaną rozpiętość cen sprzedaż/kupno (*realized bid/ask spread*) jako miarę „chwilowego" komponentu efektywnego spreadu. Zgodnie z literaturą zrealizowana rozpiętość cen sprzedaż/kupno, jako tymczasowy komponent efektywnego spreadu, może być wykorzystana do aproksymacji elastyczności rynku.

Podobne podejście zaproponował Ranaldo [2001], który aproksymował elastyczność rynku poszczególnych spółek na giełdzie w Zurychu między innymi za pomocą śródziennych stóp zwrotu z bardzo krótkich (np. 10-minutowych) przedziałów czasu. Ponadto Ranaldo szacował elastyczność, stosując również alternatywne wersje współczynnika płynności (*liquidity ratio*) oraz współczynnik wariancji (*variance ratio*). Do pracy Ranaldo nawiązał von Wyss w doktoracie z 2004 r., w którym przedstawił wyczerpujący przegląd miar płynności/niepłynności, stosowanych również w kontekście estymacji wymiarów płynności rynku. Podał inną, trzecią wersję *liquidity ratio* jako średniej dziennej stopy zwrotu ze wszystkich transakcji w danym dniu [von Wyss, 2004, s. 19].

Inną, stosunkowo często stosowaną metodą jest analiza elastyczności rynku wykorzystująca funkcje odpowiedzi na impuls (*impulse responce function*), estymowane na podstawie różnych wersji modeli typu VAR, np. [Coppejans i inni, 2004; Hmaied i inni, 2006; Lo, Hall, 2015].

Wśród oryginalnych pomysłów estymacji elastyczności rynku można wymienić propozycję Muranagi i Shimizu [Muranaga, Shimizu, 1999; Muranaga, 2000], którzy wykorzystali do oszacowania szybkości konwergencji rozpiętości cen sprzedaż/kupno, przeniesioną bezpośrednio z zastosowań fizycznych miarę szybkości powrotu odkształconej spirali do pierwotnej postaci.

Z kolei autorzy pracy [Bhattacharya i inni, 2016] zaproponowali współczynnik efektywności rynku (*market efficiency coefficient* – MEC), jako estymator elastyczności, jednak do oszacowania jego wartości wykorzystali dane dzienne, a nie śródzienne.

Nie wszystkie prace dotyczące elastyczności rynku zawierają badania empiryczne. Na przykład Foucault i inni [2005] oraz Large [2007] przedstawili propozycje pewnych modeli teoretycznych oraz odpowiednie przykłady numeryczne. Do pracy [Foucault i inni, 2005] nawiązali Kempf, Mayston i Yadav [2009], którzy przeprowadzili badania empiryczne elastyczności rynku na giełdzie we Frankfurcie.

Degryse i inni [2003] analizowali elastyczność rynku 20 spółek z giełdy w Paryżu z punktu widzenia analizy zdarzeń (*event study*), w kontekście występowania tzw. agresywnych zleceń (*aggressive orders*), czyli takich, które wpływają na zmianę ceny papieru wartościowego.

Ciekawą propozycję estymacji elastyczności przedstawili Dong i inni [2007], stosując filtr Kalmana jako podstawowe narzędzie pomiaru. Autorzy analizowali elastyczność rynku na podstawie danych o ultrawysokiej częstotliwości z 294 dni transakcyjnych, od stycznia 2000 do marca 2001 r., dla 100 spółek z giełdy nowojorskiej NYSE. Zbadali też zależności korelacyjne pomiędzy kilkoma miarami płynności oraz elastycznością. Ponadto wykorzystali elastyczność jako zmienną objaśniającą w modelu wyceny papierów wartościowych.

W pracy z 2015 r. Kim i Kim zastosowali dyskretną transformatę Fouriera (*Discrete Fourier Transform* – DFT) do konstrukcji estymatora elastyczności, przekształcając logarytmy cen akcji do postaci spektralnej w dziedzinie częstotliwości. Autorzy podkreślili, że zaproponowany estymator może służyć do pomiaru szybkości ruchu powrotnego ceny papieru wartościowego do poziomu równowagi.

2. Estymacja elastyczności rynku na giełdzie warszawskiej

2.1. Opis bazy danych

Baza danych spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie S.A., w okresie od 3 stycznia 2005 r. do 30 czerwca 2015 r., objęła grupę 147 firm. Uzasadnienie wyboru metody tworzenia bazy zawierającej stałą liczbę spółek w całym analizowanym okresie można znaleźć np. w pozycji [Olbryś, 2014]. Poniższy, skrócony opis bazy danych przedstawiono na podstawie pracy [Nowak, Olbryś, 2015, s. 69–70].

W celu identyfikacji firm dużych, średnich i małych spółki posortowano według wartości rynkowej MV w ostatnim dniu roboczym grudnia. W każdym roku wyodrębniono trzy grupy (BIG, MEDIUM, SMALL), wykorzystując wartości percentyli 70% oraz 30% [Fama, French, 1993]. Następnie wybrano te spółki, które w całym analizowanym okresie pozostawały w tej samej grupie. Ostatecznie wszystkie warunki spełniły 53 spółki. Otrzymano w ten sposób trzy reprezentatywne grupy spółek. Odpowiednio, w grupie BIG znalazło się 27 spółek, w grupie MEDIUM – 18 spółek, natomiast w grupie SMALL – 8 spółek. Przeprowadzenie badania w wymienionych grupach umożliwiło analizę elastyczności rynku pod kątem występowania efektu wielkości spółki.

Dodatkowym celem pracy była analiza wrażliwości uzyskanych wyników na wybór okresu badania, z uwzględnieniem całej próby statystycznej

(styczeń 2005 – czerwiec 2015) oraz trzech jednakowo licznych podokresów [Nowak, Olbryś, 2015, s. 70]:

1) przed kryzysem: 6 września 2005 r. – 31 maja 2007 r.,
2) kryzys: 1 czerwca 2007 r. – 27 lutego 2009 r.,
3) po kryzysie: 2 marca 2009 r. – 19 listopada 2010 r.

Wymienione przedziały czasu ustalono na podstawie wyników pracy [Olbryś, Majewska, 2015]. Długości podokresów przed i po kryzysie zostały dopasowane do długości okresu kryzysu, w celu uzyskania porównywalności wyników. Liczebność całej próby statystycznej wyniosła 2626, natomiast każdy z podokresów zawierał po 436 dni.

2.2. Zrealizowana rozpiętość cen sprzedaż/kupno jako estymator elastyczności rynku

Jak podkreślono w pierwszej części, zrealizowana rozpiętość cen sprzedaż/kupno może być wykorzystana jako estymator elastyczności rynku. W niniejszej pracy zastosowano zrealizowaną rozpiętość cen sprzedaż/kupno zaproponowaną w artykule [Goyenko i inni, 2009, s. 156] i przekształconą do postaci rozpiętości procentowej (%RealS):

$$\%RealS_t = \begin{cases} 200 \cdot \ln \dfrac{P_t}{P_{t+5}}, \text{ gdy transakcja } t \text{ jest sklasyfikowana jako } buy \\ 200 \cdot \ln \dfrac{P_{t+5}}{P_t}, \text{ gdy transakcja } t \text{ jest sklasyfikowana jako } sell \end{cases} \quad (1)$$

gdzie P_t oznacza cenę zamknięcia transakcji t, natomiast cena P_{t+5} jest ceną zamknięcia transakcji o numerze $t+5$. Jako ultrakrótka logarytmiczna stopa zwrotu estymator (1) może być dowolnego znaku oraz jest równy zero, gdy $P_t = P_{t+5}$. Ze wzoru (1) wynika, że do oszacowania tego estymatora niezbędna jest informacja na temat stron inicjujących poszczególne transakcje, z wyróżnieniem transakcji inicjowanych przez nabywcę (*buy*) lub sprzedawcę (*sell*). Przez stronę inicjującą daną transakcję rozumiemy inwestora, który jako ostatni złożył zlecenie przed transakcją, która doszła do skutku [Odders-White, 2000, s. 262]. Niestety, szczegółowe dane transakcyjne dotyczące ofert kupna i sprzedaży nie są archiwizowane przez giełdę warszawską. Istnieje jednak możliwość pośredniej identyfikacji stron inicjujących transakcje z wykorzystaniem odpowiednich algorytmów klasyfikacji, które są stosowane na świecie od lat 90. XX wieku. Najbardziej znane procedury to: 1) reguła cenowa, 2) reguła tikowa, 3) algorytm LR [Lee, Ready, 1991] oraz 4) algorytm EMO [Ellis i inni, 2000]. W prezentowanym badaniu wykorzystano algorytm LR, ponieważ jest on najczęściej stosowany w badaniach empirycznych na rynkach kapitałowych w różnych krajach, a co więcej w pracy [Olbryś, Mursztyn, 2015] pokazano, że metoda ta skutecznie klasyfikuje transakcje na giełdzie warszawskiej, niezależnie od wielkości spółki i wyboru okresu badania. Ponadto Theissen

[2001] wykorzystał algorytm LR do estymacji zrealizowanej rozpiętości cen sprzedaż/kupno na giełdzie we Frankfurcie i potwierdził wystarczającą precyzję klasyfikacji stron inicjujących transakcje.

Po obliczeniu wartości estymatora (1) dla poszczególnych transakcji w ciągu dnia szacowana jest dzienna procentowa zrealizowana rozpiętość cen sprzedaż/kupno jako wartość średnia procentowych zrealizowanych rozpiętości wszystkich transakcji w danym dniu, ważona wolumenem transakcji. Ponadto dzienny procentowy zrealizowany spread %RealS jest zdefiniowany jako równy zero, w przypadku gdy wszystkie transakcje w danym dniu pozostają niesklasyfikowane po zastosowaniu algorytmu Lee-Ready.

2.3. Pomiar elastyczności rynku z wykorzystaniem współczynnika płynności

Zgodnie z literaturą elastyczność rynku można również szacować, stosując różne wersje tzw. współczynnika płynności (*liquidity ratio*), np. [Ranaldo, 2001, s. 313; von Wyss, 2004, s. 18–19]. Zatem w prezentowanym badaniu jako alternatywny estymator elastyczności wykorzystano zmodyfikowaną wersję współczynnika płynności z pracy [von Wyss, 2004]. Przekształcono wzór do postaci procentowej, zastosowano logarytmiczną stopę zwrotu z transakcji t oraz średnią ważoną wolumenem poszczególnych transakcji w ciągu dnia zamiast zwykłej średniej arytmetycznej. Ostateczny wzór na dzienny procentowy współczynnik płynności %LR ma postać (2):

$$\%LR = 100 \cdot \frac{\sum_{t=1}^{N} \left| r_t \right| \cdot V_t}{V} \tag{2}$$

gdzie $r_t = \ln \frac{P_t}{P_{t-1}}$ oznacza logarytmiczną stopę zwrotu z transakcji t, P_t jest ceną zamknięcia transakcji t, V_t jest wolumenem transakcji t, $V = \sum_{t=1}^{N} V_t$ jest skumulowanym dziennym wolumenem wszystkich transakcji, natomiast N oznacza liczbę transakcji w danym dniu (z wyłączeniem pierwszej transakcji).

Wartość estymatora (2) jest zawsze nieujemna. Dzienny procentowy współczynnik płynności %LR definiujemy jako równy zero, w przypadku gdy skumulowany dzienny wolumen transakcji w danym dniu (czyli mianownik wzoru 2) jest równy zero.

2.4. Wyniki empiryczne zrealizowanej rozpiętości cen sprzedaż/kupno oraz współczynnika płynności na giełdzie warszawskiej

W pierwszej kolejności wyznaczono wartości średnie alternatywnych estymatorów elastyczności rynku (1) oraz (2), w trzech grupach spółek,

w czterech analizowanych okresach obejmujących całą próbę statystyczną oraz trzy wyróżnione podokresy. Wyniki zestawiono w tablicy 1.

Tablica 1. Wartości średnie dziennej procentowej zrealizowanej rozpiętości cen sprzedaż/kupno oraz dziennego procentowego współczynnika płynności dla spółek z grup BIG, MEDIUM i SMALL, w podokresach

	%RealS (wzór 1)				
	BIG	P_1	P_2	P_3	P_4
1	**BHW**	0,09	0,14	0,11	0,11
2	**BPH**	0,14	0,07	0,16	0,20
3	**BNP**	0,03	0,01	0,06	–0,001
4	**BOS**	0,04	0,01	0,007	0,06
5	**BDX**	0,11	0,03	0,15	0,10
6	**BZW**	0,07	0,09	0,08	0,07
7	**DBC**	0,09	0,12	0,07	0,12
8	**ECH**	0,16	0,10	0,17	0,22
9	**GTN**	0,11	0,15	0,08	0,07
10	**GTC**	0,09	0,11	0,05	0,08
11	**ING**	0,10	0,13	0,09	0,09
12	**KTY**	0,13	0,22	0,13	0,09
13	**KGH**	0,02	0,02	0,03	0,03
14	**LPP**	0,09	0,05	0,19	0,12
15	**MBK**	0,07	0,13	0,06	0,08
16	**MIL**	0,11	0,15	0,13	0,10
17	**MOL**	0,07	0,17	0,14	0,04
18	**NET**	0,12	0,15	0,13	0,13
19	**OPL**	0,04	0,04	0,04	0,04
20	**ORB**	0,13	0,17	0,10	0,17
21	**PEO**	0,03	0,03	0,04	0,04
22	**PKN**	0,02	0,02	0,03	0,02
23	**PKO**	0,03	0,04	0,03	0,03
24	**STP**	0,14	0,16	0,22	0,16
25	**SNS**	0,12	0,12	0,20	0,20
26	**TVN**	0,08	0,09	0,08	0,08
27	**ZWC**	0,03	0,03	0,04	0,03
	średnia	**0,08**	**0,09**	**0,10**	**0,10**
	MEDIUM	P_1	P_2	P_3	P_4
1	**ALM**	0,22	0,26	0,20	0,23
2	**AMC**	0,17	0,21	0,25	0,17
3	**ATG**	0,12	0,03	-0,003	0,04
4	**ATM**	0,16	0,19	0,25	0,16
5	**CNG**	0,14	0,17	0,08	0,18

	%LR (wzór 2)				
	BIG	P_1	P_2	P_3	P_4
1	**BHW**	0,24	0,18	0,35	0,28
2	**BPH**	0,41	0,20	0,39	0,37
3	**BNP**	0,66	0,56	1,30	0,31
4	**BOS**	0,55	0,51	0,61	0,60
5	**BDX**	0,35	0,56	0,37	0,27
6	**BZW**	0,20	0,22	0,21	0,17
7	**DBC**	0,47	0,32	0,59	0,50
8	**ECH**	0,42	0,40	0,46	0,41
9	**GTN**	0,21	0,19	0,20	0,18
10	**GTC**	0,23	0,21	0,21	0,21
11	**ING**	0,29	0,32	0,37	0,30
12	**KTY**	0,37	0,31	0,49	0,36
13	**KGH**	0,09	0,09	0,12	0,11
14	**LPP**	0,47	0,50	0,68	0,52
15	**MBK**	0,20	0,24	0,21	0,19
16	**MIL**	0,24	0,27	0,26	0,21
17	**MOL**	0,60	0,32	0,78	0,83
18	**NET**	0,20	0,20	0,27	0,22
19	**OPL**	0,13	0,12	0,13	0,13
20	**ORB**	0,52	0,37	0,57	0,64
21	**PEO**	0,12	0,13	0,16	0,13
22	**PKN**	0,11	0,07	0,15	0,12
23	**PKO**	0,10	0,10	0,13	0,11
24	**STP**	0,58	0,51	0,64	0,51
25	**SNS**	0,26	0,32	0,40	0,32
26	**TVN**	0,18	0,17	0,21	0,19
27	**ZWC**	0,46	0,40	0,79	0,51
	średnia	**0,32**	**0,29**	**0,41**	**0,32**
	MEDIUM	P_1	P_2	P_3	P_4
1	**ALM**	0,69	0,61	0,68	0,49
2	**AMC**	0,39	0,43	0,61	0,34
3	**ATG**	0,90	0,72	1,15	1,07
4	**ATM**	0,67	0,49	0,63	0,56
5	**CNG**	0,56	0,40	0,59	0,48

	%RealS (wzór 1)				
	BIG	P_1	P_2	P_3	P_4
6	**COL**	0,17	0,04	0,20	0,19
7	**IND**	0,09	0,10	0,08	0,20
8	**IPL**	0,11	0,17	0,08	0,10
9	**LTX**	0,15	0,17	0,18	0,19
10	**MCI**	0,16	0,20	0,15	0,12
11	**MNI**	0,21	0,23	0,20	0,15
12	**PEK**	0,14	0,23	0,22	0,07
13	**PUE**	0,05	0,02	0,06	0,08
14	**SKA**	0,02	0,04	0,08	– 0,02
15	**STF**	0,13	0,18	0,19	0,17
16	**STX**	0,19	0,17	0,15	0,18
17	**TIM**	0,11	0,20	0,15	0,06
18	**VST**	0,20	0,07	0,20	0,19
	średnia	**0,14**	**0,15**	**0,16**	**0,14**
	SMALL	P_1	P_2	P_3	P_4
1	**APL**	0,17	0,31	0,27	0,29
2	**BDL**	0,25	0,22	0,29	0,24
3	**EFK**	0,14	0,32	0,21	0,005
4	**ENP**	0,24	0,46	0,28	0,29
5	**KMP**	0,26	0,38	0,33	0,40
6	**MZA**	0,20	0,41	0,27	0,11
7	**PLA**	0,15	0,26	0,20	0,24
8	**SME**	0,15	0,45	0,04	0,12
	średnia	**0,20**	**0,35**	**0,24**	**0,21**

	%LR (wzór 2)				
	BIG	P_1	P_2	P_3	P_4
6	**COL**	0,44	0,53	0,68	0,28
7	**IND**	0,90	0,85	1,48	0,90
8	**IPL**	0,63	0,48	0,78	0,56
9	**LTX**	0,35	0,37	0,37	0,37
10	**MCI**	0,27	0,32	0,28	0,22
11	**MNI**	0,45	0,34	0,40	0,33
12	**PEK**	0,73	0,54	0,84	0,90
13	**PUE**	0,85	0,86	1,10	0,88
14	**SKA**	0,76	0,68	1,05	0,78
15	**STF**	0,61	0,34	0,58	0,59
16	**STX**	0,35	0,28	0,27	0,28
17	**TIM**	0,77	0,48	0,77	0,71
18	**VST**	0,47	0,47	0,67	0,30
	średnia	**0,60**	**0,51**	**0,72**	**0,56**
	SMALL	P_1	P_2	P_3	P_4
1	**APL**	0,69	0,66	0,97	0,84
2	**BDL**	0,98	0,42	0,49	0,42
3	**EFK**	1,06	0,78	1,26	0,97
4	**ENP**	0,96	0,77	1,07	0,77
5	**KMP**	0,83	0,99	0,89	0,85
6	**MZA**	0,98	0,76	1,28	1,01
7	**PLA**	0,78	0,58	0,84	0,68
8	**SME**	1,16	1,05	1,10	1,13
	średnia	**0,93**	**0,75**	**0,99**	**0,83**

Objaśnienia: Oznaczenia spółek w postaci trzyliterowych symboli. Spółki wymieniono według pełnych nazw w kolejności alfabetycznej. Oznaczenia okresów: P_1 – cały okres, styczeń 2005 – czerwiec 2015 (2626 obserwacji), P_2 – podokres przed kryzysem, wrzesień 2005 – maj 2007 (436 obserwacji), P_3 – podokres kryzys, czerwiec 2007 – luty 2009 (436 obserwacji), P_4 – podokres po kryzysie, marzec 2009 – listopad 2010 (436 obserwacji).
Źródło: Opracowanie własne.

Zgodnie z literaturą procentowa zrealizowana rozpiętość cen sprzedaż/kupno %RealS mierzy szybkość ruchu powrotnego ceny papieru wartościowego (*price reversal component*), np. [Huang, Stoll, 1996]. Średnie dzienne wartości tego estymatora są zwykle dodatnie. Otrzymane wyniki empiryczne potwierdziły ten fakt dla prawie wszystkich spółek, poza nielicznymi wyjątkami (tab. 1). Ponadto nie zaobserwowano efektu wielkości spółki, ponieważ zarówno wysokie, jak i niskie (co do modułu) wartości indykatora (1) wystąpiły we wszystkich grupach spółek. Wyniki okazały się

odporne na wybór okresu badania oraz nie stwierdzono, aby okres kryzysu wyróżniał się pod względem wartości %RealS na tle pozostałych okresów.

Podobnie jak w przypadku estymatora (1), konstrukcja procentowego współczynnika płynności %LR (wzór 2) opiera się na ultrakrótkiej stopie zwrotu z pojedynczych transakcji. Średnie dzienne wartości tego estymatora są dodatnie. Wyniki zestawione w tablicy 1 pozwalają stwierdzić brak wyraźnego efektu wielkości spółki oraz odporność wartości estymatora (2) na wybór okresu badania. Analogicznie jak w przypadku estymatora (1) nie stwierdzono, aby okres kryzysu wyróżniał się pod względem wartości %LR na tle pozostałych okresów.

2.5. Testowanie stabilności zależności korelacyjnych pomiędzy estymatorami elastyczności rynku w podokresach

Kolejnym etapem badania była weryfikacja hipotezy badawczej zakładającej stabilność zależności korelacyjnych pomiędzy estymatorami elastyczności. W tym celu testowano równość macierzy korelacji w podokresach: (kryzys, przed kryzysem) oraz (kryzys, po kryzysie). Sformułowano następującą hipotezę zerową:

$$H_0 : P_K = P_{PK} \tag{3}$$

wobec hipotezy alternatywnej:

$$H_0 : P_K \neq P_{PK} \tag{4}$$

gdzie P_K, P_{PK} oznaczają macierze korelacji w podokresach kryzysu (P_K) oraz przed lub po kryzysie (P_{PK}). Statystyczna weryfikacja hipotezy zerowej została przeprowadzona z wykorzystaniem modyfikacji testu Larntza–Perlmana [1985] równości macierzy zawierających współczynniki korelacji Pearsona, w dwóch jednakowo licznych podokresach ($n_K = n_{PK} = n = 436$ obserwacji). Procedura testowania równości macierzy o jednakowym wymiarze (ale niekoniecznie symetrycznych) została zaproponowana w pracy [Nowak, Olbryś, 2015] i składała się z następujących etapów:

1. Przekształcenie wszystkich elementów obu macierzy $\hat{P}_K = (\hat{r}_{ij}^K)$ oraz $\hat{P}_{PK} = (\hat{r}_{ij}^{PK})$, zawierających korelacje empiryczne, za pomocą transformacji z-Fishera [1921], określonej wzorem (5):

$$\hat{z}_{ij} = \frac{1}{2}\ln\left(\frac{1+\hat{r}_{ij}}{1-\hat{r}_{ij}}\right) \tag{5}$$

2. Obliczenie wartości statystyki testowej T, danej wzorem (6):

$$T = \sqrt{\frac{n-3}{2}} \cdot \max_{\substack{1\le i\le p_1 \\ 1\le j\le p_2}} \left|\hat{z}_{ij}^K - \hat{z}_{ij}^{PK}\right| \tag{6}$$

gdzie p_1, p_2 oznaczają odpowiednio liczbę wierszy (p_1) oraz kolumn (p_2) analizowanych macierzy.

3. Dla ustalonego poziomu istotności α, odrzucenie hipotezy zerowej (3), gdy wartość statystyki testowej $T > b_\alpha$, gdzie $b_\alpha > 0$ spełnia następującą równość:

$$[\Phi(b_\alpha) - \Phi(-b_\alpha)]^{p_1 \cdot p_2} = 1 - \alpha \tag{7}$$

gdzie $\Phi(b_\alpha)$ jest wartością dystrybuanty rozkładu normalnego standardowego, natomiast $p_1 \cdot p_2$ jest liczbą wszystkich współczynników korelacji w danej macierzy.

Wyniki przedstawione w tablicy 2 wskazują, że należy odrzucić hipotezę zerową (3) o równości macierzy współczynników korelacji pomiędzy estymatorami elastyczności rynku w podokresach: (kryzys, przed kryzysem) oraz (kryzys, po kryzysie).

Tablica 2. Wyniki zmodyfikowanego testu Larntza–Perlmana równości macierzy współczynników korelacji pomiędzy estymatorami elastyczności

Testowane podokresy	Statystyka testowa T	Wartość krytyczna b_α (5%)	Statystyka testowa T	Wartość krytyczna b_α (1%)
P_3, P_2	5,73	3,30 (H_1)	5,73	3,73 (H_1)
P_3, P_4	4,26	3,30 (H_1)	4,26	3,73 (H_1)

Objaśnienia: jak do tablicy 1.
Źródło: Opracowanie własne.

Zróżnicowanie zależności korelacyjnych pomiędzy procentową zrealizowaną rozpiętością cen sprzedaż/kupno %RealS a procentowym współczynnikiem płynności %LR pokazują również tablice 3 i 4. W tablicy 3 zestawiono podstawowe statystyki opisowe transformacji z-Fishera (wzór 5) współczynników korelacji Pearsona badanych estymatorów w trzech grupach spółek. Natomiast tablica 4 prezentuje procent istotnych statystycznie współczynników korelacji w grupach i podokresach. Jest on stosunkowo wysoki w większości przypadków, jednak zgodnie z oczekiwaniami powinien być wyższy, ponieważ oba estymatory szacują ten sam wymiar płynności oraz mają podobną konstrukcję opartą na ultrakrótkiej, śródziennej stopie zwrotu.

Tablica 3. Podstawowe statystyki opisowe transformacji *z*-Fishera (5) współczynników korelacji Pearsona pomiędzy estymatorami elastyczności

		Min	Max	Mediana	Średnia	Odchylenie standardowe
BIG (27 firm)	P_1	0,008	0,258	0,117	0,108	0,059
	P_2	–0,069	0,377	0,156	0,149	0,115
	P_3	–0,042	0,287	0,074	0,095	0,079
	P_4	–0,074	0,417	0,069	0,093	0,099
MEDIUM (18 firm)	P_1	0,023	0,383	0,079	0,106	0,084
	P_2	–0,006	0,398	0,233	0,220	0,126
	P_3	–0,045	0,528	0,136	0,162	0,144
	P_4	0,020	0,329	0,108	0,133	0,097
SMALL (8 firm)	P_1	0,006	0,143	0,099	0,090	0,040
	P_2	0,190	0,420	0,266	0,296	0,084
	P_3	0,044	0,284	0,137	0,144	0,083
	P_4	0,057	0,221	0,089	0,113	0,058

Objaśnienia: jak do tablicy 1.
Źródło: Opracowanie własne.

Tablica 4. Procent istotnych statystycznie współczynników korelacji pomiędzy estymatorami elastyczności

	Korelacje (%RealS, %LR)			
	P_1	P_2	P_3	P_4
BIG	88,89%	62,96%	40,74%	37,04%
MEDIUM	83,33%	83,33%	72,22%	50,00%
SMALL	87,50%	100,0%	62,50%	50,00%

Objaśnienia: Wartość krytyczna współczynnika korelacji dla całej próby statystycznej P_1 (2626 obserwacji) wynosi 0,038, natomiast w przypadku podokresów P_2, P_3 i P_4 (436 obserwacji) jest równa 0,094. Pozostałe objaśnienia jak do tablicy 1.
Źródło: Opracowanie własne.

Zakończenie

Głównym celem pracy była analiza empiryczna elastyczności rynku (*market resiliency*) jako jednego z wymiarów płynności Giełdy Papierów Wartościowych w Warszawie SA. Dodatkowym celem było badanie wrażliwości uzyskanych wyników na wybór okresu, z uwzględnieniem całej próby statystycznej (styczeń 2005–czerwiec 2015) oraz trzech jednakowo licznych podokresów: przed kryzysem, kryzys, po kryzysie. Badanie objęło 53 spółki podzielone według wartości rynkowej na grupy firm dużych, średnich i małych. Dokonano aproksymacji elastyczności rynku poszczególnych papierów wartościowych z wykorzystaniem sekundowych danych transakcyjnych oraz dwóch alternatywnych estymatorów elastyczności:

1) procentowej zrealizowanej rozpiętości cen sprzedaż/kupno (%RealS) oraz 2) procentowego współczynnika płynności (%LR). Ponadto przeprowadzono analizę zależności korelacyjnych pomiędzy alternatywnymi estymatorami elastyczności i zweryfikowano, z wykorzystaniem modyfikacji testu Larntza–Perlmana [1985], hipotezę badawczą zakładającą stabilność tych zależności w czasie.

Wyniki empiryczne pozwoliły sformułować następujące wnioski:

1. Uzyskane wartości estymatorów elastyczności nie ujawniły wyraźnego efektu wielkości spółki oraz wykazały odporność na wybór okresu badania.
2. Stwierdzono konieczność odrzucenia hipotezy zerowej o stabilności zależności korelacyjnych pomiędzy estymatorami elastyczności w podokresach: (kryzys, przed kryzysem) oraz (kryzys, po kryzysie).
3. Otrzymano stosunkowo wysoki procent istotnych statystycznie współczynników korelacji pomiędzy alternatywnymi estymatorami elastyczności w grupach i podokresach, jednak, zgodnie z oczekiwaniami, powinien on być wyższy. Obserwacja ta potwierdza znany z literatury fakt, że elastyczność rynku jest najtrudniejszym do oszacowania wymiarem płynności i nie ma jednego, powszechnie akceptowanego estymatora tego wymiaru.

Badania elastyczności rynku będą kontynuowane. W dalszych analizach tego wymiaru płynności na giełdzie warszawskiej planowane jest zastosowanie dyskretnej transformaty Fouriera w celu oszacowania estymatora elastyczności z wykorzystaniem śróddziennych danych transakcyjnych. Będzie to propozycja modyfikacji metody przedstawionej w pracy [Kim, Kim, 2015]. Dotychczasowe wyniki badań empirycznych są obiecujące, jednak znajdują się obecnie w fazie wstępnej.

Literatura

Alan N.S., Hua J., Peng L., Schwartz R.A. (2015), *Stock resiliency and expected returns*, Working Paper.

Bernstein P. L. (1987), *Liquidity, stock markets, and market makers*, „Financial Management", Vol. 16, No. 2.

Będowska-Sójka B. (2014), *Liquidity dynamics around jumps. The evidence from the Warsaw Stock Exchange*, Available at SSRN 2500715.

Bhattacharya S.N., Sengupta P., Bhattacharya M., Roychoudhury (2016), *Multidimensional liquidity: Evidence from the Indian stock market*, „Applied Finance Letters", Vol. 5, No. 2.

Coppejans M., Domowitz I., Madhavan A. (2004), *Resiliency in an automated auction*, Working Paper.

Degryse H., de Jong F., van Ravenswaaij M., Wuyts G. (2003), *Aggressive orders and the resiliency of a limit order market*, Working Paper.

Doman M. (2011), *Mikrostruktura giełd papierów wartościowych*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.

Dong J., Kempf A., Yadav P.K. (2007), *Resiliency, the neglected dimension of market liquidity: empirical evidence from the New York Stock Exchange*, Working Paper, Available at SSRN 967262.

Ellis K., Michaely R., O'Hara M. (2000), *The accuracy of trade classification rules: Evidence from Nasdaq*, „Journal of Financial and Quantitative Analysis", Vol. 35, No. 4.

Engle R., Lange J. (1997), *Measuring, forecasting and explaining time varying liquidity in the stock market*, NBER Working Paper 6129.

Fama E.F., French K.R. (1993), *Common Risk Factors in the Returns on Stocks and Bonds*, „Journal of Financial Economics", Vol. 33, No. 1.

Fisher R.A. (1921), *On the "Probable Error" of a Coefficient of Correlation Deduced From a Small Sample*, „Metron", No. 1.

Foucault T., Kadan O., Kandel E. (2005), *Limit order book as a market for liquidity*, „Review of Financial Studies", Vol. 18, No. 4.

Foucault T., Pagano M., Röell A. (2013), *Market liquidity. Theory, evidence, and policy*, Oxford University Press.

Glosten L.R. (1987), *Components of the bid-ask spread and the statistical properties of transaction prices*, „Journal of Finance", Vol. 42, No. 4.

Goyenko R.Y., Holden C.W., Trzcinka C.A. (2009), *Do liquidity measures measure liquidity?*, „Journal of Financial Economics", Vol. 92, No. 2.

Harris L. (2003), *Trading and exchanges: Market microstructure for practitioners*, Oxford University Press.

Hasbrouck J., Seppi D.J. (2001), *Common factors in prices, order flows, and liquidity*, „Journal of Financial Economics", Vol. 59, No. 3.

Hmaied D.M., Grar A., Sioud O.B. (2006), *Dynamics of market liquidity of Tunisian stocks: an analysis of market resiliency*, „Electronic Markets", Vol. 16, No. 2.

Huang R.D., Stoll H.R. (1996), *Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE*, „Journal of Financial Economics", Vol. 41, No. 2.

Huang R.D., Stoll H.R. (1997), *The components of the bid-ask spread: A general approach*, „Review of Financial Studies", Vol. 10, No. 4.

Jankowski R., Olbryś J. (2015), *Wymiary płynności rynku papierów wartościowych*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego", nr 854, „Finanse, Rynki Finansowe, Ubezpieczenia", nr 73.

Kempf A., Mayston D., Yadav P.K. (2009), *Resiliency in limit order book markets: A dynamic view of liquidity*, Working Paper.

Kim J., Kim Y. (2015), *Transitory price, resiliency, and the cross-section of stock returns*, Working Paper.

Kyle A.S. (1985), *Continuous auctions and insider trading*, „Econometrica", Vol. 53, No. 6.

Large J. (2007), *Measuring the resiliency of an electronic limit order book*, „Journal of Financial Markets", Vol. 10, No. 1.

Lin W. T., Sun D. S., Tsai S.-C. (2012), *Does trading remove or cause friction?*, „Emerging Markets Finance & Trade", Vol. 48 (Supp. 2).

Lo D.K., Hall A.D. (2015), *Resiliency of the limit order book*, „Journal of Economic Dynamics & Control", Vol. 61.

Muranaga J. (2000), *Dynamics of market liquidity of Japanese stocks: An analysis of tick-by-tick data of the Tokyo Stock Exchange,* Working Paper, Bank of Japan.

Muranaga J., Shimizu T. (1999), *Market microstructure and market liquidity*, Working Paper, Bank of Japan.

Nowak S., Olbryś J. (2015), *Zależności korelacyjne w szeregach stóp zwrotu spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie S.A.*, „Zarządzanie i Finanse", nr 4, cz. 2.

Odders-White E.R. (2000), *On the occurrence and consequences of inaccurate trade classification*, „Journal of Financial Markets", Vol. 3, No. 3.

Olbryś J. (2014), *Wycena aktywów kapitałowych na rynku z zakłóceniami w procesach transakcyjnych*, Difin, Warszawa.

Olbryś J. (2017), *Interaction between market depth and market tightness on the Warsaw Stock Exchange: A preliminary study*, w: K. Jajuga, L. Orlowski, K. Staehr (eds.) *Contemporary Trends and Challenges in Finance*. „Springer Proceedings in Business and Economics", Springer, Cham.

Olbryś J., Majewska E. (2015), *Bear market periods during the 2007–2009 financial crisis: Direct evidence from the Visegrad countries*, „Acta Oeconomica", Vol. 65, No. 4.

Olbryś J., Mursztyn M. (2015), *Comparison of selected trade classification algorithms on the Warsaw Stock Exchange*, „Advances in Computer Science Research", No. 12.

Olbryś J., Mursztyn M. (2016), *Głębokość rynku jako jeden z wymiarów płynności Giełdy Papierów Wartościowych w Warszawie S.A.*, „Finanse, Rynki Finansowe, Ubezpieczenia", nr 79.

Olbryś J., Mursztyn M. (2017a), *Pomiar naprężenia rynku na Giełdzie Papierów Wartościowych w Warszawie SA*, „Finanse, Rynki Finansowe, Ubezpieczenia", nr 85.

Olbryś J., Mursztyn M. (2017b), *Dimensions of market liquidity: The case of the Polish stock market*, w: N. Tsounis, A. Vlachvei (eds.) *Advances in Applied Economic Research*, „Springer Proceedings in Business and Economics", Springer, Cham.

Ranaldo A. (2001), *Intraday market liquidity on the Swiss Stock Exchange*, „Swiss Society for Financial Market Research", Vol. 15, No. 3.

Wong J., Fung L. (2002), *Liquidity of the Hong Kong stock market since the Asian financial crisis*, „Proceedings of the Third Joint Central Bank Research Conference", Bank for International Settlements, Basel, Switzerland.

Wyss von R. (2004), *Measuring and predicting liquidity in the stock market*, Dissertation No. 2899, University of St. Gallen.

Użyte skróty

%RealS – procentowa zrealizowana rozpiętość cen sprzedaż/kupno (*percentage realized bid/ask spread*)

%LR – procentowy współczynnik płynności (*percentage liquidity ratio*)

Streszczenie

Zgodnie z literaturą uzupełnieniem badania płynności/niepłynności aktywów kapitałowych może być analiza wymiarów płynności jako dodatkowych

charakterystyk rynku giełdowego. Do wymiarów płynności najczęściej zaliczane są: naprężenie, głębokość oraz elastyczność (sprężystość) rynku [Kyle, 1985]. Głównym celem pracy była analiza empiryczna elastyczności rynku (*market resiliency*) jako jednego z wymiarów płynności Giełdy Papierów Wartościowych w Warszawie SA. Dodatkowym celem było badanie wrażliwości uzyskanych wyników na wybór okresu, z uwzględnieniem całej próby statystycznej (styczeń 2005–czerwiec 2015) oraz trzech jednakowo licznych podokresów: przed kryzysem, kryzys, po kryzysie. Badanie objęło 53 spółki podzielone według wartości rynkowej na grupy firm dużych, średnich i małych. Dokonano aproksymacji elastyczności rynku poszczególnych papierów wartościowych z wykorzystaniem sekundowych danych transakcyjnych oraz dwóch alternatywnych estymatorów elastyczności. Wyniki empiryczne nie ujawniły wyraźnego efektu wielkości spółki oraz wykazały odporność na wybór okresu badania. Ponadto przeprowadzono podstawową analizę zależności korelacyjnych pomiędzy alternatywnymi estymatorami elastyczności rynku i zweryfikowano hipotezę badawczą zakładającą stabilność tych zależności w czasie. Według wiedzy autorów analogiczne badanie nie było dotąd prowadzone na giełdzie warszawskiej.

Słowa kluczowe
wymiary płynności rynku, elastyczność rynku, zrealizowana rozpiętość cen sprzedaż/kupno, współczynnik płynności, Giełda Papierów Wartościowych w Warszawie SA

Resiliency as one of market liquidity dimensions on the Warsaw Stock Exchange (Summary)

According to the literature, liquidity in a financial market is not a one-dimensional variable but it includes several dimensions. The majority of researchers distinguish between three dimensions of market liquidity as special liquidity characteristics: market depth, market tightness, and market resiliency. The main goal of this paper was to examine market resiliency employing two alternative proxies of resiliency based on intraday data 'rounded to the nearest second', in the group of 53 companies listed on the Warsaw Stock Exchange. The whole sample covered period January 2005–December 2014 and three adjacent sub–periods of equal size: the pre–crisis, crisis, and post–crisis periods. The financial crisis on the WSE was formally established as the period June 2007–February 2009. The additional goal of the paper was to test the research hypothesis concerning stability of correlations between two proxies of market resiliency in the subperiods specified. The empirical results turned out to be robust to the choice of the period and rather did not depend on a firm size. To the best of authors' knowledge, the results for the WSE are novel and have not been presented in the literature thus far.

Keywords
dimensions of market liquidity, market resiliency, realized bid/ask spread, liquidity ratio, Warsaw Stock Exchange

Magdalena Gostkowska-Drzewicka*
Ewa Majerowska**

Oddziaływanie wybranych czynników kształtujących strukturę kapitału spółek notowanych na GPW w Warszawie

Wstęp

Pomimo niemal sześćdziesięciu lat badań nie udało się w pełni wyjaśnić, jak ukształtować optymalną strukturę kapitału, ani w sposób jednoznaczny określić, jakie czynniki o tym decydują. W literaturze przedmiotu wskazuje się różnorodne determinanty wpływające na decyzje w zakresie pozyskiwania kapitału. Stanowią one przedmiot wielu badań, na podstawie których określono kierunek ich oddziaływania i siłę wpływu na poziom dźwigni finansowej. Celem artykułu jest empiryczna weryfikacja czynników kształtujących strukturę kapitału spółek notowanych na GPW w Warszawie w latach 2000–2015. Sformułowana hipoteza zakłada, że w badanych podmiotach istnieją znaczące determinanty struktury kapitału, a siła ich wzajemnego oddziaływania zmienia się w czasie.

Realizacja tak sformułowanego celu i weryfikacja hipotezy badawczej wymagała zastosowania analizy głównych składowych. Jako zmienne diagnostyczne przyjęto: tempo wzrostu, płynność finansową, inwestycyjną tarczę podatkową, rentowność, wielkość i strukturę aktywów. Obliczenia zostały wykonane za pomocą programu Statistica.

Artykuł składa się ze wstępu, czterech części i zakończenia. W części pierwszej opisano najważniejsze czynniki wpływające na decyzje dotyczące pozyskiwania kapitału. Część druga zawiera charakterystykę struktury kapitału badanych przedsiębiorstw. Kolejna część ma charakter metodyczny. Zawiera opis metody głównych składowych. W kolejnej części zaprezentowano wyniki badań, a w zakończeniu dokonano ich podsumowania.

1. Czynniki struktury kapitału w literaturze przedmiotu

Pojęcie struktury kapitału jest wieloznaczne i różnie interpretowane. Przykładowo Jerzemowska [2013, s. 164] wskazuje, że najogólniej jest ona

* Dr, Katedra Finansów Przedsiębiorstw, Wydział Zarządzania, Uniwersytet Gdański, ul. Armii Krajowej 101, 81-824 Sopot, mgostkowska@wzr.ug.edu.pl
** Dr, Katedra Ekonometrii, Wydział Zarządzania, Uniwersytet Gdański, ul. Armii Krajowej 101, 81-824 Sopot, ewa.majerowska@ug.edu.pl

utożsamiana ze strukturą pasywów bilansu, tj. ze strukturą finansowania. Natomiast Duliniec [1998, s. 14] definiuje tę kategorię w odmienny sposób. Autorka podkreśla, że struktura kapitału odzwierciedla kapitał własny oraz zobowiązania długo- i krótkoterminowe, z wyłączeniem zobowiązań wobec dostawców, zobowiązań z tytułu podatków oraz wynagrodzeń. Takie podejście oznacza, że struktura kapitału różni się od struktury pasywów bilansu, ponieważ uwzględnia wyłącznie zobowiązania oprocentowane. W innym ujęciu struktura kapitału wyraża wyłącznie podział kapitału stałego (długoterminowego) na kapitał własny i zobowiązania długoterminowe. Kategoria ta bywa również rozumiana jako kombinacja dłużnych i właścicielskich papierów wartościowych wyemitowanych przez przedsiębiorstwo [Duliniec, 1998, s. 12–13].

W literaturze przedmiotu wskazuje się różnorodne czynniki kształtujące strukturę kapitału. Wśród nich najczęściej wskazywane są:

- wzrost,
- wielkość,
- struktura aktywów,
- rentowność,
- płynność finansowa,
- inwestycyjna tarcza podatkowa.

Przedsiębiorstwa o dużym potencjale wzrostu (rozwoju) dążą do utrzymania niskiej dźwigni finansowej. Realizacja możliwości rozwojowych wiąże się bowiem z wyższym ryzykiem i oczekiwanymi kosztami trudności finansowych, co skutkuje ograniczaniem zadłużenia. Należy podkreślić, że duży potencjał wzrostu często znajduje odzwierciedlenie we wzroście wskaźnika wartości rynkowej do wartości księgowej, co sprzyja korzystaniu z zewnętrznych źródeł kapitału własnego (emisja akcji), a nie z kapitału obcego [Duliniec, 2015, s. 80].

Duże przedsiębiorstwa wykazują wyższy poziom zadłużenia. Podmioty te zazwyczaj dywersyfikują działalność, dzięki czemu mają niewielkie ryzyko utraty płynności. Prowadzi to do zminimalizowania kosztów trudności finansowych i umożliwia zwiększenie dźwigni finansowej [Duliniec, 2015, s. 80]. Podobnego zdania są Rajan, Zingales [1995, s. 1421–1460], Barclay, Smith [1996, s. 4–17], Degryse i inni [2012, s. 432–447], Komera, Lukose [2015, s. 807–823], którzy podkreślają dodatni związek pomiędzy wielkością przedsiębiorstwa i poziomem zadłużenia. Warto zwrócić uwagę na fakt, że kierunek tego związku zależy od charakteru długu. Powiązania dodatnie występują w przypadku zadłużenia długoterminowego, a ujemne – krótkoterminowego [Bevan, Danbolt, 2002, s. 159–170]. Duże podmioty mają lepszy dostęp do rynku finansowego i są w stanie

negocjować korzystniejsze dla siebie warunki, na jakich zaciągają zobowiązania, niż firmy małe [Ferri, Jones, 1979, s. 631–644; Serrasqueiro i inni, 2016, s. 13–28].

Z badań przeprowadzonych w krajach wysoko rozwiniętych wynika, że przedsiębiorstwa dysponujące znacznymi zasobami aktywów rzeczowych mają wyższy poziom dźwigni finansowej [Rajan, Zingales, 1995, s. 1421–1460; Titman, Wessels, 1988, s. 1–19]. Aktywa rzeczowe, stanowiąc zabezpieczenie zobowiązań, umożliwiają bowiem zmniejszenie bezpośrednich kosztów bankructwa, co sprzyja większemu zadłużeniu [Chaklader, Chawla, 2016, s. 267–277].

Przedsiębiorstwa osiągające wysokie wskaźniki rentowności wykazują zazwyczaj niższe wskaźniki zadłużenia. Podmioty te mają bowiem większe możliwości finansowania rozwoju ze źródeł wewnętrznych. W rezultacie wykazują mniejsze zapotrzebowanie na kapitał obcy [Myers, Majluf, 1984, s. 187–222; Titman, Wessels, 1988, s. 1–19; Rajan, Zingales, 1995, s. 1421–1460; Bevan, Danbolt, 2002, s. 159–170].

Jak wspomniano powyżej, przedsiębiorstwa generujące wysokie zyski są w stanie w znacznym stopniu finansować działalność operacyjną i rozwojową ze źródeł wewnętrznych. Prowadzi to do zmniejszenia zapotrzebowania na kapitał, a zatem podmioty o wysokiej płynności finansowej wykazują niską dźwignię finansową [Mazur, 2007, s. 495–514; De Jong i inni, 2008, s. 1954–1969; Bevan, Danbolt, 2002, s. 159–170].

Efekt tarczy (oszczędności) podatkowej jest jedną z przyczyn prowadzących do wzrostu zadłużenia [Modigliani, Miller, 1963, s. 433–443]. Jednakże postępowanie takie jest atrakcyjne wyłącznie dla przedsiębiorstw, które generują dochód umożliwiający uzyskanie korzyści podatkowych, lecz nie wykazują innych kosztów, które działają podobnie jak tarcza podatkowa. Wśród kosztów takich wskazuje się przede wszystkim amortyzację, która w relacji do aktywów ogółem wyznacza poziom tzw. inwestycyjnej tarczy podatkowej. W przeciwieństwie do tarczy podatkowej prowadzi ona do redukcji zadłużenia [DeAngelo, Masulis, 1980, s. 3–29; Jordan i inni, 1998, s. 1–27].

2. Struktura kapitału w badanych przedsiębiorstwach

Podmiotem analizy są spółki notowane nieprzerwanie na rynku głównym Giełdy Papierów Wartościowych w Warszawie w latach 2000–2015 (tab. 1).

Tablica 1. Liczba spółek notowanych na GPW w latach 2000–2015 (stan na dzień 8 lutego 2017)

Wyszczególnienie	Liczba spółek ogółem	Liczba spółek zakwalifikowanych do analizy	Liczba spółek, które zostały odrzucone
Spółki z WIG 20	6	4	2
Spółki z WIG 40	17	14	3
Spółki z WIG 80	22	20	2
Pozostałe spółki	72	52	20
Razem	117	90	27

Źródło: Opracowanie własne na podstawie danych GPW.

W latach 2000–2015 na GPW w Warszawie w sposób nieprzerwany było notowanych 117 spółek, w tym 6 z WIG-u 20, 17 z WIG-u 40, 22 z WIG-u 80 i 72 spółki niezaliczone do żadnego z wymienionych indeksów. Spośród powyższych spółek do badań przyjęto wyłącznie podmioty z sektora niefinansowego. Z tego względu z próby wyłączono 16 spółek. Ponadto odrzucono podmioty, które nie przedstawiły kompletnych sprawozdań finansowych w badanym okresie, tj. 11 spółek. Ostatecznie do badań zakwalifikowało się 90 przedsiębiorstw, tj. 77% wstępnie wytypowanych podmiotów.

W tablicy 2 zaprezentowano wartości wskaźnika ogólnego zadłużenia badanych przedsiębiorstw w latach 2000–2015.

Tablica 2. Wskaźnik ogólnego zadłużenia w spółkach notowanych na GPW w latach 2000–2015

Lata	Wskaźnik ogólnego zadłużenia (zobowiązania ogółem / aktywa ogółem)	Lata	Wskaźnik ogólnego zadłużenia (zobowiązania ogółem / aktywa ogółem)
2000	0,50	2008	0,41
2001	0,52	2009	0,37
2002	0,56	2010	0,36
2003	0,63	2011	0,38
2004	0,63	2012	0,39
2005	0,64	2013	0,38
2006	0,46	2011	0,43
2007	0,38	2015	0,38
		Średnio	0,46

Źródło: Opracowanie własne na podstawie bazy danych Notoria.

Wskaźnik ogólnego zadłużenia obliczony dla zbiorowości spółek notowanych na GPW w latach 2000–2015 przyjmował wartości świadczące o niewielkim stopniu ich zadłużenia. Udział zobowiązań w finansowaniu aktywów ogółem podlegał pewnym fluktuacjom i kształtował się na poziomie wynoszącym od 0,36 (2010) do 0,64 (2005). W latach 2000–2005 wartości wskaźnika rosły systematycznie od 0,5 do 0,64. Oznacza to wzrost uzależnienia finansowego badanych podmiotów od kapitałów obcych. Jednakże należy podkreślić, że pomimo tych tendencji wartości wskaźnika nie przekraczały 0,67, tj. górnej granicy wyznaczającej poziom akceptowalnego ryzyka finansowego związanego z koniecznością spłaty kredytów wraz z odsetkami [Gabrusewicz, 2014, s. 342]. Począwszy od roku 2006 zauważyć można odwrócenie dotychczasowych tendencji. Wartości wskaźnika ogólnego zadłużenia w latach 2006–2015 przyjmowały wartości wynoszące poniżej 0,5. Niski poziom wskaźnika świadczy o niewielkim zadłużeniu analizowanych podmiotów oraz ich dużej niezależności finansowej. Należy podkreślić, że po pierwsze, konserwatyzm w zakresie kształtowania struktury finansowej skutkuje zwiększeniem odporności analizowanych spółek na ryzyka wynikające ze zmian o charakterze koniunkturalnym. Po drugie, mogą one liczyć na pozytywną ocenę kredytodawców. Preferują oni podmioty o niskim wskaźniku zadłużenia, ponieważ zapewnia on im poczucie większego bezpieczeństwa i zwiększa szanse na odzyskanie zainwestowanego kapitału w przypadku likwidacji przedsiębiorstwa.

Opisywane tendencje znajdują pokrycie w badaniach struktury kapitału przedsiębiorstw polskich przeprowadzonych przez innych autorów. Z badań Marzec [2010, s. 143–153] wynika, że kapitały własne są najważniejszym źródłem finansowania działalności, w tym inwestycji, podmiotów z sektora MSP w Polsce. Do takich samych wniosków doszła Janus [2006, s. 69–78]. Również Lisińska [2012, s. 449–458], Barburski [2014, s. 117–136] oraz Wrońska-Bukalska [2014, s. 393–402], wskazują kapitał własny jako podstawowe źródło finansowania przedsiębiorstw polskich. W przypadku badanych firm potwierdzenie znalazła teoria kolejności źródeł finansowania, według której przedsiębiorstwa preferują wewnętrzne źródła finansowania. Odróżnia je to od podmiotów niemieckich czy portugalskich, które w znacznie większym stopniu wykorzystywały kapitały zewnętrzne, a zwłaszcza obce, w finansowaniu aktywów. Podobna sytuacja została zaobserwowana w Czechach w latach 2007–2011, gdzie w większości przedsiębiorstw motoryzacyjnych kapitały obce stanowiły główne źródło finansowania aktywów [Konečný, 2013, s. 83–97].

3. Opis metody badawczej

Do analizy badanych zmiennych wykorzystać można metodę głównych składowych, będącą rozszerzeniem analizy czynnikowej. Celem metody głównych składowych jest wyodrębnienie najistotniejszych czynników spośród zestawu zmiennych pierwotnych, pozwala ona zatem na przekształcenie dowolnej liczby zmiennych na zbiór niezależnych składowych. Każda ze składowych wyjaśnia część całkowitej zmienności, którą reprezentuje wariancja.

Niech X_1, X_2,…X_k oznacza zbiór k zmiennych oraz niech macierz wariancji i kowariancji pomiędzy tymi zmiennymi nie posiada macierzy odwrotnej. Oznacza to, że kolumny macierzy wariancji i kowariancji są liniowo zależne względem siebie. Idea metody głównych składowych polega na takim przekształceniu wyjściowych zmiennych, aby wyeliminować efekt współliniowości. Tworzony zestaw nowych zmiennych, zwanych głównymi składowymi, stanowi liniową kombinację zmiennych postaci [Aczel, 1989, s. 951]:

$$\begin{aligned} Y_1 &= a_{11}X_1 + a_{21}X_2 + \ldots + a_{k1}X_k \\ Y_2 &= a_{12}X_1 + a_{22}X_2 + \ldots + a_{k2}X_k \\ &\vdots \\ Y_k &= a_{1k}X_1 + a_{2k}X_2 + \ldots + a_{kk}X_k \end{aligned} \tag{1}$$

gdzie a_{ij} oznacza stałe wartości parametrów poszczególnych równań, przy czym $i = 1, 2,\ldots,k$ oraz $j = 1, 2,\ldots,k$.

Jak można zauważyć, nowo utworzone zmienne są niezależne względem siebie. Pierwsza składowa wyjaśnia największą część wariancji, druga mniejszą itd. Liczbę głównych składowych ustala się za pomocą odpowiedniego kryterium. Najważniejsze kryteria to [Gatnar, Walesiak, 2004, s. 208]: kryterium istotności testu statystycznego, kryterium bazujące na wartościach własnych, kryterium połowy, kryterium osypiska, kryterium wyjaśnionej wariancji i kryterium heurystyczne. W niniejszej pracy posłużono się kryterium wartości własnej, która zakłada, że składowe, które pozostawia się w dalszej analizie, to te, dla których wartości własne są większe niż 1 dla całkowitej macierzy korelacji. Kryterium to wskazuje na liczbę czynników wspólnych w macierzy korelacji.

Analiza czynnikowa stosowana jest do wyodrębniania liniowych kombinacji czynników wspólnych oraz specyficznych dla każdej wyjściowej zmiennej. Można zatem zapisać, że [Aczel, 1989, s. 951]:

$$\begin{aligned} X_1 &= b_{11}y_1 + b_{21}Y_2 + \ldots + b_{m1}Y_m + U_1 \\ X_2 &= b_{12}Y_1 + b_{22}Y_2 + \ldots + b_{m2}Y_m + U_2 \\ &\vdots \\ X_k &= b_{1k}Y_1 + b_{2k}Y_2 + \ldots + b_{mk}Y_m + U_k \end{aligned} \tag{2}$$

gdzie Y_i dla $i = 1, 2,\ldots,m$ stanowią czynniki wspólne, natomiast U_j dla $j = 1, 2,\ldots,k$ czynnik specyficzny dla zmiennej X_j, zaś parametry b_{ij} zwane są ładunkami czynników.

Do wyznaczenia wartości poszczególnych ładunków wykorzystano metodę centroidalną. Podobnie jak w przypadku metody głównych składowych, poszczególne czynniki wyjaśniają maksymalną część wariancji zmiennych wyjściowych. Pierwszy z czynników stanowi kombinację liniową zmiennych, wyjaśniającą maksymalną część całkowitej wariancji, drugi – maksymalną część, która pozostała po wyodrębnieniu pierwszego czynnika. Czynniki wspólne wyodrębniane są poprzez zsumowanie korelacji danej zmiennej z wszystkimi pozostałymi zmiennymi i podzieleniu jej przez sumę wszystkich współczynników z macierzy korelacji [Gatnar, Walesiak, 2004, s. 202]. Standardowo ładunki można uznać za znaczące, jeśli ich wartość bezwzględna przekracza 0,7. Główna różnica pomiędzy metodą głównych składowych a analizą czynnikową polega na tym, że pierwsza pozwala na przekształcenie zmiennych, druga na ich redukcję.

4. Wyniki badań

Wyniki niniejszych badań empirycznych stanowią próbę określenia czynników wpływających na strukturę kapitału w spółkach notowanych na GPW w latach 2000–2015. Każdy rok objęty badaniem oceniano oddzielnie, dzięki czemu określono, które czynniki miały istotny wpływ na dobór źródeł finasowania w poszczególnych latach.

Do badań przyjęto czynniki struktury kapitału, jakie najczęściej wskazywane są w literaturze przedmiotu, a mianowicie:

X1 – tempo wzrostu (GR), stopa wzrostu przychodów ze sprzedaży w odniesieniu do roku poprzedniego,

X2 – płynność finansowa (LIQ), relacja aktywów obrotowych do zobowiązań bieżących,

X3 – inwestycyjna tarcza podatkowa (NDTS), relacja amortyzacji i aktywów ogółem,

X4 – rentowność (ROE), relacja zysku netto do kapitału własnego ogółem,

X5 – wielkość (SIZE), stopa wzrostu aktywów ogółem w odniesieniu do roku poprzedniego,

X6 – struktura aktywów (TANG), relacja rzeczowych aktywów trwałych do aktywów ogółem.

Zgodnie z procedurą przedstawioną w części 3 niniejszej pracy, uwzględniając największe i różne co do wielkości wartości własne otrzymanych nowych parametrów, dla każdego roku wyodrębniono dwie składowe główne o największych wariancjach. Każda z tych składowych wyjaśnia określoną część wariancji całego układu, a zarazem odzwierciedla istotne

statystycznie czynniki kształtujące strukturę kapitału w badanych podmiotach (tab. 3).

Tablica 3. Główne składowe obliczone dla czynników struktury kapitału w spółkach notowanych na GPW w latach 2000–2015

Rok	Składowa	GR	LIQ	NDTS	ROE	TANG	SIZE	Skum. war. (%)
2015	1	0,5647	–0,0647	–0,3660	0,3720	0,2943	0,4910	28,78
	2	–0,1242	–0,3359	0,2546	–0,0732	0,4917	–0,3755	49,72
2014	1	0,5056	–0,0628	–0,1025	0,3242	0,3405	0,6482	28,08
	2	–0,2779	–0,1186	0,3968	–0,0099	**0,8755**	–0,3095	49,95
2013	1	0,2264	–0,1486	0,6581	0,0786	**0,7872**	0,5500	30,77
	2	0,1381	0,1209	–0,4376	0,1166	–0,5456	**0,8494**	52,54
2012	1	0,4674	0,4817	0,5047	–0,0665	0,4600	–0,3656	33,28
	2	–0,5628	–0,5115	0,6780	0,0936	**0,7156**	0,2257	58,06
	3	0,2967	0,2938	0,0062	0,3337	–0,0062	0,2568	76,02
2011	1	–0,0225	–0,5035	**0,7646**	0,5072	**0,7194**	0,0619	32,87
	2	0,2101	–0,4249	–0,4423	0,4377	–0,5094	0,2992	57,55
2010	1	**0,7678**	–0,3622	0,5848	0,0806	0,6342	0,1857	32,07
	2	–0,6528	0,0659	0,6682	–0,1270	0,6216	–0,4441	58,83
	3	0,0017	0,3796	–0,0186	0,2646	0,0186	0,1455	75,61
2009	1	–0,2157	–0,3600	0,6943	0,2970	**0,7221**	0,0705	31,39
	2	–0,0505	–0,1859	–0,5150	0,4166	–0,4824	0,3444	53,46
2008	1	0,1213	–0,3191	0,5146	0,1087	0,6569	0,3077	33,12
	2	0,6338	0,1935	–0,5025	0,1742	–0,5466	0,4345	54,62
2007	1	–0,3407	0,1637	0,4255	0,0046	0,6390	0,2479	34,63
	2	–0,1237	0,4593	–0,6136	0,0141	–0,6642	0,6808	53,58
	3	–0,3672	–0,2647	–0,1003	0,1375	0,1003	–0,2400	70,58
2006	1	0,0866	–0,1505	**0,7296**	–0,1595	**0,9831**	–0,1755	30,67
	2	–0,3136	–0,2561	–0,0071	–0,1076	0,2076	0,2505	50,16
2005	1	–0,2121	0,2537	0,6175	0,1623	**0,7151**	0,3039	30,42
	2	0,0564	–0,6470	0,5323	0,0386	0,5073	–0,3748	52,75
	3	–0,1001	–0,2138	–0,1381	–0,1336	0,2717	0,1137	69,60
2004	1	0,2315	0,4356	**0,7588**	0,1790	0,6313	–0,1609	32,26
	2	0,0039	**0,7358**	–0,6194	0,4975	–0,3920	0,3259	52,99
	3	–0,0436	–0,5558	0,1120	0,5993	0,0542	0,1662	69,88
2003	1	–0,1081	–0,2642	**0,7684**	–0,1127	0,6411	–0,2505	27,32
	2	–0,4102	–0,0968	0,1968	0,1910	–0,5521	–0,0393	46,98
	3	0,1110	0,1489	0,2809	0,2968	0,2968	0,0210	63,85

Rok	Składowa	GR	LIQ	NDTS	ROE	TANG	SIZE	Skum. war. (%)
2002	1	–0,4849	0,0053	**0,7959**	0,1802	0,4640	0,0913	28,60
	2	0,2130	0,0498	0,6675	0,1360	0,0303	–0,6707	49,71
2001	1	0,6222	0,2846	0,4400	–0,1439	0,5402	0,6281	30,76
	2	–0,2738	–0,0966	0,6573	0,2579	0,5527	–0,5817	58,15
2000	1	0,4011	**–0,7074**	0,4999	0,1782	0,5242	0,3907	31,05
	2	0,2639	–0,2158	**–0,7374**	0,0419	–0,4531	0,6688	56,40
	3	0,1082	–0,6846	0,0198	–0,1749	–0,0198	–0,6179	74,13

Źródło: Opracowanie własne.

Jak wynika z tablicy 3, w większości analizowanych okresów udało się wyodrębnić dwie główne składowe, w pozostałych trzy. W pierwszej składowej dominuje struktura aktywów (TANG) i nieodsetkowa tarcza podatkowa (NDTS). Pojawiły się one jako zmienne dominujące aż pięciokrotnie. Pierwsza składowa łącznie objaśniała dyspersję badanej zbiorowości przedsiębiorstw z umiarkowaną siłą, co odzwierciedla wynoszący około 30% udział w całkowitej wariancji. Oprócz powyższych zmiennych pierwsza składowa uwzględniła jednokrotnie wzrost (GR) w roku 2010 i płynność (LIQ) w roku 2000.

W ramach drugiej składowej, wyodrębnionej dla każdego roku, znaczącymi okazały się struktura aktywów (TANG) w latach 2012 i 2014, nieodsetkowa tarcza podatkowa (NDTS) w roku 2000, płynność (LIQ) w roku 2004 i wielkość (SIZE) w roku 2013. Składowe te wyjaśniły umiarkowane zróżnicowanie badanej zbiorowości, co znalazło odzwierciedlenie w ich skumulowanej wartości udziału w wariancji wynoszącej co najmniej około 50%.

Dzięki objęciu badaniem każdego roku z osobna możliwe było prześledzenie zmian, jakie zachodziły w sile oddziaływania badanych zmiennych na strukturę kapitału w zbiorowości spółek notowanych na GPW. Z analizy tej wynika, że siła ta była zróżnicowana w poszczególnych latach. W latach 2001, 2007, 2008 i 2015 otrzymane wyniki wskazują na brak znaczących czynników wpływających na strukturę kapitału. Z kolei w latach 2000, 2004, 2006, 2011 i 2013 w ustalonych postaciach składowych znalazły się jako dominujące po dwie z tych determinant.

Oceniając otrzymane wyniki z punktu widzenia kształtowania struktury kapitału, można zauważyć, że w badanym okresie kluczową rolę odgrywała nieodsetkowa tarcza podatkowa i struktura aktywów. Oznacza to, że wśród badanych przedsiębiorstw dominowały podmioty odnoszące korzyści podatkowe z tytułu amortyzacji, a nie kosztów finansowych.

Wniosek ten koresponduje z przedstawionymi w części 2 niniejszej pracy spostrzeżeniami odnośnie do wielkości zadłużenia badanych podmiotów. Spółki te charakteryzowały się bowiem stosunkowo niskimi wartościami wskaźnika zadłużenia w całym badanym okresie. Należy zwrócić uwagę, że zasoby środków trwałych pozostających w dyspozycji tych przedsiębiorstw stanowiły nie tylko podstawę odpisów amortyzacyjnych, ale i potencjalne zabezpieczenie zobowiązań, co sprzyjało ewentualnemu zadłużeniu. Wyjaśnia to istotną rolę struktury aktywów jako czynnika wpływającego na decyzje dotyczące doboru źródeł finansowania.

Jak już wspomniano, znaczenie płynności finansowej, jako czynnika kształtującego strukturę kapitału, było znaczące dwukrotnie: w roku 2000 w ramach pierwszej składowej i w roku 2004 w ramach drugiej składowej. W kolejnych latach czynnik ten w żadnym przypadku nie przyjął wartości powyżej 0,7 lub poniżej – 0,7. Wskazuje to na marginalizację znaczenia płynności finansowej w procesie kształtowania struktury kapitału. W tym samym kontekście znikoma jest także rola wzrostu i wielkości. Oba czynniki były znaczące zaledwie jeden raz, odpowiednio w ramach pierwszej i drugiej składowej. Jednakże w przypadku wielkości warto zwrócić uwagę na wysoki wynoszący blisko 53% poziom wyjaśnienia wariancji przez główne składowe w roku 2013.

Zakończenie

Z przeprowadzonych badań wynika, że średnia wartość wskaźnika ogólnego zadłużenia w badanym okresie ukształtowała się na poziomie 0,46. Oznacza to, że kapitał własny był głównym źródłem finansowania spółek notowanych na GPW w latach 2000–2015, a zatem struktura kapitału badanych podmiotów może zostać określona mianem konserwatywnej. Najważniejszymi czynnikami kształtującymi strukturę kapitału w badanych podmiotach okazały się struktura aktywów i nieodsetkowa tarcza podatkowa. Pierwsza z tych determinant istotnie oddziaływała na strukturę kapitału badanych przedsiębiorstw w pięciu przypadkach w ramach pierwszej składowej i w dwóch w ramach drugiej. Z kolei nieodsetkowa tarcza podatkowa okazała się dominująca pięć razy w przypadku pierwszej składowej i jeden raz – drugiej. Pozostałymi zmiennymi oddziałującymi na strukturę kapitału były płynność finansowa, wzrost i wielkość. Znaczenie tych determinant było jednak znikome. Siła oddziaływania badanych zmiennych na decyzje w zakresie kształtowania struktury kapitału w zbiorowości spółek notowanych na GPW była zróżnicowana w poszczególnych latach. W latach 2001, 2007, 2008 i 2015 nie odnotowano dominujących czynników wpływających na strukturę kapitału. Z kolei w latach 2000, 2004, 2006, 2011 i 2013 znaczącymi okazały się po dwie z tych

determinant. Podsumowując, hipoteza badawcza zakładająca, że ,,istnieją znaczące determinanty struktury kapitału, a siła ich oddziaływania zmienia się w czasie" została potwierdzona.

Literatura

Aczel A.D. (1989), *Complete Business Statistics*, Richard D. IRWIN, INC., Boston.

Barclay M.J., Smith C.W. (1996), *On financial architecture: leverage, maturity and priority*, ,,Journal of Applied Corporate Finance", Vol. 8, No. 4.

Barburski J. (2014), *Kapitały własne jako podstawa bezpieczeństwa działalności gospodarczej na przykładzie przedsiębiorstw WIG 20*, ,,Finanse, Rynki Finansowe, Ubezpieczenia", nr 67.

Bevan A.A., Danbolt J. (2002), *Capital structure and its determinants in the United Kingdom – A decompositional analysis*, ,, Applied Financial Economics", Vol. 12, No. 3.

Chaklader B., Chawla D. (2016), *A study of Determinants of Capital Structure through Panel Data Analysis of Firms Listed in NSE CNX 500*, ,,Vision", Vol. 20, No. 4.

DeAngelo H., Masulis R.W. (1980), *Optimal capital structure under corporate and personal taxation*, ,,Journal of Financial Economics", Vol. 8, No. 1.

De Jong A., Kabir R., Nguyen T.T. (2008), *Capital structure around the world: The roles of firm – and country – specific determinants*, ,,Journal of Banking & Finance", Vol. 32, No. 9.

Degryse H., de Goeij P., Kappert P. (2012), *The impact of firm and industry characteristics on small firms' capital structure*, ,,Small Business Economics", Vol. 38, No. 4.

Duliniec A. (2015), *Wybór źródeł finansowania a optymalna struktura kapitału w przedsiębiorstwie*, ,,Finanse, Rynki Finansowe, Ubezpieczenia", nr 74 (855).

Duliniec A. (1998), *Struktura i koszt kapitału w przedsiębiorstwie*, PWN, Warszawa.

Ferri M., Jones W. (1979), *Determinants of financial structure: A new methodological approach*, ,,Journal of Finance", Vol. 34, No. 3.

Gabrusewicz W. (2014), *Analiza finansowa przedsiębiorstwa. Teoria i zastosowanie*, PWE, Warszawa.

Gatnar E., Walesiak M. (2004), *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.

Janus A. (2006), *Kapitał własny jako źródło finasowania działalności małych i średnich przedsiębiorstw*, ,,Folia Oeconomica", nr 200.

Jerzemowska M. (2013), *Kształtowanie struktury kapitału w przedsiębiorstwie*, w: *Analiza ekonomiczna w przedsiębiorstwie*, Jerzemowska M. (red.), PWE, Warszawa.

Jordan J., Lowe J., Taylor P. (1998), *Strategy and financial policy in UK small firms*, ,,Journal of Business Finance & Accounting", Vol. 25, No. 1–2.

Komera S., Lukose J. (2015), *Capital structure choice, information asymmetry, and debt capacity: evidence from India*, ,,Journal of Economics and Finance", Vol. 39, No. 4.

Konečný Z. (2013), *Golden Rules of Financing Related to the Life Cycle of Czech Automotive Firms*, ,,Journal of Competitiveness", Vol. 5, No. 2.

Lisińska K. (2012), *Struktura kapitałowa przedsiębiorstw produkcyjnych w Polsce, Niemczech i Portugalii*, ,,Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", nr 27.

Marzec J. (2010), *Złote reguły finansowania w praktyce małych i średnich przedsiębiorstw w Polsce*, ,,Ekonomiczne Problemy Usług", nr 51.

Mazur K. (2007), *The Determinants of Capital Structure Choice: Evidence from Polish Companies*, ,,International Advances in Economic Research", Vol. 13, No. 4.

Modigliani F., Miller M.H. (1963), *Corporate Income Taxes and the Cost of Capital: A Correction*, ,,American Economic Review", Vol. 53, No. 3.

Myers S.C., Majluf N.S. (1984), *Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have*, ,,Journal of Financial Economics", Vol. 13, No. 2.

Serrasqueiro Z., Matias F., Salsa L. (2016), *Determinants of capital structure: New evidence from Portuguese small firms*, ,,Dos Algarves: A Multidisciplinary e-Journal", Vol. 28.

Rajan R. G., Zingales L. (1995), *What do we know about capital structure? Some evidence from international data*, ,,Journal of Finance", Vol. 50, No. 5.

Titman S.,Wessels R. (1988), *The determinants of capital structure choice*, ,,Journal of Finance", Vol. 43, No. 1.

Wrońska-Bukalska E. (2014), *Zróżnicowanie branżowe poziomu i struktury kapitału własnego*, ,,Annales Universitatis Mariae Curie-Skłodowska", nr XLVIII.

Streszczenie

Celem artykułu jest empiryczna weryfikacja czynników kształtujących strukturę kapitału spółek notowanych na GPW w Warszawie w latach 2000–2015. Zakłada się, że istnieją znaczące determinanty struktury kapitału, a siła ich oddziaływania zmienia się w czasie. Realizacja tak sformułowanego celu badań i weryfikacja hipotezy badawczej wymagała zastosowania analizy głównych składowych. Obliczenia wykonano za pomocą programu Statistica. Jako zmienne diagnostyczne przyjęto: tempo wzrostu, płynność finansową, inwestycyjną tarczę podatkową, rentowność, wielkość i strukturę aktywów. Najważniejszymi czynnikami kształtującymi strukturę kapitału w badanych podmiotach okazały się struktura aktywów i nieodsetkowa tarcza podatkowa. Pozostałymi zmiennymi oddziałującymi na strukturę kapitału były płynność finansowa, wzrost i wielkość. Znaczenie tych determinant było jednak znikome. Siła oddziaływania badanych zmiennych na decyzje w zakresie pozyskiwania kapitału w zbiorowości spółek notowanych na GPW była zróżnicowana w poszczególnych latach.

Słowa kluczowe

nieodsetkowa tarcza podatkowa, struktura aktywów, wskaźnik zadłużenia, dźwignia finansowa, analiza składowych

The influence of selected factors of capital structure of companies listed on the Warsaw Stock (Summary)

The aim of the article is an empirical verification of factors shaping the capital structure of companies listed on the Warsaw Stock Exchange in 2000–2015. It is assumed that there are relevant determinants of the capital structure, and the strength of their influence changes over time. The hypothesis was verified using principal components method. Calculations were made with help of the Statistica software. The diagnostic variables were: growth rate, financial liquidity, non-debt (investment) tax shields, profitability, size and asset structure. The most important factors shaping the capital structure of the analysed companies were the asset structure and the non-debt tax shields. The other variables shaping the capital structure were liquidity, growth and size. The significance of these determinants was, however, minimal. The impact of the analysed variables on capital structure of companies listed on WSE varied in different years.

Keywords

non-debt tax shields, asset structure, debt ratio, leverage, principal components method

Błażej Kochański*

20+ years of Polish residential real estate prices – a house price index proxy

Introduction

Trends in house prices are followed by governments and private sector analysts as well as households for various reasons, including, inter alia, economic growth assessment, shaping monetary policy, estimation of wealth. In many economies house price indices have been available for many years. In Poland, the official index recently introduced by the Central Statistical Office (GUS) goes back in time only to 2010.

The aim of this paper is to propose a house price index proxy going back further to history – the quarterly series of the proposed index goes back to 2006 and the series with yearly frequency goes back to 1995. In order to attain this goal various data sources are gathered – some of them are unfortunately fragmented and methodology/assumptions behind them are not always clear. However, as they show quite similar trends, it is possible to use them to propose an index (or, at least, a "proxy" of an index), which may be improved when additional data are found or collected.

In the first section of the paper current practices in the area of house prices measurement are shortly discussed; in the second section current indices for Poland are presented. Next, the available time series on Polish residential property prices are described. Fourth section contains a proposal for the house price index proxy and shows that it fits well to the "cloud" of available data. Fifth section discusses two "false trails", which may be mistakenly taken as time series representing trends in Polish residential real estate prices.

1. Residential property price indices – current practices

Investors, financial institutions, governments, regulatory bodies as well as households and academic analysts are interested in house prices changes in time. Residential real estate price indices are constructed all around the world as an aid in tracking the trends of house prices. Such indices have many uses including economic growth assessment, aid in shaping monetary policy/financial stability policy, household wealth estimation,

* PhD, Department of Economic Science, Faculty of Management and Economics, Gdańsk University of Technology, ul. Gabriela Narutowicza 11/12, 80-233 Gdańsk, bkochanski@zie.pg.gda.pl

deflator in the national accounts, inter-area and international comparisons. [*Handbook*, 2013, p. 16]. Individual citizens look at the indices when making decisions about buying/selling houses and flats or taking mortgages. Banks use such indices to assess portfolio credit risk and perform stress tests [Schuermann, 2014]. It is difficult to make optimal economic decisions regarding real estate and household budget management and it is impossible to understand real estate and credit market without summarized information on tendencies in house prices.

Variety of approaches and data sources are used to construct residential price indices. Data on house prices come from brokers, registers, newspapers, real estate agents, online intermediaries, mortgage lenders, lawyers, notaries or land registers. The indices may be based on transaction price, but also on initial or final asking price, the offer price or an expert valuation. Various methods are used to construct a summarising information: starting from simple measures of central tendency (mean or median) with stratification/mix adjustment methods through hedonic regression methods, repeat sales methods (with a famous Case-Shiller index as a most prominent example) and methods based on appraisal data. Also the index may cover combination of land, flats or houses, newly built or existing, stock of residential real estate or stream of houses/flats being sold. An informative summary on all of above aspects is provided by Eurostat [*Handbook*, 2013].

2. Indices currently available in Poland

Currently in Poland there are two main residential indices publicly available:

- a Central Statistical Office (GUS) House Price Index (HPI) covering the whole territory of Poland, published first in 2015, going back to 2010,
- NBP hedonic index, covering big cities whose data go back to the third quarter of 2006.

In the case of Polish Central Statistical Office index, the data come from administrative source: the Register of Real Estates Prices and Values established in all 350+ powiats (counties) in Poland. The methodology and details are described by R. Rechnio [2016]. The calculations of the index are based on the information on flats (single family houses are excluded). Data are compiled quarterly on the basis of transaction prices as specified in the notarial deeds, covers both newly built houses as well as transactions between households. The transaction prices include the value of land. Only market prices are taken into account, non-market transactions are excluded from the calculation. GUS obtains its values for the country as a whole, NUTS 2 regions and separately for the biggest cities. The new

and existing flat indices are available along with the total index. The HPI is a chain-linked Laspeyres-type price index; in order to control for changes in the mix, the stratification method is used. The choice of this method was dictated by the availability of data – the lack of sufficiently detailed description of properties made hedonic approach impossible.

Interestingly, the HPI is not part of standard set of macroeconomic variables available at GUS webpage at the moment. The most reliable and up-do-date source for this index is Eurostat database [http://appsso.eurostat.ec.europa.eu/nui, accessed: 22.09.2017].

NBP's hedonic index is based on hedonic methodology and is based on the BaRN database created by the National Bank of Poland. The data on transaction prices from the secondary housing market are gathered on voluntary basis from intermediaries and developers. The BaRN database covers both existing and new dwellings, and stores two kinds of prices (offer and transactional) for 17 cities. There are 80+ data series showing average prices per square metre available at the NBP's webpage [http://www.nbp.pl/home.aspx?f=/publikacje/rynek, accessed: 22.09.2017]. The house price index based on transaction prices was built with use of hedonic regression for the seven biggest cities [Widłak, 2013].

3. Poland. Going back further with data

The officially quoted Polish HPI from the Central Statistical Office (GUS) goes back only to 2010. It means it covers neither the outbreak of the global financial crisis (2008–2009) nor the time of housing market boom in Poland (2005–2007). NBP data start mid-2006 (NBP), so the the boom period is covered only partially and the preceding period is not covered at all. This creates problems for some uses of the indices, when time series required need to be longer. For example the series are too short to use them as a deflator for the prices in longer term econometric analysis or to apply them in a bank's portfolio stress testing (which would need the boom, the preceding period and the crisis data).

Analysts or investors in need for longer time series may turn to some other data sources. Those sources are, however, quite limited, partial (cover just several years), the methodology behind them is not always clear, the underlying data and assumptions are not disclosed.

Table 1 summarizes information on available data on house prices including the frequency of the data (data points may be monthly, quarterly and yearly) and period covered; several other aspects are discussed in the remarks section.

AMRON/SARFiN report is the report prepared in the Centre of Banking Law and Information created by Polish Banks Association. This quarterly

report contains information on average transaction prices per square metre in several cities (including Warszawa, Białystok, Katowice, Wrocław, Gdańsk, Kraków, Poznań, Łódź). For further calculations the index based on the simple average of these time series will be used.

Expander is a financial intermediary and mortgage broker company in Poland. Regular report on house prices prepared by Expander in cooperation with szybko.pl real estate listing website and Metrohouse real estate agency has been regularly published on the Web. Data frequency is monthly. The report contains average offer prices in several cities based on mortgage and real estate intermediaries data going back to 2005. Also, transaction prices are available starting in 2011.

Before official introduction of official Polish House Price Index, information on Polish house prices was published on Eurostat webpage as part of the framework of "alert mechanism" within macroeconomic imbalance procedures of EU. The data are not available online anymore, but it was available at European Commission page [ec.europa.eu/economy_finance/indicators/economic_reforms, accessed: 9.09.2013] and kept by the author. The data was available with yearly frequency for years 2005–2012.

Unfortunately, none of the sources enumerated above covers the period before 2005. The time series going back before that year are much more fragmented and less documented. For example, M. Żeliński [2007] prepared, based on the data available to him from Reas and Rednet consulting companies, a time series showing average per square metre house price in six biggest Polish cities and a resulting weighted price. Another such example is a report prepared by a representative of a real estate appraisal agency on taksator.pl [Zarzycki, 2007]. It contains average transaction price per square metre from Wrocław area calculated based on the data from commercial database built based on collected data from notaries. AD Drągowski, a real estate agency, was also trying to fill the gap resulting from lack of house price index. In the report published on their webpages in the past [Drągowski, 2007], they shared their indirect "index", based on information about average intermediary commission in real estate transactions. The geographical area covered is unclear, most probably the data come mainly from Warsaw.

Another source covering period before 2007 was the data presented in NBP's publication [Łaszek et al., 2009], where one of the charts presented average market price per square metre in Warsaw. According to the authors the series was "based on GUS, NBP and PONT Info data".

Table 1. Sources of data on Polish house prices trends

Data source	Data frequency	Period covered	Remarks
GUS (Central Statistical Office)	quarterly	since 2010	Stratification method, based on government real estate prices register.
NBP (National Bank of Poland)	quarterly	since 3rd quarter of 2006	Hedonic index based on the secondary market data (7 cities).
AMRON-SARFiN (report by the Polish Banks Association)	quarterly	since 1st quarter of 2006	Average transaction prices in several Polish cities based mainly on Polish interbank data.
Expander offer prices (Expander/ szybko.pl/ Metrohouse report on real estate prices)	monthly	since 2005	Average offer prices in several cities based on mortgage and real estate intermediaries data.
Expander transaction prices	quarterly	since 2011	Average transaction prices in several cities based on mortgage and real estate intermediaries data.
Eurostat – experimental index, data were available as SB_A_NOMHOUSE variable in an "Alert Mechanism Framework"	yearly	2005–2012	No explanation regarding methodology, not available online at the moment.
Reas/RedNet – data collected from consulting companies reports by M. Żeliński	yearly	2000–2007	Data on average prices of square metre in newly built flats.
Taksator.pl – analysis prepared by real estate appraisal agency representative – D. Zarzycki	quarterly	1st quarter 2004 – 2nd quarter 2007	Average transaction prices in Wrocław agglomeration from commercial register of notary data.
AD Drągowski – real estate price index prepared by real estate agency	yearly	1992–2007	Index based on information on average intermediary commission in real estate transactions, probably mainly in Warsaw.
Research paper by Łaszek/ Widłak/ Augustyniak – Average prices in Warsaw [Łaszek et al., 2008, p. 39]	yearly	1995–2007	Average market price per square metre in Warsaw ("based on GUS, NBP and PONT Info data"), no further details available.

Source: Own elaboration.

4. House Price Index proxy construction

The data described in previous section and in Table 1 may be used to derive a house price index proxy. Assumptions behind the proposed proxy are following:

- starting point would be existing House Price Index presented by Eurostat/GUS for period starting in the first quarter of 2010, this also means that index=100 would represent average prices in 2010[1]
- available information from other sources will be used in order to go back in history as far as to 1995.

Such an index would have to be called a "proxy" as it would be based on fragmented and not fully compatible data.

First step to build an extension of existing GUS index is possible thanks to the hedonic index proposed by NBP. In order to rescale the latter to fit the former, linear regression with zero intercept is used (period 2010–2016 was used). This is practically equivalent to shifting the NBP index so that it fits average GUS index. Thanks to this move the quarterly GUS index can be extended back in time to cover the period starting in the third quarter of 2006. Similar approach may be used to further augment the index proxy back to the beginning of 2006 using AMRON/SARFiN data. With those two uncomplicated steps, we are able to arrive at an index proxy of quarterly frequency covering period of 11 years. The results of such procedure are shown in Table 2.

Table 2. HPI and HPI proxy – quarterly data

Quarter	Index	Quarter	Index	Quarter	Index
2006Q1	62,92**	2010Q1	99,05	2014Q1	91,99
2006Q2	66,29**	2010Q2	100,25	2014Q2	94,02
2006Q3	69,21*	2010Q3	100,39	2014Q3	93,42
2006Q4	81,93*	2010Q4	100,32	2014Q4	93,63
2007Q1	91,80*	2011Q1	100,25	2015Q1	93,79
2007Q2	104,75*	2011Q2	100,90	2015Q2	95,10
2007Q3	109,82*	2011Q3	99,80	2015Q3	95,28
2007Q4	108,75*	2011Q4	99,34	2015Q4	94,57
2008Q1	107,05*	2012Q1	98,00	2016Q1	94,69
2008Q2	106,45*	2012Q2	97,35	2016Q2	95,48
2008Q3	104,35*	2012Q3	96,02	2016Q3	97,27
2008Q4	105,13*	2012Q4	94,96	2016Q4	98,34
2009Q1	99,75*	2013Q1	92,51		

[1] Please note, that Eurostat changed the basis of their index recently, now 2015 being the base (index=100) year. The proxy index may be rescaled accordingly.

Quarter	Index	Quarter	Index	Quarter	Index
2009Q2	98,18*	2013Q2	92,33		
2009Q3	99,31*	2013Q3	92,26		
2009Q4	99,73*	2013Q4	92,31		

Legend: *proxy based on extrapolation using regression based on NBP index **proxy based on further extrapolation based on AMRON data
Source: Own elaboration.

Unfortunately, the data available for the years before 2006 are in most cases of yearly frequency, that is why it is not possible to provide a quarterly proxy. The proxy for years 1995–2005 will be thus available only with yearly frequency. The proxy was constructed as follows: 1) yearly time series were produced from quarterly proxy obtained so far, 2) available time series described in Table 1 were adjusted to produce a fixed-base indices, 3) individual fixed-base indices were shifted so that their 2007 value is equal in each case to the value of the proxy index for that year, 4) a simple average was taken from all indices which were available for the given year.

The proxy calculated in such a process is presented in Table 3.

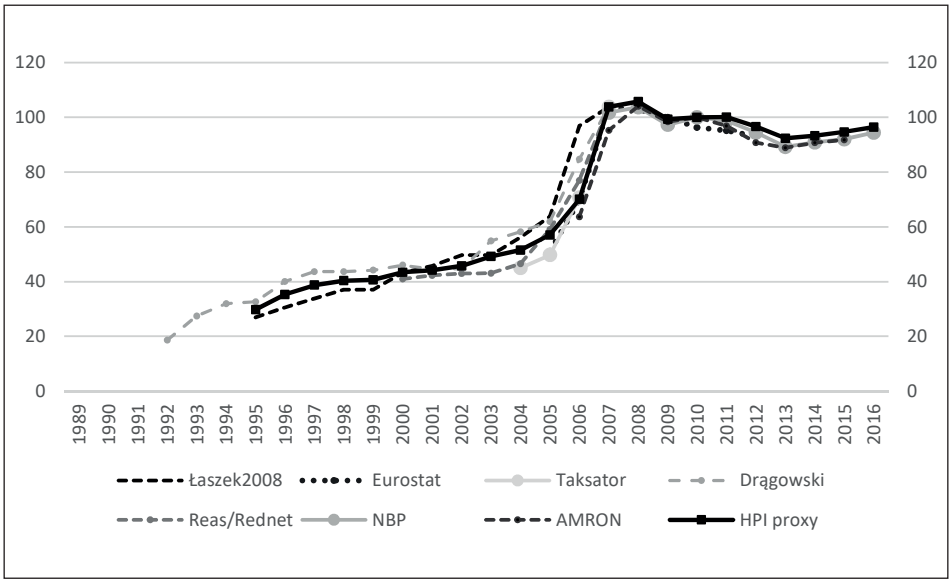
Table 3. HPI and HPI proxy – yearly data

Year	Index	Year	Index	Year	Index
1995	29,80**	2003	49,23**	2011	100,07
1996	35,33**	2004	51,49**	2012	96,58
1997	38,73**	2005	57,10**	2013	92,35
1998	40,35**	2006	70,09*	2014	93,27
1999	40,61**	2007	103,78*	2015	94,69
2000	43,38**	2008	105,75*	2016	96,45
2001	44,17**	2009	99,24*		
2002	45,76**	2010	100,00		

Legend: * proxy based on extrapolation using regression based on NBP and AMRON index
** proxy based on further extrapolation based on other data from Table 1.
Source: Own elaboration.

Figure 1 presents a proposed yearly HPI index proxy plotted against source data. It should be noted that all the data on the chart have been translated to the form of fixed-based indices and rescaled adequately. As it can be intuitively read from the chart, the proposed index proxy fits quite well into the "cloud" of available data.

Figure 1. Proposed HPI proxy against source indices



Source: Own elaboration.

5. Polish residential real estate data – "false trails"

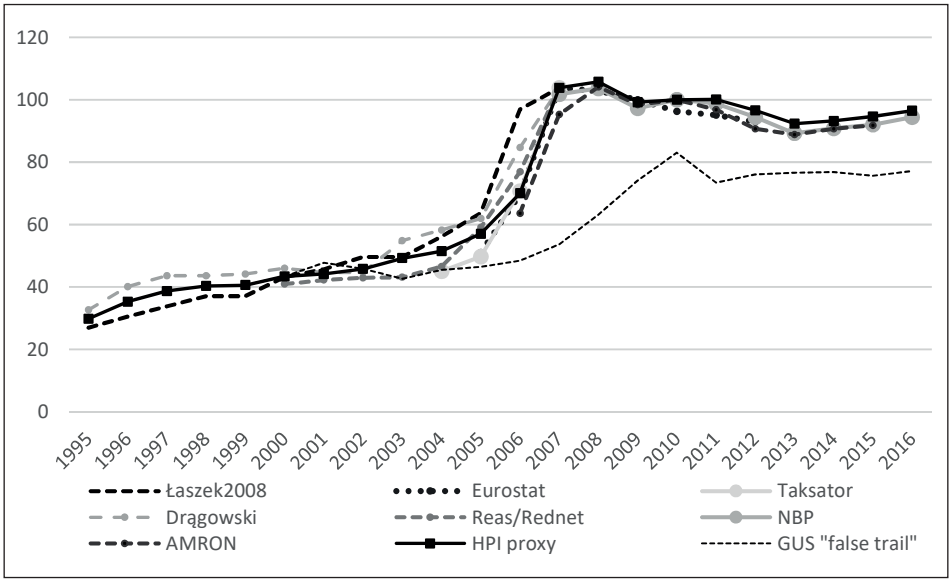
An analyst or researcher looking for data on Polish house prices can easily stumble on data that resembles a house price index, but is inconsistent with majority of other sources. Indices based on such data will be called a "false trail" for the purposes of this paper.

One of such "false trails" is the information on "the price of a square meter of usable floor space of a residential building" available on webpages of the Central Statistical Office (GUS). This times series seems quite appealing: the name suggests it could be treated to derive a simple house price index and – even more importantly – the series goes back in time as far as to the fourth quarter of 1998.

Unfortunately, GUS themselves warn that this series "cannot be used as a price of square metre of usable floor space of residential buildings completed or for analysis such prices offered in housing market" – it cannot constitute a basis for a house price index. As it is claimed on their webpage the data should be understood as "as average outlays for the whole country incurred by investors for the construction of multi-dwelling buildings per 1 square metre of the usable floor space of a building". This means that the time series, despite the name under which it appears on the official statistical internet site, does not represent the prices, but rather the construction expenses.

As it can be seen on Figure 2, the GUS construction cost follows quite a different path then all the other data on house prices. This should also reinforce the claim that it cannot be used for a house price analysis. Still, some researchers may mistakenly use this data as Polish house prices indicator, an example could be Favilukis et al. [2013].

Figure 2. Illustration of "false trail 1" – GUS construction cost



Source: Own elaboration.

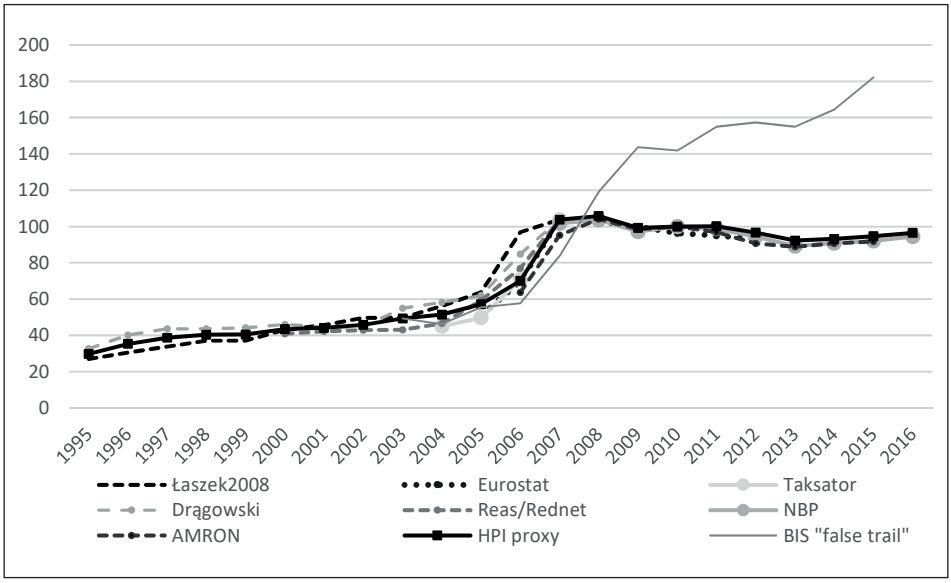
Another notable example of "false trails" are some of time series presented at the webpage of the Bank of International Settlements from Basel. As part of their statistics gathering effort, the Bank collates data on house prices for various countries. The database contains more than 300 price series for more than 50 countries [Scatigna et al., 2014]. In case of Poland, it reports – inter alia – the data on house prices and the house price indices from GUS and NBP. However, there are some data series reported for Poland in the data base whose source is not exactly clear. This is the case for the time series with yearly frequency – there are six of them, allegedly coming from the Central Statistical Office, but the link given in the database results in an error and suggests a document from 2008 while the latest data point is from 2015.

The series covers single-family houses and flats for the whole country, urban and rural areas. As a note in the database says, "a series on prices of residential buildings are not very representative, as the market of residential buildings is still quite narrow". Indeed, the series on houses as well

as the one on flats on rural areas, are quite erratic. But the two series on flats (for the urban areas and for the whole country) seem consistent (and similar). However, the confrontation of any of them with the data gathered from other sources show they are incompatible – so this is, most probably, another "false trail". Figure 3 illustrates the problem. The dynamic growth of years 2005–2007 seems to be delayed in these data (which may be explained, as a note in the database says, by a lag in registration in notary acts for newly built dwellings). More worryingly, the growth continues after 2009 when other time series flatten out. The reason may be the mix of reported transactions.

As the BIS database seems to be widely used for the real estate price analysis there is a risk that "false trails" from this database are also mistakenly taken as correctly representing the trends in the Polish market. As in the previous case, using them may seem appealing: the time series starts in 2003, several years earlier than GUS/NBP indices.

Figure 3. Illustration of "false trail 2" – BIS database time series



Source: Own elaboration.

Conclusion

The index proposed in this article may be useful for analysts looking for time series on Polish real estate prices going back in time further than 2006–2010.

The proxy index may be improved when new data on house price history, especially from before 2006, become available. It is possible that

institutions that internally store and process data on house prices (banks, notary registers, other) will be willing to share their data with the public. Also, systematic effort of gathering the data dispersed in paper sources like notary acts or banking transactions documentation could also help develop better basis for the reliable index.

The derivation methodology of this proxy index is not sophisticated and one may have legitimate objections regarding the data sources and their compatibility. However, the proxy seems to be better than its alternatives. In case of lack of such index an analyst or researcher has an option to abandon the research idea or exclude Poland from the research (which does not seem to be a constructive solution) or follow above mentioned "false trails" (which, in turn, may result in unreliable results and wrong conclusions).

References

Detailed data set on nominal residential property prices, Bank of International Settlements, www.bis.org/statistics/pp/pp_detailed.xlsx, accessed: 22.09.2017.

Drągowski L. (2009), *Wskaźnik zmian cen na rynku nieruchomości*, http://www.dragowski.pl/strona/wskaznik_zmian_cen, accessed: 28.11.2012.

ec.europa.eu/economy_finance/indicators/economic_reforms/api/data.cfm?application_name=mip&timeSeries=&group=SBA&country=&year=&format=excel&FC=1&L=1&LY=1&x=0&v=L, accessed: 9.09.2013.

Favilukis J., Kohn D., Ludvigson S.C., Van Nieuwerburgh S. (2013), *International capital flows and house prices: theory and evidence*, in: E.L. Glaeser, T. Sinai (eds.), *Housing and the Financial Crisis*, National Bureau of Economic Research, University of Chicago Press.

http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=prc_hpi_q, accessed: 22.09.2017.

http://www.nbp.pl/home.aspx?f=/publikacje/rynek_nieruchomosci/index2.html, accessed: 22.09.2017.

Handbook on Residential Property Prices Indices (RPPIs) (2013), Eurostat Methodologies & Working papers, http://ec.europa.eu/eurostat/documents/3859598/5925925/KS-RA-12-022-EN.PDF, accessed: 22.09.2017.

Łaszek J. (2004), *Sektor nieruchomości mieszkaniowych w Polsce: stan i perspektywy rozwoju*, Szkoła Główna Handlowa, Warszawa.

Łaszek J., Augustyniak H., Widłak M. (2009), *Euro a ryzyko bąbli na rynku nieruchomości mieszkaniowych*, „Materiały i Studia", nr 238, Warszawa.

Rechnio R. (2016), *The Polish experience in developing House Price Index*, Meeting of the Group of Experts on Consumer Price Indices, Geneva, http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2016/Session_5_Poland_The_Polish_experience_in_developing_HPI.pdf, accessed: 22.09.2017.

Scatigna M., Szemere R., Tsatsaronis K. (2014), *Residential property price statistics across the globe*, "BIS Quarterly Review", September 2014.

Schuermann T. (2014), *Stress testing banks*, "International Journal of Forecasting", Vol. 30.

Widłak M. (2013), *The hedonic house price index for Poland – modelling on NBP BaRN data*, Narodowy Bank Polski International Workshop, Zalesie Górne, https://www.nbp.pl/badania/konferencje/2013/realestateworkshop/pdf/hedonic_house_price_index_for_poland.pdf, accessed: 22.09.2017.

Zarzycki D. (2007), *Analiza rynku wtórnego mieszkań na terenie miasta Wrocław – styczeń 2004 r. – wrzesień 2007 r.*, http://taksator.pl/8-ze-starej-strony/41-id49.html, accessed: 22.09.2017.

Żeliński M. (2007), *Cena metra kwadratowego mieszkania wg GUS*, http://opinieekonomiczne.blox.pl/2007/11/Cena-metra-kwadratowego-mieszkania-wg-gus.html, accessed: 22.09.2017.

20+ years of Polish residential real estate prices – a house price index proxy (Summary)

Information on house prices is often considered crucial when assessing developments in economy. However, in Poland for the long time house prices haven't been tracked in an organized manner.

Recent developments, most notably house price indices developed by Poland's national bank and by the statistical office do not go back more than 10 years, therefore data series are relatively short compared to those of other macroeconomic data.

In the paper a collection of publicly available data on Polish house prices from various sources is presented. House price index estimate is then proposed which goes back to 1995 and covers periods of housing boom and the financial crisis. The index should be treated as a proxy due to the fact that the data availability is limited and the source data are based on various, sometimes not clearly articulated, methodologies. Nevertheless this proxy index may prove useful as – to the knowledge of the author – there has been no other publicly available attempt at reconstructing the latest history of residential real-estate prices in Poland. Furthermore, some researchers used to take inadequate data, like residential construction cost, as a proxy for house prices, which may have led to wrong conclusions.

The index may be further improved when new data emerge on history of house prices before 2006 and more advanced methodologies are used.

Keywords
house price index, Poland, residential real estate prices

Andrzej Karpio*
Dorota Żebrowska-Suchodolska**

Analiza porównawcza aktywnych i pasywnych strategii inwestycyjnych otwartych funduszy emerytalnych i otwartych funduszy zrównoważonych w latach 2014–2017

Wstęp

Podstawą oceny efektywności inwestycyjnej instytucji zbiorowego inwestowania najczęściej są standardowe miary w rodzaju współczynników Treynora, Sharpe'a, Jensena i inne, wraz z ich modyfikacjami lub uogólnieniami (Calmar, Omega itp.) [Perez, 2012]. Biorą one pod uwagę zarządzanie portfelami inwestycyjnymi w kontekście różnych form ryzyka, ale też różnych definicji benchmarków. Z wyjątkiem miary wykorzystującej współczynnik alfa, w rodzaju wskaźnika Jensena i jego modyfikacji, miary te dokonują oceny biernego stylu zarządzania. Aktywność zarządzających można badać, w ograniczonym zakresie, wyznaczając na przykład persystencję pozycji rankingowych utworzonych na podstawie powyższych miar w różnych okresach czasu, często odnoszących się do zmian koniunktury rynkowej. Z punktu widzenia klientów (uczestników) funduszy zbiorowego inwestowania mogą oni oczekiwać dużo więcej niż jedynie biernego „naśladowania" rynku. Wnosząc opłaty za zarządzanie, naliczane od aktywów, którymi klienci obciążani są wielokrotnie przez cały czas pozostawania w funduszu, mogą wymagać aktywnego podejścia zarządzających do inwestowania i trafnego przewidywania zachowania się zarówno rynku, jak i poszczególnych dostępnych instrumentów finansowych. Narzędzi do badania tego aspektu funkcjonowania funduszy zbiorowego inwestowania dostarczają modele Treynora-Mazuy'ego [Treynor, Mazuy, 1966] oraz Henrikssona-Mertona [Henriksson, Merton, 1981]. Umożliwiają one identyfikowanie umiejętności wykorzystania zmian koniunktury rynkowej w celu poprawy uzyskiwanej stopy zwrotu z inwestycji. Zaproponowane modele dokonują podziału stopy zwrotu pomiędzy nadwyżkę związaną z wykorzystaniem trendów rynkowych i umiejętności doboru papierów

* Dr, Katedra Ekonometrii i Statystyki, Wydział Zastosowań Informatyki i Matematyki, SGGW w Warszawie, ul. Nowoursynowska 159, 02-787 Warszawa, andrzej_karpio@sggw.pl
** Dr, Katedra Ekonometrii i Statystyki, Wydział Zastosowań Informatyki i Matematyki, SGGW w Warszawie, ul. Nowoursynowska 159, 02-787 Warszawa, zdorota@o2.pl

wartościowych do portfeli. W pierwszym przypadku miarami są odpowiednie współczynniki strukturalne w równaniach opisujących modele. Natomiast umiejętność selektywnego doboru aktywów może być mierzona wskaźnikami odpowiednio: Treynora-Mazuy'ego i Henrikssona-Mertona. Aby odseparować te dwie kategorie od siebie, należy oszacować model metodą najmniejszych kwadratów i zbadać estymatory odpowiednich współczynników strukturalnych.

Ocena efektywności instytucji zbiorowego inwestowania może dotyczyć aktywnej polityki lub pasywnej. Dlatego przedmiotem prezentowanej analizy jest próba odpowiedzi na pytanie, czy zarządzający funduszami emerytalnymi i wybranymi funduszami zrównoważonymi akcji lepiej radzą sobie z aktywnym zarządzaniem czy pasywnym. Pomiar polityki aktywnej oparty jest na wspomnianych modelach Treynora-Mazuy'ego i Henrikssona-Mertona, natomiast polityki pasywnej na współczynnikach beta w liniach charakterystycznych funduszy. W obu przypadkach wykorzystuje się benchmark odnoszący się do składów portfeli, którego konstrukcja jest zaprezentowana w niniejszej pracy.

Po zmianie przepisów dotyczących funkcjonowania funduszy emerytalnych, z punktu widzenia struktury portfeli, zaczęły one przypominać otwarte fundusze zrównoważonego wzrostu, dlatego ich efektywność porównuje się właśnie z tymi podmiotami. Wcześniej analogiczne badania dotyczyły funduszy stabilnego wzrostu, patrz na przykład [Karpio, Żebrowska-Suchodolska, 2014].

1. Konstrukcja benchmarku i założenia metodyczne

W badaniach uwzględniono okres funkcjonowania funduszy po zmianie zasad prawnych dotyczących składu portfeli funduszy emerytalnych, zatem jest to okres od 1 lutego 2014 r. do 31 stycznia 2017 r. Analizę porównawczą efektywności funduszy emerytalnych przed zmianami przepisów i po można znaleźć na przykład w pracy [Karpio, Żebrowska-Suchodolska, 2016]. W niniejszym opracowaniu pod uwagę brano miesięczne i tygodniowe zwykłe stopy zmian jednostek rozrachunkowych i jednostek uczestnictwa. Na rynku funduszy emerytalnych dochodziło do przejęć, zatem w badaniach uwzględniono jedynie podmiot przejmujący. Kryterium doboru funduszy zrównoważonych brało pod uwagę funkcjonowanie w całym okresie, w konsekwencji podmioty uwzględnione w badaniach zebrano w tablicy 1 zawierającej 12 funduszy emerytalnych i 13 zrównoważonych.

Tablica 1. Lista funduszy emerytalnych i funduszy zrównoważonych wybranych do badania

Fundusze emerytalne	Skrócone nazwy	Fundusze zrównoważone akcji	Skrócone nazwy
Aegon OFE	Aegon	Arka BZ WBK Zrównoważony	Arka
Allianz Polska OFE	Alianz	Aviva Investors Zrównoważony	Aviva
Aviva OFE	Aviva	BPH Zrównoważony	BPH
AXA OFE	Axa	Credit Agricole Dynamiczny Polski	Credit Agricole
Generali OFE	Generali	Investor Zrównoważony	Investor
MetLife OFE	Metlife	KBC Aktywny	KBC
Nationale-Nederlanden OFE	Nationale	NN Zrównoważony	Nationale
Nordea OFE	Nordea	Novo Zrównoważonego Wzrostu	Novo
OFE Pocztylion	Pocztylion	Pioneer Zrównoważony	Pioneer
OFE PZU "Złota Jesień"	PZU	PKO Zrównoważony	PKO
Pekao OFE	Pekao	PZU Zrównoważony	PZU
PKO BP Bankowy OFE	Bankowy	Skarbiec Waga Zrównoważony	Skarbiec
		UniKorona Zrównoważony	UniKorona

Źródło: Opracowanie własne.

W dalszej części artykułu używa się skróconych nazw funduszy w celu bardziej przejrzystego przedstawiania wyników.

Jak można zorientować się ze stron internetowych funduszy emerytalnych, po zmianie zasad prawnych udział akcji notowanych na rynku regulowanym na ogół przewyższał 75%. Pozostały kapitał lokowany był w wiele różnych instrumentów, w tym w obligacje komunalne lub korporacyjne, natomiast część w instrumenty płynne. Ze względu na brak powszechnie wykorzystywanego indeksu opisującego pozaakcyjne aktywa jako benchmark przyjęto WIBOR 1W, gdy obliczenia dokonywane były na tygodniowych zmianach jednostek udziałowych i WIBOR 1M, w przypadku zmian miesięcznych. Wspomniany wcześniej udział akcji w portfelach pozwala przyjąć, że procentowa zmiana czynnika rynkowego opisuje następująca kombinacja liniowa:

$$r_{M,t} = 0{,}8r_{WIG,t} + 0{,}2r_{WIBOR,t} \quad (1)$$

gdzie t jest zmienną czasową, odpowiednio równą kolejnym tygodniom lub miesiącom. Jeśli symbolem $I_{M,t}$ oznaczymy wartość czynnika rynkowego w chwili t, to rozwiązując równanie będące definicją stopy zmian:

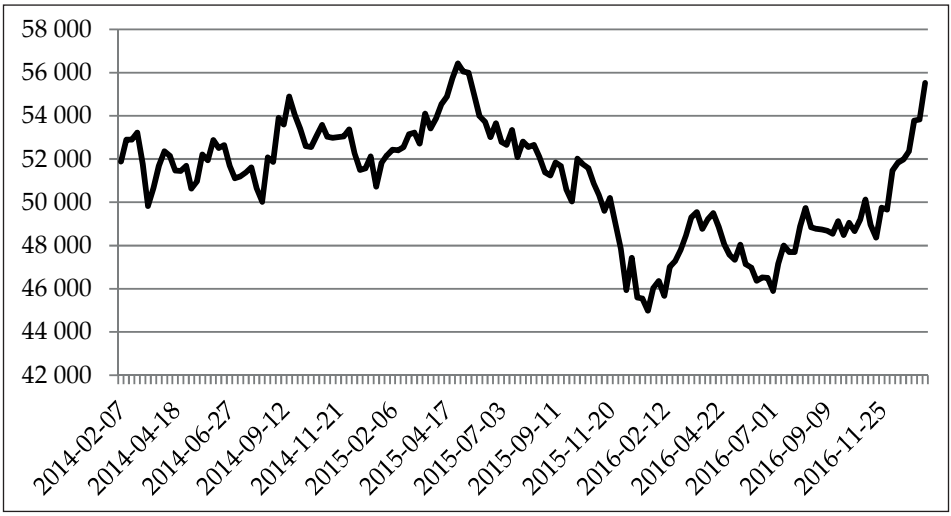
$$\frac{I_{M,t} - I_{M,t-1}}{I_{M,t-1}} = r_{M,t} \tag{2}$$

znajdujemy wartości benchmarku zależne od jego wartości początkowej I_0:

$$I_{M,t} = I_0 \Pi_{i=1}^{t}(1 + r_{M,i}) \tag{3}$$

W konsekwencji przebieg czynnika rynkowego uzyskanego na podstawie danych tygodniowych przedstawia rysunek 1, dla danych miesięcznych jest jakościowo zbliżony i nie będzie podany.

Rysunek 1. Wykres benchmarku w okresie od 1 lutego 2014 do 31 stycznia 2017 r. oparty na danych tygodniowych



Źródło: Opracowanie własne.

Równanie opisujące model Treynora–Mazuy'ego [Treynor, Mazuy, 1966] przyjęto w następującej postaci:

$$r_{A,t} - r_{f,t} = \alpha_A + \beta_A(r_{M,t} - r_{f,t}) + \delta_A(r_{M,t} - r_{f,t})^2 + \varepsilon_{A,t} \tag{4}$$

gdzie $r_{A,t}$, $r_{M,t}$, $r_{f,t}$ są stopami zwrotu w okresie t, odpowiednio: portfela funduszu A, benchmarku M oraz waloru pozbawionego ryzyka f.

W tym ostatnim przypadku, zarówno w tym miejscu, jak i w dalszej części pracy, przyjęto, że stopa ta jest równa stopie WIBOR 1W lub 1M. Wprawdzie istnieje w Polsce rynek skarbowych papierów dłużnych, ale jest on mało płynny, co łatwo zauważyć, przyglądając się notowaniom giełdowym. Natomiast dane z rynku pozagiełdowego są trudno dostępne

i nic nie wiadomo o jego płynności i wielkości. Powszechnie akceptowana interpretacja współczynników równania stwierdza, że współczynnik α_A jest miarą umiejętności zarządzania portfelem w zakresie selektywności aktywów, δ_A w zakresie stosowania market-timingu, natomiast β_A jest miarą wykorzystywania trendów rynkowych.

Model Henrikssona-Mertona [Henriksson, Merton, 1981] opisywany jest równaniem:

$$r_{A,t} - r_{f,t} = \alpha_A + \beta_A(r_{M,t} - r_{f,t}) + \gamma_A y_{M,t} + \varepsilon_{A,t} \tag{5}$$

gdzie: $y_{M,t} = \max\{0, r_{M,t} - r_{f,t}\}$. Interpretacja współczynników strukturalnych jest analogiczna do poprzedniej, z tym że odpowiednikiem δ_A jest teraz współczynnik γ_A.

W obu modelach badanie umiejętnego doboru aktywów do portfeli polega na weryfikacji hipotezy zerowej $\alpha_A = 0$, która oznacza, że zarządzający funduszem nie posiadają umiejętności przewidywania w skali mikro. Hipoteza alternatywna ma postać: $\alpha_A \neq 0$. Natomiast badanie umiejętności wyczucia rynku wymaga weryfikacji hipotezy zerowej $\delta_A = 0$ ($\gamma_A = 0$ dla modelu Henrikssona-Mertona), która oznacza, że zarządzający funduszem nie posiadają umiejętności w zakresie wyczucia rynku (trendu). Hipotezy alternatywne stwierdzają, że współczynniki te są różna od zera. W obu przypadkach poziom istotności był równy 0,05. Bardziej szczegółowe badania mogą dotyczyć hipotez alternatywnych w postaci nierówności, wówczas uzyskamy informacje o tym, czy działania zarządzających mają pozytywny czy negatywny wpływ na wartość portfela inwestycyjnego, jeśli zostanie odrzucona hipoteza zerowa.

W oparciu o oba modele definiuje się spółczynniki efektywności Treynora-Mazuy'ego oraz Henrikssona-Mertona zdefiniowane analogicznie do klasycznego współczynnika Jensena, a mianowicie:

$$TM_A = \bar{r}_A - r_f - \beta_A(\bar{r}_M - r_f) - \delta_A(\bar{r}_M - r_f)^2 \tag{6}$$

$$HM_A = \bar{r}_A - r_f - \beta_A(\bar{r}_M - r_f) - \gamma_A \bar{y}_M \tag{7}$$

Bierzemy wówczas pod uwagę średnie stopy zwrotu z udziałów w funduszach $\bar{r}_A$, średnią dla benchmarku $\bar{r}_M$, średnią $\bar{y}_M$ uwzględniającą zarówno wartości dodatnie, jak i zera ($y_{M,t} = \max\{0, r_{M,t} - r_{f,t}\}$) oraz średnią stopę pozbawioną ryzyka w okresie badań r_f. Podobnie jak dla współczynników α_A stawia się hipotezy zerowe $TM_A = 0$ i $HM_A = 0$, wobec hipotez alternatywnych jednostronnych lub dwustronnych w zależności od wymaganej precyzji badań. W niniejszym artykule, ze względu na określoną jego objętość, ograniczymy się do badania współczynników TM_A i HM_A, a nie α_A, jako miar umiejętności doboru aktywów przez zarządzających.

W obu modelach współczynnik β_A jest miarą podążania za trendem rynkowym, dlatego jest on wyznaczany ze standardowego modelu jednowskaźnikowego w postaci:

$$r_{A,t} = \alpha_A + \beta_A r_{M,t} + \varepsilon_{A,t} \tag{8}$$

Współczynnik ten opisuje pasywne strategie zarządzających. Ponieważ przyjęto założenie, że miarą aktywnej strategii są współczynniki Treynora-Mazuy'ego i Henrikssona-Mertona, więc wyraz wolny w modelu jednowskaźnikowym nie będzie badany.

2. Badanie wyczucia rynku i umiejętności doboru aktywów

W pierwszej kolejności zostaną zaprezentowane wyniki estymacji parametrów δ_A i γ_A w omówionych wyżej modelach. Tablica 2 dotyczy funduszy emerytalnych, a tablica 3 funduszy zrównoważonych.

Tablica 2. Estymatory parametrów wyczucia rynku w modelach Treynora-Mazuy'ego i Henrikssona-Mertona dla otwartych funduszy emerytalnych

Fundusz	Dane miesięczne		Dane tygodniowe	
	δ_A	γ_A	δ_A	γ_A
Aegon	–2,17100	–0,19461	1,53043	0,10515
Alianz	–1,82916	–0,11138	–0,53935	–0,01102
Aviva	–1,12196	–0,06369	0,00448	0,01420
Axa	–1,02604	–0,04165	0,56259	0,04959
Bankowy	**–3,57774**	–0,32995	0,23778	0,04265
Generali	–1,63388	–0,08655	0,84286	0,06269
Metlife	–0,43000	0,01059	–0,26774	0,00308
Nationale	–1,62159	–0,15032	0,72071	0,08470
Nordea	–1,05315	–0,05515	0,76146	0,06509
Pekao	–2,73820	–0,25232	0,66054	0,05699
Pocztylion	–2,18286	–0,12644	0,65402	0,05423
PZU	**–3,67699**	–0,37558	0,76840	0,06284

Źródło: Opracowanie własne.

Tablica 3. Estymatory parametrów wyczucia rynku w modelach Treynora-Mazuy'ego i Henrikssona-Mertona dla otwartych funduszy zrównoważonych

Fundusz	Dane miesięczne		Dane tygodniowe	
	δ_A	γ_A	δ_A	γ_A
Arka	**–3,42169**	**–0,44423**	1,29260	0,11290
Aviva	**–4,18985**	**–0,45459**	0,39995	0,03646
BPH	–3,34704	**–0,42482**	–0,17016	0,01231
Credit Agricole	–2,88016	–0,41280	0,62714	0,07401
Investor	**–8,98195**	**–1,18140**	–1,84077	–0,13729

Fundusz	Dane miesięczne		Dane tygodniowe	
	δ_A	γ_A	δ_A	γ_A
KBC	–0,33983	–0,05186	**0,90462**	**0,05187**
Nationale	–1,86431	0,06246	0,62143	0,06246
Novo	–0,88982	–0,27648	–1,10045	–0,10058
Pioneer	–2,47524	**–0,35780**	–0,59669	0,01180
PKO	–2,27950	**–0,33581**	0,34664	0,06025
PZU	–3,52380	–0,47039	0,62825	0,05933
Skarbiec	–3,33746	–0,46935	–0,54038	0,00095
UniKorona	–0,52647	–0,09779	0,00600	–0,03610

Źródło: Opracowanie własne.

Liczby pogrubione wskazują wartości istotnie różne od zera, hipoteza alternatywna była dwustronna. Gdyby nawet pokusić się o rozstrzygnięcie, czy te wartości są dodatnie czy ujemne, to widać, że prawie na pewno byłyby one ujemne, z wyjątkiem otwartego funduszu KBC, gdyż jest to jedyny przypadek dodatniego estymatora zarówno dla danych tygodniowych, jak i miesięcznych. Można zatem stwierdzić, że w przypadku współczynników wyczucia rynku fundusze inwestycyjne zrównoważone wypadają nieco gorzej niż emerytalne. Dla danych miesięcznych pojawia się więcej różnych od zera współczynników, ale są one ujemne, co oznacza „wyczucie" niezgodne z trendem. W konsekwencji w obu rodzajach podmiotów dominują wartości statystycznie równe zeru, co pozwala na stwierdzenie, że polityka aktywna mierzona współczynnikami delta i gamma w zasadzie nie istnieje, bez względu na horyzont inwestycyjny – tygodniowy i miesięczny.

Umiejętność selektywnego doboru aktywów do portfeli badano, obliczając współczynniki Treynora-Mazuy'ego (TM_A) oraz Henrikssona-Mertona (HM_A) zadane wzorami odpowiednio (6) i (7). Wyniki zawarto w tablicach 4 i 5.

Tablica 4. Wartości współczynników selektywności aktywów w modelach Treynora-Mazuy'ego i Henrikssona-Mertona dla otwartych funduszy emerytalnych

Fundusz	Dane miesięczne		Dane tygodniowe	
	TM_A	HM_A	TM_A	HM_A
Aegon	0,00058	0,00302	0,00017	–0,00050
Alianz	0,00074	0,00215	0,00020	0,00027
Aviva	0,00053	0,00134	0,00016	0,00007
Axa	0,00115	0,00168	0,00032	0,00001
Bankowy	0,00150	0,00564	0,00040	0,00013

Fundusz	Dane miesięczne		Dane tygodniowe	
	TM_A	HM_A	TM_A	HM_A
Generali	–0,00042	0,00068	–0,00007	–0,00047
Metlife	0,00158	0,00145	0,00041	0,00039
Nationale	0,00012	0,00201	0,00007	–0,00047
Nordea	0,00103	0,00173	0,00032	–0,00010
Pekao	0,00025	0,00342	0,00010	–0,00027
Pocztylion	0,00083	0,00244	0,00023	–0,00011
PZU	0,00097	0,00568	0,00029	–0,00011

Źródło: Opracowanie własne.

Tablica 5. Wartości współczynników selektywności aktywów w modelach Treynora-Mazuy'ego i Henrikssona-Mertona dla otwartych funduszy zrównoważonych

Fundusz	Dane miesięczne		Dane tygodniowe	
	TM_A	HM_A	TM_A	HM_A
Arka	–0,00024	0,00531	–0,00004	–0,00076
Aviva	–0,00079	0,00491	–0,00019	–0,00042
BPH	–0,00242	0,00289	–0,00058	–0,00066
Credit Agricole	0,00092	0,00607	0,00023	–0,00024
Investor	0,00514	0,01991	0,00130	0,00218
KBC	0,00062	0,00127	0,00017	–0,00016
Nationale	0,00124	0,00058	0,00032	–0,00008
Novo	–0,00517	–0,00174	–0,00131	–0,00067
Pioneer	–0,00379	0,00068	–0,00089	–0,00097
PKO	0,00037	0,00456	0,00011	–0,00027
PZU	–0,00061	0,00527	–0,00012	–0,00050
Skarbiec	–0,00050	0,00536	–0,00010	–0,00011
Unikorona	–0,00035	0,00087	–0,00009	0,00013

Źródło: Opracowanie własne.

Jak widać, wszystkie współczynniki są statystycznie nieistotne. Zatem w tym przypadku zgodność pomiaru umiejętności selektywnego doboru aktywów do portfeli dla funduszy emerytalnych i inwestycyjnych zrównoważonych jest jeszcze większa niż poprzednio. Współczynniki Treynora-Mazuy'ego oraz Henrikssona-Mertona wskazują jednoznacznie na brak aktywnej polityki inwestycyjnej.

3. Pasywna polityka inwestycyjna

Podążanie za trendem rynkowym polega na korelacji procentowych zmian portfela inwestycyjnego z benchmarkiem, zatem przyjęło się mierzyć je współczynnikiem beta w liniach charakterystycznych. W przypadku

badanych podmiotów czynnik rynkowy został opisany we wstępnej części pracy. Wyniki estymacji współczynników beta dla otwartych funduszy emerytalnych i otwartych funduszy zrównoważonych zebrano w tablicach 6 i 7.

Tablica 6. Estymatory współczynników beta dla otwartych funduszy emerytalnych

Fundusz	Dane miesięczne	Dane tygodniowe
Aegon	0,90210	0,80001
Alianz	0,91296	0,81931
Aviva	0,97863	0,84910
Axa	0,93625	0,80889
Bankowy	0,97520	0,83482
Generali	0,97110	0,83863
Metlife	0,99929	0,83253
Nationale	1,02836	0,87796
Nordea	0,99171	0,75070
Pekao	1,08462	0,92334
Pocztylion	1,02863	0,89235
PZU	1,03877	0,90389

Źródło: Opracowanie własne.

Tablica 7. Estymatory współczynników beta dla otwartych funduszy zrównoważonych

Fundusz	Dane miesięczne	Dane tygodniowe
Arka	0,63019	0,62143
Aviva	0,59053	0,53774
BPH	0,68068	0,62511
Credit Agricole	0,64037	0,57038
Investor	0,51390	0,52365
KBC	0,02512	0,00111
Nationale	0,60724	0,53537
Novo	1,06685	1,02399
Pioneer	0,77883	0,72463
PKO	0,52931	0,50987
PZU	0,97686	0,81857
Skarbiec	0,61520	0,64656
Unikorona	0,59693	0,53860

Źródło: Opracowanie własne.

Rezygnujemy z oznaczania parametrów istotnych statystycznie czcionką pogrubioną, gdyż wszystkie wartości podane w obu tabelach są różne od zera. Widać, że w przypadku współczynników beta w liniach charakterystycznych funduszy sytuacja wygląda zupełnie inaczej niż miało to miejsce w poprzedniej części. Fundusze emerytalne charakteryzują się współczynnikami o wartościach zbliżonych do jedynki. Fundusze zrównoważone mają wartości tej miary na ogół nieco mniejsze. W konsekwencji można stwierdzić, że obie grupy podmiotów charakteryzuje pasywna polityka inwestycyjna mierzona silną korelacją z rynkiem.

Zakończenie

Zarządzający portfelami nie są w stanie wykazać się umiejętnością selektywnego doboru walorów w celu poprawy efektywności inwestycyjnej. Pomiar aktywnego zarządzania portfelami wskazuje na brak takiej właśnie strategii. Zarządzający obu rodzajami podmiotów stosowali strategię pasywną, na co wskazuje silna (szczególnie w przypadku funduszy emerytalnych) korelacja z wykorzystywanym w pracy czynnikiem rynkowym. Z wykresu benchmarku (rys. 1) widać wyraźnie, że w badanym okresie mieliśmy do czynienia ze zmienną koniunkturą rynkową, zatem zarządzający mieli okazję do wykazania się swoimi umiejętnościami w zakresie market-timingu, za co fundusz pobiera stosowną opłatę. Niestety, wyniki pokazują, że nie skorzystali z tej okazji. Problem efektywności inwestycyjnej funduszy zbiorowego inwestowania, w tym emerytalnych, jest ciągle aktualny i rozwija się bardzo dynamicznie i prowadzi do nowych metod [Witkowska, 2016].

Literatura

Henriksson R., Merton R. (1981), *On Market Timing and Investment Performance. II Statistical Procedures for Evaluating Forecasting Skills*, "Journal of Business" Vol. 54, No. 4.

Karpio A, Żebrowska-Suchodolska D. (2014), *Efektywność inwestycyjna funduszy emerytalnych na tle FIO stabilnego wzrostu*, „Metody Ilościowe w Badaniach Ekonomicznych", nr 4.

Karpio A., Żebrowska-Suchodolska D. (2016), *Efektywność inwestycyjna polskich funduszy emerytalnych w okresie zmian zasad prawnych*, „Metody Ilościowe w Badaniach Ekonomicznych", nr 3.

Perez K. (2012), *Efektywność funduszy inwestycyjnych*, Difin, Warszawa.

Treynor J.L., Mazuy K. (1966), *Can Mutual Funds Outguess the Market?*, „Harvard Business Review", Vol. 44.

Witkowska D. (2016), *Propozycja oceny efektywności inwestycyjnej rynków funduszy emerytalnych*, „Metody Ilościowe w Badaniach Ekonomicznych", nr 2.

Streszczenie

Zmiana prawa dotyczącego składu portfeli otwartych funduszy emerytalnych, która zaczęła obowiązywać od początku 2014 r. upodobniła te podmioty do otwartych funduszy zrównoważonego wzrostu. Dlatego celem artykułu jest porównanie efektywności inwestycyjnej tych instytucji zbiorowego inwestowania w latach 2014–2017. Autorzy definiują benchmark rynkowy odzwierciedlający składy portfeli zawierających aktywa dwóch rodzajów: akcje oraz instrumenty mniej ryzykowne. Jest on podstawą tworzenia rankingów opartych na miarach: współczynnik beta w liniach charakterystycznych, współczynniki oparte na modelach Treynora-Mazuy'ego oraz Henrikssona-Mertona. Zatem dają one możliwość uzyskania odpowiedzi na pytanie o efektywność inwestycyjną związaną ze strategią pasywną, jak i wykorzystywaniem umiejętności wyczucia rynku i doboru dochodowych aktywów przez zarządzających portfelami. Wnioski wynikające z analizy dostarczają informacji o tym, które podmioty dają nadzieję na zgromadzenie większego kapitału emerytalnego, zarówno w ramach II, jak i III filara.

Słowa kluczowe

fundusz emerytalny, efektywność inwestycyjna, benchmark rynkowy, strategia pasywna, strategia aktywna

Comparison analysis of the active and passive investment strategies of open-end pension funds and open-end balanced mutual funds in the years 2014–2017 (Summary)

The change of legal rules referring to the structure of pension funds' investment portfolios which has been in force since the beginning of 2014, have made these institutions similar to the balanced open-end mutual funds. That is why the purpose of the work is the comparison of the investment effectiveness of these mutual funds in the years 2014–2017. The authors define market benchmark which reflects the structure of portfolios including the assets of two kinds: shares and less risk securities. It is the basis of the creation of rankings based on the following measures: beta coefficient in characteristic lines, coefficients based on the models of Treynor-Mazuy and Henriksson-Merton. So, they give the possibility to answer the questions about the investment effectiveness of the passive strategy and also the ability of market timing and selecting of profit assets by the portfolio managers. The conclusions which result from the analysis supply the information about the institutions which hope for accumulation of more pension capital within II and III pillar.

Keywords

pension fund, investment efficiency, benchmark, passive strategy, active strategy

Dorota Żebrowska-Suchodolska*

Podobieństwo miar opartych na momentach cząstkowych wyznaczonych dla funduszy akcyjnych w latach 2004–2015

Wstęp

Wskaźnik Sharpe'a jest jedną z najczęściej używanych miar do oceny efektywności inwestycji. Już jednak sama konstrukcja wzoru może wpływać na mylny obraz danego funduszu ze względu na przyjętą tam definicję ryzyka. Alternatywą zatem stają się wskaźniki, w których ryzyko rozumiane jest wyłącznie w kontekście straty dla inwestora. Do takich wskaźników zaliczamy m.in. miary oparte na momentach cząstkowych.

Z uwagi na wielość mierników służących do oceny efektywności inwestycji zasadne staje się pytanie, czy ważny jest wybór odpowiedniej miary. W przypadku bowiem istotnej korelacji wybranie konkretnego wskaźnika nie będzie miało tak dużego znaczenia w praktyce. W literaturze można odnaleźć m.in. prace Ellinga i Schuhmachera [2006], Kooli, Morin i Sedzro [2005] czy Wiesinger [2010], w których autorzy wskazują na istnienie licznych zależności między wskaźnikami służącymi do oceny efektywności inwestycji. Elling i Schuhmacher [2005] potwierdzili występowanie istotnej dodatniej korelacji między rankingami 2763 funduszy hedgingowych wyznaczonymi przy użyciu 13 miar. Podobne wyniki otrzymali Kooli, Morin i Sedzro [2005], badając 675 funduszy hedgingowych. Również badania Wiesinger [2010] wskazują na to, iż wiele miar jest ze sobą skorelowanych.

Z uwagi na to, iż wcześniejsze badania autora dotyczyły miar będących modyfikacją wskaźnika Sharpe'a, artykuł skupia się na kolejnej grupie miar – opartych na momentach cząstkowych. Posłużą one do zbudowania rankingów i w efekcie zbadania zależności między nimi. Celem pracy jest próba odpowiedzi na pytanie, czy wybór miary spośród wskaźników opartych na momentach cząstkowych ma istotne znaczenie dla inwestora.

1. Dane i metodologia badań

Obawa inwestorów przed możliwą stratą spowodowała, iż zaczęto konstruować różne narzędzia zwracające szczególną uwagę na odchylenie *in minus*. Należą do nich m.in. miary oparte na momentach cząstkowych.

* Dr, Katedra Ekonometrii i Statystyki, Wydział Zastosowań Informatyki i Matematyki, SGGW w Warszawie, ul. Nowoursynowska 159, 02-787 Warszawa, zdorota@o2.pl

Dolny moment cząstkowy (*lower partial moment) n*-tego rzędu opisany jest wzorem [Kaplan, Knowles, 2004, s. 3]:

$$LPM_n(m) = \frac{1}{T}\sum_{t=1}^{T} \max[m - r_t, 0]^n \tag{1}$$

gdzie *m* jest minimalną stopą zwrotu możliwą do zaakceptowania przez inwestora. Za *m* może być przyjęta średnia stopa zwrotu, wartość zero lub stopa zwrotu wolna od ryzyka.

Dolny moment cząstkowy jest podstawą takich miar, jak wskaźnik Omega, Omega-Sharpe'a, *gain-loss*, Sortino, Ziemby czy M²-Sortino.

Wskaźnik Omega został zaproponowany przez Keatinga i Shadwicka w 2002 r. Ma on następującą postać:

$$Omega(m) = \frac{\bar{r}_t - m}{LPM_1(m)} + 1 \tag{2}$$

gdzie $\bar{r}_t$ jest średnią stopą zwrotu funduszu.

Wskaźnik Omega jest powiększonym o jeden stosunkiem nadwyżkowej stopy zwrotu do ujemnego odchylenia. Im wartość wskaźnika jest wyższa, tym inwestowanie w dany fundusz jest bardziej opłacalne. Jeśli za *m* zostanie przyjęta średnia stopa zwrotu, wskaźnik Omega osiągnie wartość 1.

Modyfikacją miernika Omega jest wskaźnik Omega-Sharpe'a opisany następującym wzorem [Bacon, 2009, s. 7]:

$$Omega\text{-}Sharpe(m) = \frac{\bar{r}_t - m}{LPM_1(m)} \tag{3}$$

Przypomina on swoim zapisem wskaźnik Sharpe'a w liczniku i wskaźnik Omega w mianowniku.

Szczególnym przypadkiem wskaźnika Omega dla $m = 0$ jest wskaźnik Bernardo-Ledoita, nazywany też wskaźnikiem *gain-loss* [Bernardo-Ledoit, 2000]:

$$gain\text{-}loss(0) = \frac{\bar{r}_t}{LPM_1(0)} + 1 \tag{4}$$

Na dolnym momencie cząstkowym rzędu drugiego oparty jest wskaźnik Sortino podany przez Sortino i van der Meera w 1991 r.:

$$Sortino(m) = \frac{\bar{r}_t - m}{\sqrt{LPM_2(m)}} \tag{5}$$

Koncentruje się on na stopach zwrotu poniżej średniej, przyjmując za miarę ryzyka semiwariancję. Wskaźnik Sortino jest więc stosunkiem zysku do ryzyka rozumianego jako odchylenie standardowe poniżej średniej

stopy zwrotu. Im niższa jest wartość wskaźnika Sortino, tym inwestycja jest bardziej podatna na straty. Należy zatem unikać inwestycji o małych wartościach tego miernika. Ujemny wynik wskaźnika Sortino świadczy o tym, iż portfel nie osiągnął minimalnej akceptowalnej stopy zwrotu.

Modyfikacją wskaźnika Sortino jest miernik M^2-Sortino bazujący na zapisie wskaźnika M^2, który w odróżnieniu od wskaźnika Sharpe'a uważa się za lepszą miarę w interpretacji dla inwestorów [Bacon, 2008, s. 102–103].

$$M^2 - Sortino = \overline{r_t} + Sortino \cdot (\sigma_m - \sigma) \tag{6}$$

gdzie: σ_m – odchylenie standardowe benchmarku, σ – odchylenie standardowe stopy zwrotu funduszu.

Miarą opartą na ujemnej wariancji jest podany przez Ziembę [2005] miernik bazujący na wskaźniku Sharpe'a:

$$Ziemba(0) = \frac{\overline{r_t} - r_f}{\sqrt{2LPM_2(0)}} \tag{7}$$

Takie rozumienie ryzyka powoduje, że uznawany jest on za lepszą miarę niż wskaźnik Sharpe'a do wskazywania bardziej efektywnych aktywów.

Górny moment cząstkowy (*higher partial moment*) jest postaci [Eling, Schuhmacher, 2006, s. 431]:

$$HPM_n(m) = \tfrac{1}{T}\sum_{t=1}^{T} \max[r_t - m, 0]^n \tag{8}$$

Występuje on we wzorach wskaźnika potencjału nadwyżkowej stopy zwrotu (*upside potential ratio*) *UPR* i wskaźnika *d*.

Wskaźnik potencjału nadwyżkowej stopy zwrotu *UPR* został zaproponowany przez Sortino, van der Meer i Platinga w 1999 r.:

$$UPR(m) = \frac{HPM_1(m)}{\sqrt[2]{LPM_2(m)}} \tag{9}$$

Jest on modyfikacją wskaźnika Sortino skupiającym się zarówno na stopach powyżej, jak i poniżej progu *m*. Taki zapis wzoru powoduje, iż dodatkowo nagradzani są ci, którzy wypracowują stopę zwrotu powyżej określonego progu *m*.

Wskaźnik *d* opisany jest wzorem [Lavinio, 1999]:

$$d - ratio = \frac{n_u \cdot LPM_1(0)}{n_d \cdot HPM_1(0)} \tag{10}$$

gdzie: n_d – liczba stóp zwrotu dodatnich, n_u – liczba stóp zwrotu ujemnych.

Przyjmuje on wartości od zera do nieskończoności. Swoim zapisem przypomina wskaźnik *gain-loss*, ale tu dodatkowo brana jest pod uwagę ilość dodatnich i ujemnych stóp zwrotu. Mniejsze wartości wskaźnika lepiej

świadczą o funduszu. Wartość zero wskaźnik osiąga wtedy, gdy w szeregu nie ma stóp zwrotu ujemnych, a nieskończoność – gdy nie występują stopy zwrotu większe od 0.

Badaniami objęto 16 funduszy akcyjnych funkcjonujących na polskim rynku w latach 2004–2015, podzielonych na krótsze podokresy (tj. 2-, 3-, 4- i 5-letnie). Były to następujące fundusze: Arka BZWBK Akcji, BPH Akcji, BPH Akcji Dynamicznych Spółek, CU Akcji Polskich, DWS Akcji, DWS Akcji Plus, DWS Top 25, ING Akcji, Legg Mason Akcji, Millennium Akcji, Pioneer Akcji Polskich, PKO/CS Akcji, PZU Akcji Krakowiak, SEB 3, Skarbiec Akcja, UniKorona Akcje. W przyjętym okresie badań niektóre z funduszy zmieniły swoje nazwy, co było związane z różnego rodzaju konsolidacjami i przejęciami. Z uwagi jednak na wcześniejsze badania autora w pracy stosowane będzie nazewnictwo początkowe funduszu.

Na podstawie miesięcznych zwykłych stóp zwrotu wyznaczono wskaźniki oparte na dolnym i górnym momencie cząstkowym, tj. Omega, Omega-Sharpe'a Sortino, M^2-Sortino *gain-loss*, Ziemby, *UPR* i *d*. Wyznaczone wartości wskaźników były podstawą zbudowania rankingów w poszczególnych podokresach. Chcąc odpowiedzieć na pytanie, czy decyzja o wyborze konkretnej miary ma jakiekolwiek znaczenie, zweryfikowano najpierw test rang Friedmana, stawiając hipotezę o równości średnich rang dla kolejnych miar wobec hipotezy alternatywnej, że nie wszystkie średnie rang są sobie równe. Statystyka testowa ma przy tym postać [Friedman, 1937]:

$$\chi^2 = \frac{12}{nk(k+1)} \sum_{j=1}^{k} R_j^2 - 3n(k+1) \tag{11}$$

gdzie: R_j – suma rang j-tej zmiennej, n – liczba obserwacji, k – liczba zmiennych.

Statystyka ma rozkład χ^2 z n–1 stopniami swobody.

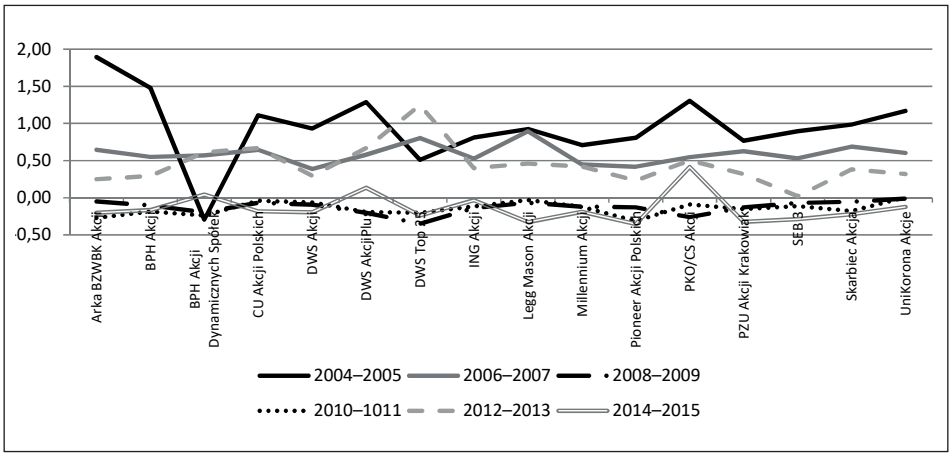
W przypadku braku podstaw do odrzucenia hipotezy zerowej wyznaczono współczynnik podobieństwa tau-Kendalla i zbadano jego istotność. Postawiono przy tym hipotezę, że współczynnik korelacji rangowej $r_ł$ jest nieistotny statystycznie wobec hipotezy alternatywnej, że współczynnik korelacji $r_ł$ jest istotny statystycznie. We wszystkich testach przyjęto poziom istotności 0,05.

2. Wyniki

Dolny i górny moment cząstkowy stały się podstawą wskaźników, tj. Omega, Omega-Sharpe'a Sortino, M^2-Sortino *gain-loss*, Ziemby, *UPR* i *d*. Wyznaczono je dla funduszy akcyjnych w podokresach 2-, 3-, 4- i 5-letnich. Jako minimalną akceptowalną przez inwestora stopę zwrotu przyjęto wartość zero. Takie zdefiniowanie *m* powoduje, iż wartości wskaźników Omega

i *gain-loss* stają się identyczne. Wyniki dla wskaźnika Sortino w okresach dwuletnich przedstawiono na rysunku 1.

Rysunek 1. Wartości wskaźnika Sortino w podokresach dwuletnich



Źródło: Opracowanie własne.

Najwyższe wartości wskaźnika Sortino spośród podokresów dwuletnich fundusze osiągnęły w okresie dobrej koniunktury giełdowej przypadającej na lata 2004–2005, a najniższe w latach będących początkiem kryzysu światowego, tj. 2008–2009. Spośród wszystkich funduszy najwyższą wartość odnotował fundusz Arka BZWBK Akcji w okresie 2004–2005, a najniższą DWS Top 25 w latach 2008–2009. Zysk z inwestycji kompensujący inwestorowi poniesione ryzyko ma miejsce jedynie w pojedynczych przypadkach. Wartości powyżej jedności wystąpiły bowiem dla sześciu funduszy w latach 2004–2005 i jednego w okresie 2012–2013. Biorąc pod uwagę podokresy trzyletnie, najwyższe wartości wskaźnika Sortino wystąpiły w podokresie 2004–2006, a podokresem, który plasował się na końcu pod względem osiągniętych wyników, był okres 2007–2009. Jedynie w okresie 2004–2006 wyniki wskaźnika Sortino przewyższały wartość jeden. W podziale na podokresy czteroletnie najwyższe wartości pod względem wskaźnika fundusze osiągnęły w latach 2004–2007 (w tym podokresie tylko jeden fundusz osiągnął wartość przewyższającą jedność), a wśród podokresów pięcioletnich – w latach 2009–2013. Największe spadki odnotowane w latach 2008–2009 tak znacząco wpłynęły na wyniki, że wśród podokresów pięcioletnich gorsze wyniki odnotowano dla lat 2004–2008 niż 2009–2013.

Podobne wyniki, tylko znajdujące się w przedziałach o innym rozstępie, fundusze osiągnęły w przypadku wskaźnika Omega, Ziemby i M^2-Sortino. Dla wszystkich podokresów wartości wskaźnika Omega-Sharpe'a były

pomniejszeniem o jeden wyników otrzymanych dla wskaźnika Omega, a wyniki wskaźnika *gain-loss* były identyczne jak w przypadku wskaźnika Omega.

Biorąc pod uwagę wskaźnik *UPR*, przewaga odchyleń dodatnich wystąpiła dla większości funduszy w okresie 2004–2005, 2006–2007, 2004–2006 oraz 2004–2007. Wyniki, jakie fundusze uzyskały w przypadku wskaźnika *d*, były wręcz odwrotne w stosunku do pozostałych wskaźników, co wynika z konstrukcji tego miernika.

Wyznaczone wskaźniki stały się podstawą utworzenia rankingów funduszy opartych na momentach cząstkowych. Tablica 1 przedstawia rankingi otrzymane przy użyciu miar Omega, Omega-Sharpe'a, Sortino, M^2-Sortino, *gain-loss*, Ziemby, *UPR* i *d* dla okresów dwuletnich. Z uwagi na to, iż rankingi wskaźników Omega-Sharpe'a, *gain-loss* były identyczne jak dla wskaźnika Omega, zostały one pominięte w tablicy 1.

Tablica 1. Rankingi w podokresach dwuletnich

	2004–2005						2006–2007					
	Omega	*Sortino*	*Ziemba*	*M2-Sortino*	*UPR*	*d*	*Omega*	*Sortino*	*Ziemba*	*M2-Sortino*	*UPR*	*d*
Arka BZWBK Akcji	1	1	1	1	1	16	4	5	5	2	8	15
BPH Akcji	2	2	2	2	2	14	12	10	10	9	10	4
BPH Akcji Dynamicznych Spółek	16	16	16	16	16	1	11	9	7	16	7	5
CU Akcji Polskich	6	6	6	5	6	15	5	4	4	10	4	9
DWS Akcji	9	8	9	8	8	7	16	16	16	14	16	1
DWS AkcjiPlus	3	4	3	4	4	13	8	8	9	8	9	12
DWS Top 25	15	15	15	15	15	6	1	2	2	12	2	14
ING Akcji	11	11	11	12	12	5	13	13	12	13	11	3
Legg Mason Akcji	8	9	8	9	10	8	2	1	1	4	1	13
Millennium Akcji	14	14	14	14	14	2	14	14	14	11	14	8
Pioneer Akcji Polskich	13	12	12	13	11	4	15	15	15	15	15	2
PKO/CS Akcji	4	3	4	3	3	12	9	11	11	5	13	6
PZU Akcji Krakowiak	12	13	13	11	13	10	6	6	6	3	5	7
SEB 3	10	10	10	10	9	3	10	12	13	7	12	11
Skarbiec Akcja	7	7	7	7	7	9	3	3	3	1	3	16
UniKorona Akcje	5	5	5	6	5	11	7	7	8	6	6	10

	2008–2009						2010–2011					
	Omega	*Sortino*	*Ziemba*	*M2-Sortino*	*UPR*	*d*	*Omega*	*Sortino*	*Ziemba*	*M2-Sortino*	*UPR*	*d*
Arka BZWBK Akcji	2	2	2	2	3	12	15	15	15	15	14	3
BPH Akcji	8	8	8	9	6	13	10	11	12	11	10	7
BPH Akcji Dynamicznych Spółek	13	13	13	13	8	2	14	14	14	14	15	2
CU Akcji Polskich	4	5	3	4	5	8	3	3	2	3	4	9
DWS Akcji	7	7	7	7	7	9	4	4	4	4	3	15
DWS AkcjiPlus	14	14	14	14	14	3	12	12	11	12	12	4
DWS Top 25	16	16	16	16	15	1	13	13	13	13	13	10
ING Akcji	12	12	12	12	10	5	6	7	5	6	7	12
Legg Mason Akcji	6	4	6	6	12	14	2	2	3	2	2	16
Millennium Akcji	9	9	10	10	13	7	8	8	7	8	6	11
Pioneer Akcji Polskich	11	11	9	8	9	6	16	16	16	16	16	1
PKO/CS Akcji	15	15	15	15	16	4	5	5	6	5	5	13
PZU Akcji Krakowiak	10	10	11	11	11	11	9	9	9	9	9	8
SEB 3	5	6	5	5	4	10	7	6	8	7	8	5
Skarbiec Akcja	3	3	4	3	2	15	11	10	10	10	11	6
UniKorona Akcje	1	1	1	1	1	16	1	1	1	1	1	14

	2012–2013						2014–2015					
	Omega	*Sortino*	*Ziemba*	*M2-Sortino*	*UPR*	*d*	*Omega*	*Sortino*	*Ziemba*	*M2-Sortino*	*UPR*	*d*
Arka BZWBK Akcji	14	14	14	14	15	7	10	10	13	10	10	10
BPH Akcji	12	13	12	11	13	9	6	6	6	6	5	5
BPH Akcji Dynamicznych Spółek	4	4	4	4	4	13	3	3	3	3	3	14
CU Akcji Polskich	2	2	3	2	3	15	7	7	8	7	8	13
DWS Akcji	13	12	13	13	11	2	9	9	12	9	9	11
DWS AkcjiPlus	3	3	2	3	2	14	2	2	2	2	2	15
DWS Top 25	1	1	1	1	1	16	12	12	10	12	12	7
ING Akcji	7	8	7	8	9	8	4	4	4	4	4	12
Legg Mason Akcji	6	6	6	6	6	10	14	15	15	14	13	1
Millennium Akcji	8	7	8	7	7	6	8	8	9	8	6	6

Pioneer Akcji Polskich	15	15	15	15	14	3	16	16	16	16	16	2
PKO/CS Akcji	5	5	5	5	5	12	1	1	1	1	1	16
PZU Akcji Krakowiak	10	10	10	12	12	11	15	14	14	15	14	3
SEB 3	16	16	16	16	16	1	13	13	11	13	15	4
Skarbiec Akcja	9	9	9	9	8	5	11	11	7	11	11	9
UniKorona Akcje	11	11	11	10	10	4	5	5	5	5	7	8

Źródło: Opracowanie własne.

Rankingi funduszy wyznaczone dla podokresów dwuletnich były bardzo zbliżone do siebie, a w wielu przypadkach nawet identyczne. Największe różnice w rankingach zauważono w przypadku rang otrzymanych przy użyciu wskaźnika *d* oraz *UPR*. Podobnie sytuacja wyglądała w przypadku podokresów trzyletnich, czteroletnich i pięcioletnich. W celu uogólnienia wyników zastosowano test Friedmana. Wartości statystyki dla wszystkich podokresów zaprezentowano w tablicy 2.

Tablica 2. Wartości statystyki testowej testu Fischera w poszczególnych podokresach

Podokres	χ^2
2004–2005	0,534
2006–2007	4,111
2008–2009	1,202
2010–2011	1,012
2012–2013	0,369
2014–2015	0,658
2004–2006	1,268
2007–2009	1,998
2010–2012	0,709
2013–2015	2,453
2004–2007	10,088
2008–2011	0,538
2012–2015	8,781
2004–2008	0,942
2009–2013	0,303

Źródło: Opracowanie własne.

We wszystkich badanych podokresach brak jest podstaw do odrzucenia hipotezy zerowej o równości średnich rang dla kolejnych miar. W następnym kroku wyznaczono zatem współczynnik podobieństwa tau-Kendalla.

Zbliżone rankingi otrzymane przy użyciu miar opartych na momentach cząstkowych, a w wielu przypadkach nawet identyczne, potwierdzone zostały wysokimi wartościami współczynnika tau-Kendalla. Wartości współczynnika dla podokresów dwuletnich przedstawia tablica 3.

Tablica 3. Współczynnik tau-Kendalla wyznaczony dla wskaźników w podokresach dwuletnich

	2004–2005					2006–2007					2008–2009				
	Sortino	*Ziemba*	*M2–Sortino*	*UPR*	*d*	*Sortino*	*Ziemba*	*M2–Sortino*	*UPR*	*d*	*Sortino*	*Ziemba*	*M2–Sortino*	*UPR*	*d*
Omega	**0,95**	**0,98**	**0,93**	**0,92**	**–0,75**	**0,90**	**0,85**	**0,50**	**0,78**	**–0,70**	**0,97**	**0,95**	**0,95**	**0,68**	**–0,72**
Sortino		**0,97**	**0,95**	**0,97**	**–0,70**		**0,95**	**0,43**	**0,88**	**–0,60**		**0,92**	**0,92**	**0,65**	**–0,75**
Ziemba			**0,92**	**0,93**	**–0,73**			**0,38**	**0,90**	**–0,55**			**0,97**	**0,70**	**–0,67**
M2–Sortino				**0,92**	**–0,75**				0,32	**–0,50**				**0,70**	**–0,67**
UPR					**–0,67**					**–0,52**					**–0,53**
	2010–2011					2012–2013					2014–2015				
	Sortino	*Ziemba*	*M2–Sortino*	*UPR*	*d*	*Sortino*	*Ziemba*	*M2–Sortino*	*UPR*	*d*	*Sortino*	*Ziemba*	*M2–Sortino*	*UPR*	*d*
Omega	**0,97**	**0,92**	**0,98**	**0,93**	**–0,72**	**0,97**	**0,98**	**0,95**	**0,88**	**–0,78**	**0,98**	**0,85**	**1**	**0,92**	**–0,72**
Sortino		**0,92**	**0,98**	**0,90**	**–0,68**		**0,95**	**0,95**	**0,92**	**–0,75**		**0,87**	**0,98**	**0,90**	**–0,73**
Ziemba			**0,93**	**0,88**	**–0,67**			**0,93**	**0,90**	**–0,77**			**0,85**	**0,77**	**–0,63**
M2–Sortino				**0,92**	**–0,70**				**0,90**	**–0,73**				**0,92**	**–0,72**
UPR					**–0,75**					**–0,67**					**–0,63**

Źródło: Opracowanie własne.

Otrzymane wyniki wskazują na silne podobieństwo miar wyznaczonych dla podokresów dwuletnich. Jedyna wartość nieistotna statystycznie pojawiła się pomiędzy miarami *UPR* i M²-Sortino w latach 2006–2007.

W przypadku podokresów trzyletnich wartości nieistotne statystycznie wystąpiły między wskaźnikiem M²-Sortino a wszystkimi pozostałymi miarami w latach 2004–2006 i 2007–2009 oraz pomiędzy wskaźnikiem *d* a M²-Sortino w podokresie 2010–2012. W podziale na okresy czteroletnie wszystkie wartości były istotne statystycznie. W podokresach pięcioletnich współczynnik nieistotny statystycznie wystąpił jedynie pomiędzy miernikami *UPR* i *d* w okresie 2009–2013.

Zakończenie

Miary oparte na momentach cząstkowych są jedną z grup miar nieklasycznych stanowiących alternatywę miar klasycznych. Brak założeń co do normalności rozkładu, ale też i inne postrzeganie ryzyka sprawia, że są one często wybierane do oceny efektywności inwestycji różnych aktywów. Badania dotyczące tej właśnie grupy wskazują na silne podobieństwo wskaźników między sobą. Można jednak dopatrzeć się wskaźników, które w pojedynczych podokresach są nieistotne statystycznie i różnią się od pozostałych. Najbardziej dostrzegalny jest tu wskaźnik M^2-Sortino, a po nim mierniki *UPR* i *d*. Warto zatem przy ocenie efektywności inwestycji podeprzeć się również którąś z miar nieskorelowanych z pozostałymi, aby ta ocena była pełna i miarodajna.

Literatura

Bacon C.R. (2008), *Practical portfolio performance measurement and attribution*, John Wiley & Sons, Hoboken.

Bacon C.R., (2009), *How sharp is the Sharpe-ratio? , Risk-adjusted Performance Measures*, StatPro, http://www.statpro.com/resources/whitepapers/how_sharp_is_the_sharp_ratio.aspx.

Bernardo A.E. (2000), *Gain, Loss and Asset Pricing*, „Journal of Political Economy" Vol. 108, No. 1.

Eling M., Schuhmacher F. (2006), *Hat die Wahl des Performancemases einen Einfluss auf die Beurteilung von Hedgefonds-Indizes?*, „Kredit und Kapital", Vol. 39, No. 3.

Friedman M. (1937), *The use of ranks to avoid the assumption of normality implicit in the analysis of variance*, Journal of the American Statistical Association. American Statistical Association, Vol. 32, No. 200.

Kaplan, P., Knowles, J. (2004), *Kappa: A Generalized Downside Risk-Adjusted Performance Measure*, „Journal of Performance Measurement", Vol. 8, No. 3.

Keeting C., Shadweek W.F. (2002), *A Universal Performance Measure*, „Journal of Performance Measurement", Vol. 6, No. 3.

Kooli M., Morin F., Sedzro K. (2005), *Evaluation des Mesures de Performance des Hedge Funds*, Paper presented at the Annual Conference of the French Association of Finance, June.

Lavinio S. (1999), *The Hedge Fund Handbook*, McGraw-Hill.

Sortino, F. Price, L. (1994), *Performance Measurement in a Downside Risk Framework*, „Journal of Investing", Vol. 3, No. 3.

Shadwick W., Keating C. (2002), *A universal performance measure*, „Journal of Performance", Vol. 6, No. 3.

Sortino F.A., van der Meer R. (1991), *Downside Risk. Capturing What's at Stake in Investments Situations*, „Journal of Portfolio Management", Vol. 17, No. 4.

Sortino F.A., van der Meer R., Plantinga A. (1999), *The Dutch Triangle*, „Journal of Portfolio Management", Vol. 26, No. 1.

Wiesinger A. (2010), *Risk-Adjusted Performance Measurement–State of the Art*, Bachelor s Thesis.

Ziemba W.T. (2005), *The Symmetric Downside-Risk Sharpe Ratio*, Working Paper.

Streszczenie

Praca porównuje wyniki otrzymane przy użyciu wybranych miar opartych na górnym i dolnym momencie cząstkowym oraz bada zależności między nimi. Badania dotyczą wskaźników: Omega, Omega-Sharpe'a, gain-loss, Sortino, Ziemby, M^2-Sortino, *UPR*, *d* . Zostały one wyznaczone dla 16 funduszy akcyjnych w okresie 2004–2015. Cały okres badań podzielono na krótsze podokresy (2-, 3-, 4- i 5-letnie). Otrzymane wyniki wskazują na silne podobieństwo miar opartych na momentach cząstkowych.

Słowa kluczowe

efektywność inwestycji, górny moment cząstkowy, dolny moment cząstkowy

The similarity of measures based on partial moments for mutual funds in the years 2004–2015 (Summary)

The paper compares the results obtained using the selected upper and lower partial moments and examines the relationships between them. Omega, Omega--Sharpe, gain-loss, Sortino, Ziemby, M^2-Sortino, *UPR*, d. They were designated for 16 equity funds in the years 2004–2015. The whole study period was divided into shorter sub-periods (2-, 3-, 4 and 5 years). The results show a strong similarity of measures based on partial moments.

Keywords

investment efficiency, higher partial moment, lower partial moment

Stanisław Wieteska*
Małgorzata Jeziorska**
Piotr Majewski***

Identyfikacja pola ubezpieczeniowego i ocena ryzyka eksploatacji samochodowych pojazdów elektrycznych w ubezpieczeniu autocasco

Wstęp

Już wiele lat temu dostrzeżono, że rozwój motoryzacji na bazie silników benzynowych i oleju napędowego przyczynia się do emisji wielu szkodliwych dla środowiska substancji. Emisja dwutlenku węgla powoduje stopniowe ocieplanie się klimatu wraz ze wszystkimi negatywnymi konsekwencjami tego zjawiska. Aby powstrzymać szkodliwą emisję, dąży się do ograniczenia substancji produkowanych przez pojazdy samochodowe. Jednym ze sposobów jest rozwój produkcji pojazdów mechanicznych o napędzie elektrycznym.

W rezolucji z dnia 6 maja 2010 r. w sprawie pojazdów elektrycznych Parlament Europejski zwraca się do Komisji państw członkowskich o ustanowienie warunków niezbędnych dla istnienia jednolitego rynku pojazdów elektrycznych przy zagwarantowaniu skutecznej koordynacji polityki na szczeblu UE, aby uniknąć negatywnych skutków społecznych związanych z przejściem na zdekarbonizowany system transportu oraz aby uniknąć współistnienia niezgodnych systemów i standardów. W rezolucji podkreśla się, że rozwój pojazdów elektrycznych powinien znaleźć miejsce w przyszłej polityce zrównoważonej mobilności. Gwałtowna ekspansja pojazdów elektrycznych wskazywana jest przez wiele źródeł jako jeden z wiodących trendów rozwoju gospodarek na świecie [Deloitte Insight, 2018; PWC, 2016].

Samochód na prąd jest wyzwaniem cywilizacyjnym. Otwiera nowy rozdział w dziejach transportu i przemysłu samochodowego [PWC, 2017;

* Prof. dr hab., Katedra Ekonomii, Wydział Nauk Społecznych, Uniwersytet Jana Kochanowskiego w Kielcach Filia w Piotrkowie Trybunalskim, ul. J. Słowackiego 114/118, 97-300 Piotrków Trybunalski, s.wieteska@unipt.pl
** Dr inż., Instytut Finansów, Wydział Ekonomiczno-Socjologiczny, Uniwersytet Łódzki, ul. Polskiej Organizacji Wojskowej 3/5, 90-255 Łódź, malgorzata.jeziorska@uni.lodz.pl
*** Dr, Katedra Finansów i Rachunkowości, Wydział Finansów i Zarządzania w Toruniu, Wyższa Szkoła Bankowa w Toruniu, ul. Młodzieżowa 31a, 87-100 Toruń, piotr.majewski@wsb.torun.pl

Burns, 2013]. Potwierdzają to także badania opinii menedżerów odpowiedzialnych za rozwój przemysłu motoryzacyjnego [KPMG, 2018].

W Strategii Rozwoju Transportu do 2020 r. opracowanej przez Ministerstwo Infrastruktury transportowi samochodami elektrycznymi poświęca się za mało miejsca. W dokumencie tym zauważa się jedynie, że Polska wciąż jest białą plamą w obszarze tworzonej infrastruktury zasilania energetycznego pojazdów samochodowych o alternatywnym napędzie: sieci ładowania bateryjnych samochodów elektrycznych (BEV) i sieci tankowania wodoru do samochodów napędzanych paliwami ogniowymi (FCV) [*Elektromobilność*, 2011, s. 11–17].

Postępując zgodnie z wytycznymi Unii Europejskiej, aby wspomóc rozwój ekosystemu elektromobilności w Polsce, Ministerstwo Energii przygotowało Program Rozwoju Elektromobilności przewidziany na lata 2016–2025. Plan ten zakłada, że po polskich drogach w 2025 r. będzie jeździć milion samochodów elektrycznych. Tak dynamiczny rozwój wymagać będzie reakcji zakładów ubezpieczeń w zakresie oferowanej ochrony ubezpieczeniowej i procesu taryfikacji. Przegląd literatury wskazuje na nieliczne publikacje opisujące to zagadnienie.

Celem artykułu jest identyfikacja pola ubezpieczeniowego i ocena wybranych aspektów ryzyka związanego z eksploatacją pojazdów mechanicznych o napędzie elektrycznym w ubezpieczeniu autocasco. W związku z różnicami związanymi ze specyfiką eksploatacji tego typu pojazdów proces dostosowania będzie konieczny.

Artykuł napisano w oparciu o załączoną literaturę przedmiotu. Autorzy nie specjalizują się w elektromobilności, lecz z uwagi na zajmowanie się ubezpieczeniami majątkowymi zwracają uwagę na ten nowy, przyszłościowy obszar ryzyka.

1. Pojęcie samochodu elektrycznego (EV). Skala produkcji, stan świadomości

Zgodnie z najbardziej ogólną definicją przez pojazd elektryczny należy rozumieć pojazd posiadający jako napęd jeden lub więcej silników elektrycznych lub silniki trakcyjne. W ramach tej definicji mieszczą się samochodowe pojazdy elektryczne, elektrowozy, elektryczne zespoły trakcyjne, tramwaje, trolejbusy, składy metra, skutery, rowery elektryczne, quady oraz wózki: transportowe, turystyczne, rekreacyjne [Kozłowski i inni, 2011, s. 299].

Norma PN EN 61851-1 elektrycznym pojazdem drogowym (EV) określa wszystkie pojazdy drogowe, łącznie z zawierającymi połączenie wtykowe hybrydowymi pojazdami drogowymi (PHEV), które całkowicie lub częściowo pobierają energię z baterii pokładowych. W dalszej części

pracy skupiono się na samochodowych pojazdach osobowych o napędzie elektrycznym.

Poniżej zaprezentowano dane dotyczące rejestracji elektrycznych samochodów osobowych (EV) w roku 2016. Liderem na rynku europejskim w analizowanym okresie była Norwegia z 29% udziałem pojazdów tego typu w liczbie wszystkich nowo zarejestrowanych samochodów. Kraj ten jest przykładem państwa, w którym zastosowano wiele instrumentów organizacyjnych oraz ekonomicznych zachęt do zakupu pojazdów elektrycznych [Makolska-Tenold, 2015, s. 53, 58–64]. Prawie połowa krajów UE cechowała się co najwyżej 0,5% udziałem w liczbie wszystkich nowo zarejestrowanych samochodów. Najwięcej osobowych samochodów elektrycznych zarejestrowano w Niemczech (0,8% z 3 351 607), Wielkiej Brytanii (1,4% z 2 692 786), Francji (1,4% z 2 015 177), Włoszech (0,2% z 1 824 968) i Hiszpanii (0,3% z 1 147 007). W Polsce udział ten kształtował się na poziomie 0,1% [ACEA, 2017a].

Skalę rejestracji nowych osobowych samochodów elektrycznych przedstawia tablica 1.

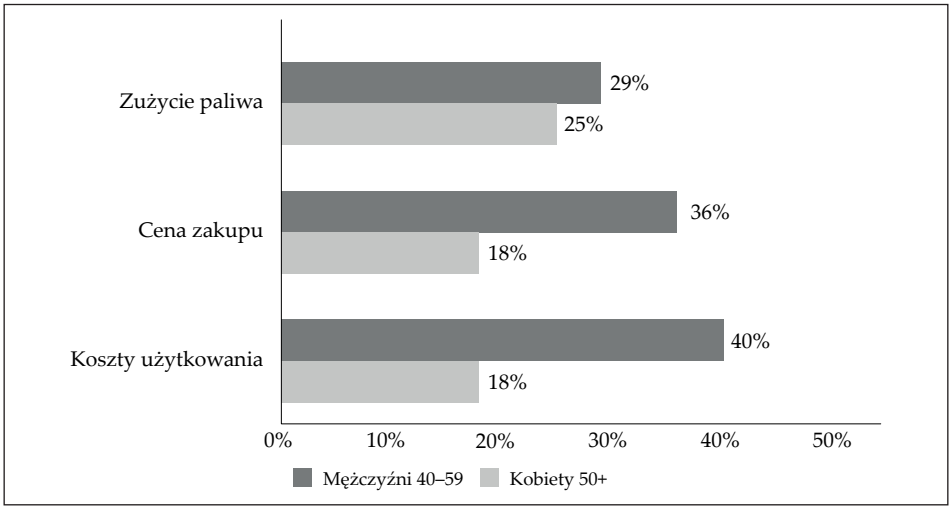
Tablica 1. Rejestracja osobowych samochodów elektrycznych w latach 2015–2017

Rodzaj samochodu	Rok		
	2015	2016	2017
Battery electric vehicles (BEV)	70	108	439
Plug-in hybrid electric vehicles (PHEV)	221	404	585

Źródło: [ACEA, 2016; 2017b].

Przeprowadzone na zlecenie innogy Polska badanie opinii na temat zainteresowania Polaków rynkiem samochodów elektrycznych wskazały, że przy cenie samochodu elektrycznego wyższej od ceny auta spalinowego będzie on najczęściej wybierany przez kobiety w wieku powyżej 50 lat (37% badanych w tej grupie wiekowej) mieszkające w miastach i mężczyzn w przedziale wiekowym 40–59 lat ze średnich i dużych aglomeracji (40% badanych w tej grupie wiekowej). Badanie pokazało dodatkowo, że wraz ze spadkiem ceny pojazdów elektrycznych wzrastać będzie zainteresowanie wśród osób młodszych, z wykształceniem średnim oraz tych zamieszkujących miasta do 100 tys. mieszkańców [*Kto w Polsce chce kupić samochód elektryczny?*, 2017] – rys. 1.

Rysunek 1. Kryteria najczęściej wskazywane przy wyborze samochodu przez grupy najbardziej zainteresowane elektromobilnością



Źródło: [*Kto w Polsce chce kupić samochód elektryczny?*, 2017].

Badania przeprowadzone przez Instytut Transportu Samochodowego nad potencjalnym rozpoznaniem zainteresowania pojazdami elektrycznymi w Polsce prowadzą do następujących wniosków [Waśkiewicz, 2014, s. 10922–10928]:

- brak kampanii informacyjnej, a także praktycznego dostępu do samochodów elektrycznych;
- brak infrastruktury technicznej do ładowania akumulatorów i baterii elektrycznych;
- zbyt wysoka cena i brak dotacji do zakupu samochodów (brak niskooprocentowanych kredytów);
- brak bezpłatnych miejsc parkingowych, oddzielnych pasów ruchu i dostępu do centrum miast tylko dla pojazdów przyjaznych środowisku;
- ok. 67% respondentów wskazuje, że zachętą do kupna samochodów powinny być niskie stawki ubezpieczeniowe.

Rozwiązania części z powyżej wskazanych czynników blokujących rozwój elektromobilności w Polsce przynosi ustawa z dnia 11 stycznia 2018 r. o elektromobilności i paliwach alternatywnych [ustawa, 2018]. Wspomniany akt prawny przewiduje szereg korzyści ekonomicznych, takich jak zwolnienie z akcyzy na zakup osobowych pojazdów elektrycznych, podwyższenie stawek odpisów amortyzacyjnych. Ponadto kierujący pojazdami elektrycznymi uzyskują możliwość poruszania się po pasach drogowych dla autobusów oraz dodatkowe miejsca parkingowe. Gminy

otrzymują podstawę prawną do tworzenia stref czystego transportu przeznaczonych dla pojazdów przyjaznych środowisku. Warto wspomnieć, że art. 2 pkt 12 wspomnianej ustawy zawęża definicję samochodowego pojazdu elektrycznego do pojazdu w rozumieniu art. 2 pkt 33 ustawy z dnia 20 czerwca 1997 r. – Prawo o ruchu drogowym, wykorzystującego do napędu wyłącznie energię elektryczną akumulowaną przez podłączenie do zewnętrznego źródła zasilania[1].

W zakresie infrastruktury dla samochodów elektrycznych celem Ministerstwa Energii jest, aby do roku 2020 w 32 aglomeracjach powstało 6,4 tys. publicznych punktów ładowania i ponad 400 punktów szybkiego ładowania [Polityka Insight, 2017].

Wiele problemów generować może jednak istniejąca obecnie infrastruktura energetyczna, w większości licząca sobie kilkadziesiąt lat i niedostosowana w pełni do obciążenia, jakie będzie generowała masowa konieczność ładowania akumulatorów pojazdów napędzanych prądem.

Warto podkreślić, że rozpowszechnienie pojazdów elektrycznych przyczynia się także do zmniejszenia zewnętrznych kosztów transportu drogowego [Gis i inni, 2013, s. 287–291].

2. Przykłady pojazdów mechanicznych o napędzie elektrycznym

Z danych Instytutu Transportu Samochodowego wynika, że przeciętny samochód w Polsce przejeżdża 8500 km rocznie – przeciętnie 23 km dziennie [*Plan rozwoju elektromobilności w Polsce*, 2016, s. 12].

Opracowany w 2017 r. przez PSPA i Obserwatorium Rynku Paliw Alternatywnych raport prezentuje szczegółowe zestawienie wraz z charakterystykami dostępnych w Polsce pojazdów wykorzystujących tylko i wyłącznie energię elektryczną zmagazynowaną w bateriach. W opracowaniu wymieniono dziewięć samochodów miejskich (Volkswagen e-Up!, Renault TWIZY, Renault ZOE R90 22kWh, Renault ZOE R90 Z.E 40, Renault ZOE Q90 Z.E 40, BMW i3, Smart forfour electric drive, Smart fortwo electric drive, Smart fortwo cabrio electric drive) i cztery samochody kompaktowe (Nissan LEAF 24 kWh, Nissan LEAF 30 kWh, Hyundai IONIQ Electric, Volkswagen e-Golf). Najmniejszym deklarowanym zasięgiem wśród aut w pełni elektrycznych dysponuje Renault TWIZY, które na jednym ładowaniu może pokonać 100 km. Do 160 km, zgodnie z deklaracjami producenta, można przejechać modelami Smarta. Nissan LEAF, w zależności od pojemności akumulatorów, może przejechać od 199 do 250 km, Hyundai IONIQ electric 280 km, Volkswagen e-Golf 300 km, a BMW i3 312 km.

[1] Art. 2 pkt 33 ustawy z dnia 20 czerwca 1997 r. – Prawo o ruchu drogowym przez pojazd samochodowy określa pojazd silnikowy, którego konstrukcja umożliwia jazdę z prędkością przekraczającą 25 km/h.

Nawet ponad 400 km na jednym ładowaniu można pokonać Renault ZOE R90 Z.E. 40 [*Pojazdy elektryczne dostępne w Polsce*, 2017].

W tablicy 2 zaprezentowano maksymalną moc silnika i maksymalną osiąganą prędkość opisywanych pojazdów.

Tablica 2. Parametry silnika i osiągi wybranych pojazdów elektrycznych

Marka/model	Moc maksymalna [KM]	Prędkość maksymalna [km/h]
Renault TWIZY	17	80
Volkswagen e-up	82	130
Smart forfour, Smart fortwo	82	130
Renault ZOE	88/92/92	135
Nissan LEAF	109	144
Hyundai IONIQ electric	120	165
Volkswagen e-Golf	136	150
BMW i3	170	150

Źródło: [*Pojazdy elektryczne dostępne w Polsce*, 2017].

Ceny zaprezentowanych pojazdów wg raportu PSPA i Obserwatorium Rynku Paliw Alternatywnych zaczynają się od 50 tys. zł. (Renault TWIZY). Więcej należy zapłacić za samochody niemieckiego producenta Daimler AG – Smart (ok 100 tys. zł.) i Volkswagena e-up! (ok. 115 tys. zł). Za ok. 130 tys. zł. można nabyć Renault ZOE i Nissana LEAF. Górny pułap to Hyundai IONIQ electric, BMW i3 i Volkswagen e-Golf oferowane za ok. 160 tys. zł [*Pojazdy elektryczne dostępne w Polsce*, 2017].

Warto zwrócić uwagę, że postęp techniczny i technologiczny dotyka także nowych generacji pojazdów elektrycznych [Tronowicz, 2016, s. 14–15]. Możemy tutaj wymienić:

- pojazdy autonomiczne samoprowadzące [Narożny, 2016, s. 61–64],
- samochody z funkcjami inteligentnymi [Miller, 2016, s. 187–211].

3. Różnice między pojazdami o napędzie spalinowym a pojazdami elektrycznymi

W rejonach eksploatacji samochodów elektrycznych wyraźnie poprawia się czystość powietrza, szczególnie w dużych aglomeracjach miejskich. Należy jednak pamiętać, że w warunkach polskich przeważającą metodą produkcji energii elektrycznej jest technologia oparta na spalaniu węgla, która generuje poważne zanieczyszczenie atmosfery. Zatem jeśli nie nastąpi w Polsce znacząca modernizacja sektora energetycznego, konieczność produkcji dodatkowej energii do ładowania akumulatorów przyczyni się do dalszej degradacji obszarów, na których zlokalizowane są elektrownie.

Nie zmienia się zanieczyszczenie spowodowane ogumieniem. Pojazdy elektryczne cechuje większe bezpieczeństwo z uwagi na brak układu paliwowego (zbiornika na paliwo).

Istotną różnicą między samochodami o napędzie spalinowym a pojazdami elektrycznymi są akumulatory. Akumulatory kwasowo-ołowiowe w samochodach elektrycznych zastąpione zostały akumulatorami: NiCd, NiMH, $LiFePO_4$ [Tomczuk, 2009, s. 25–27]. W elektrycznych pojazdach samochodach stosowane są duże akumulatory litowo-jonowe o wysokiej gęstości energetycznej (waga od 250 kg) [Kotowski, 2011, s. 44–45]. Są to instalacje bardzo kosztowne, odporne na niskie i wysokie temperatury (–30°C ÷ 40°C) [Kotowski, Konopka, 2013, s. 17–19].

Akumulatorom do samochodów elektrycznych stawia się takie wymogi, jak krótki czas ładowania, daleki przebieg, długa żywotność, wygoda w eksploatacji. Aby ww. charakterystyki osiągnąć, pracuje się nad wykorzystaniem tzw. materiałów o nadprzewodności elektrycznej, dwukierunkowego systemu ładowania [Gallardo-Lozano i inni, 2011, s. 111–116], odwracalnych ogniw litowych [Świętosławski, 2016, s. 67–72], a także nad bezprzewodowym ładowaniem [*DELPHI – Bezprzewodowe ładowanie samochodów elektrycznych stało się rzeczywistością*, 2010, s. 54; *Bezprzewodowe ładowanie akumulatorów*, 2010, s. 38–40].

Do głównych zalet samochodów elektrycznych zaliczyć możemy:
- zero emisji dwutlenku węgla, brak emisji substancji szkodliwych;
- brak w konstrukcji samochodu: systemów wydechowych, systemów chłodzenia, systemów przekładni wielostopniowych;
- mniejszą awaryjność i dłuższą żywotność (to tylko hipoteza, ze względu na skomplikowaną elektronikę układów nadzorujących pracę silnika i zarządzanie energią akumulatorów);
- brak zjawiska tarcia wewnątrz silnika;
- brak skrzyni biegów;
- możliwość poruszania się po terenach wrażliwych ekologicznie, miastach uzdrowiskowych;
- cichą pracę silnika;
- brak płynów eksploatacyjnych;
- spełnienie rygorystycznej normy emisji spalin Euro 6;
- prostą obsługę serwisową;
- zwiększenie sprawności energetycznej;
- dużo większy moment obrotowy;
- łatwość ładowania, możliwy odzysk energii w czasie hamowania;
- stabilność cen energii, obniżkę kosztów eksploatacji;
- brak zbiornika paliwowego, większe bezpieczeństwo (np. przeciwpożarowe).

W przypadku jazdy samochodem z ciągłym przyspieszaniem, np. w terenach górskich na dużych wysokościach, rozrzedzone powietrze utrudnia pracę silników spalinowych, zaś elektryczne pojazdy są obojętne na te okoliczności. Dodatkowo możliwe jest ładowanie prądem pojazdów z fotowoltaiki i elektrowni wiatrowych.

Samochody elektryczne nie są wolne od wad. Największą z nich jest cena. Wysokie koszty nabycia powodują obniżony popyt. Do innych wad zaliczyć można:

- ograniczoną dostępność;
- stosunkowo mały zasięg na jednym ładowaniu;
- ograniczony zasięg w warunkach niskiej temperatury;
- czas ładowania do kilkunastu godzin;
- brak wiedzy o bezpieczeństwie samochodów w czasie wypadków komunikacyjnych[2];
- za małą moc baterii zasilających;
- wysoki koszt magazynowania baterii.

Mamy do czynienia z nową kulturą jazdy [Teodorczuk, 2011, s. 56–57] i kulturą mobilności. Należy zwrócić uwagę, że obecne modele samochodów w coraz większym stopniu wyposażone są w układy elektroniczne dążące do bezpiecznego prowadzenia pojazdów. Rozwój samochodów idzie w kierunku przejęcia przez układy i systemy elektryczne całego pojazdu [Czarnigowski, 2009, s. 116–120]. To powoduje z jednej strony większy komfort użytkowania, a z drugiej – wymusza ciągłe uzupełnianie wiedzy o unowocześnianych pojazdach, zwłaszcza wśród pracowników serwisów obsługi [Jankowski, 2012, s. 4].

4. Zagrożenia towarzyszące eksploatacji samochodów elektrycznych

Zagrożenia towarzyszące eksploatacji samochodów elektrycznych umownie można podzielić na dwie grupy – naturalne i antropogeniczne. Wśród ekstremalnych zagrożeń naturalnych wyróżnić można takie zjawiska, jak: powódź, wiatr, gołoledź, deszcz nawalny, gradobicie, wyładowania atmosferyczne, osuwiska, promieniowanie naturalne. Pośród zagrożeń antropogenicznych wymienić należy: wypadki, kolizje komunikacyjne, kradzieże, awarie techniczne.

Biorąc pod uwagę powyższe oraz rozważania z poprzednich punktów tego artykułu, zwróćmy uwagę na niektóre – zdaniem autorów – najważniejsze zagrożenia, które mogą towarzyszyć eksploatacji samochodów elektrycznych, wynikające z ich odmiennej budowy i konstrukcji.

[2] Europejskie testy bezpieczeństwa zderzeniowego wykazały porównywalne bezpieczeństwo z samochodami o napędzie spalinowym.

Wysoka cena samochodów elektrycznych może być przyczynkiem do ich kradzieży. Dla celów bezpieczeństwa koniecznością jest wyposażenie ich w najnowsze systemy przeciwkradzieżowe, co wpływa dodatkowo na wzrost ceny zakupu [*Nowoczesne samochody…*, 2002, s. 8–9].

Wspomnieć należy o ryzyku związanym z konstrukcją pojazdów elektrycznych. Wpływ lokalizacji i zabezpieczenia modułów na ich uszkodzenie w trakcie wypadku lub kolizji jest niezwykle istotny. Negatywne skutki mogą różnić się w zależności od siły uderzeniowej, okoliczności wypadku, miejsca zderzenia (np. boczne, czołowe, dachowanie). Równie ważne jest zachowanie się pojazdu elektrycznego w czasie uderzenia mechanicznego.

Biorąc pod uwagę specyfikę konstrukcyjną i eksploatacyjną pojazdu elektrycznego, należy wspomnieć o ryzyku porażenia prądem. Podczas kolizji (wypadku) może dojść do porażenia pasażerów i ratowników medycznych[3]. Istotna jest zatem ochrona przeciwporażeniowa oraz ochrona przed prądem przetężeniowym w instalacjach zasilających stacje ładowania pojazdów elektrycznych [Tymosiak, Sułkowski, 2013, s. 48–50]. Zagadnienie to reguluje norma PN-HD 60364-7-772.2012 – Instalacje elektryczne niskiego napięcia Część 7-722 Wymagania dotyczące specjalnych instalacji lub lokalizacji. Zasilanie pojazdów elektrycznych. Norma przewiduje samoczynne wyłączenie zasilania przy wykorzystaniu wyłącznika różnicowo-prądowego. Opracowane zostały również normy dotyczące zasad bezpieczeństwa w czasie ładowania akumulatorów w samochodach o napędzie elektrycznym [norma PN-HD 60364-7-722, 2012].

W akumulatorach stosowane są roztwory chemiczne szkodliwe dla człowieka i środowiska. W przypadku kolizji (wypadku) z udziałem samochodu elektrycznego ww. roztwory mogą wypłynąć na zewnątrz.

Należy także zwrócić uwagę, że urządzenia elektryczne mogą być narażone na oddziaływanie środowiskowe (deszcz, mróz, śnieg, uderzenia mechaniczne itp.) [norma PN-EN 62262, 2005].

Należy przewidywać, że z chwilą pojawienia się kilkudziesięciu tysięcy pojazdów elektrycznych przy jednoczesnym ładowaniu akumulatorów powstaną problemy z obciążeniem sieci przesyłowych. Działania w tym zakresie powinny promować ładowanie nocne [Majchrzak, 2011, s. 8].

Cicha praca silnika pojazdu elektrycznego może powodować zwiększoną liczbę wypadków z udziałem pieszych.

Część szkód powodowanych ww. zagrożeniami może zostać skompensowana odszkodowaniem w ramach ubezpieczenia autocasco. Część z ww. zagrożeń dotyczy także odpowiedzialności cywilnej posiadacza pojazdu elektrycznego, a nie mieści się w obecnym (ustawowym) rozumieniu

[3] Napięcie baterii akumulatorów wynosi ok. 400 V – niemal dwukrotnie więcej niż w gniazdkach elektrycznych.

ubezpieczenia OC p.p.m. Rodzi to potrzebę rozszerzenia zakresu ochrony i nowych rozwiązań.

5. Ubezpieczenia autocasco. Założenia ogólne

Ubezpieczenie autocasco pokrywa koszty naprawy pojazdu uszkodzonego wskutek wszelkiego rodzaju zdarzeń losowych, a także zapewnia rekompensatę w razie jego całkowitego zniszczenia lub kradzieży. Zakresem autocasco objęte są szkody polegające na uszkodzeniu, zniszczeniu lub utracie pojazdu lub jego części, wskutek zajścia wypadku ubezpieczeniowego w okresie ubezpieczenia. Wypadek ubezpieczeniowy to zdarzenie niezależne od woli ubezpieczonego lub osoby uprawnionej do korzystania z pojazdu powodujące szkodę objętą zakresem ubezpieczenia. Ubezpieczenie autocasco jest produktem dobrowolnym, każdy zakład ubezpieczeń kształtuje warunki ubezpieczenia wg własnych zasad. W szczególności określa zakres odpowiedzialności i czynniki uwzględniane w procesie taryfikacji.

Częścią wspólną ogólnych warunków ubezpieczenia autocasco jest odpowiedzialność za:

- szkody spowodowane siłami mechanicznymi pochodzącymi od innego pojazdu, osoby, zwierzęcia, przedmiotu;
- uszkodzenia spowodowane przez osoby trzecie;
- szkody wywołane zewnętrznymi czynnikami naturalnymi, tj. pożar, wybuch, zatopienie, powódź, zalanie, grad, lawina, huragan, piorun;
- nagłe działania czynnika termicznego lub chemicznego pochodzącego z zewnątrz pojazdu;
- kradzież pojazdu lub jego części.

Z zakresu ubezpieczenia wyłącza się zdarzenia losowe wynikające z przyczyn wewnętrznych pojazdu. Zakres odpowiedzialności nie obejmuje szkód związanych z wadą pojazdu, powstałych z przyczyn eksploatacyjnych, będących konsekwencją wykonanych napraw.

Zakłady ubezpieczeń podkreślają, że warunkiem koniecznym uznania szkody jest istnienie przyczyny zewnętrznej. Samoistny charakter uszkodzenia wyklucza pokrycie szkody. Warto podkreślić, że w przypadku pojazdów elektrycznych występuje ryzyko pożaru w wyniku zwarcia instalacji w pojeździe.

Katalog wyłączeń jest dużo szerszy. W ogólnych warunkach ubezpieczenia powtarza się zapis o braku odpowiedzialności zakładu ubezpieczeń w przypadku niefrasobliwość i rażącego niedbalstwa.

W przypadku zajścia zdarzenia losowego ujętego w OWU zakład ubezpieczeń wypłaca odszkodowanie w granicach sumy ubezpieczenia odpowiednio:

– w kwocie pieniężnej koniecznej do przywrócenia pojazdu mechanicznego do stanu, w jakim znajdował się on przed zajściem wypadku,
– w kwocie będącej równowartością wartości pojazdu, jeżeli pojazd został zniszczony całkowicie lub skradziony.

Suma ubezpieczenia powinna odzwierciedlać wartość przedmiotu ubezpieczenia. Najdroższą częścią samochodów elektrycznych są silniki elektryczne i akumulatory gromadzące prąd elektryczny. Akumulatory najmniej tracą na wartości, jeśli są ciągle doładowywane. Są one swoistym zasobnikiem energii [Fice, Setlak, 2013, s. 25–38]. Akumulatory w samochodach elektrycznych są narażone na głębokie rozładowania, co zmniejsza ich trwałość. Z tego względu pojazdy elektryczne charakteryzuje odmienna niż spalinowe podatność na utratę wartości w funkcji czasu.

W przypadku ubezpieczenia autocasco jednym z czynników wpływających na wysokość składki za ubezpieczenie jest koszt części i napraw pojazdu. Mała popularność samochodów elektrycznych skutkuje brakiem powszechnej dostępności niektórych części, brak jest również tańszych zamienników (oryginalne części muszą być sprowadzane). Dodatkowo napraw podejmują się tylko nieliczne warsztaty. Taki stan rzeczy skutkuje większymi składkami za ubezpieczenie.

Wachlarz zdarzeń objętych ochroną ubezpieczeniową w ramach autocasco jest różny w poszczególnych zakładach ubezpieczeń. W przypadku pojazdów elektrycznych powinien jednak uwzględniać odmienne niż w przypadku aut spalinowych zakresy uszkodzeń, technologii napraw i wynikający z tego poziom kosztów naprawy podzespołów. Może się zdarzyć, że miejsca umieszczenia przez projektantów przewodów elektrycznych nie zawsze są optymalne i narażają je na liczne uszkodzenia [Trawiński, Dyga, 2015, s. 5]. Specyficzny charakter pojazdów elektrycznych rodzi także problem wyceny pozostałości przy szkodzie całkowitej, a przepisy z zakresu ekologii utrudnią zbyt uszkodzonych pojazdów na aukcjach i giełdach.

Zakończenie

Pojazdy elektryczne mają wiele zalet, charakteryzują się wysoką sprawnością, ich silniki pracują cicho i płynnie, mogą odzyskiwać energię podczas hamowania, nie emitują szkodliwych substancji podczas jazdy. Mimo że w najbliższej dekadzie samochody z silnikiem wewnętrznego spalania będą dominowały, udział samochodów elektrycznych będzie wzrastał [*Elektromobilność*, 2011, s. 3].

Prognozowany wzrost udziału samochodów elektrycznych w polskim parku samochodowym będzie stanowił wyzwanie dla zakładów ubezpieczeń. Zmiana struktury jeżdżących po polskich drogach pojazdów wpłynie

na proces wyceny ubezpieczenia. Zakłady ubezpieczeń będą dostosowywać proces taryfikacji i zakres ochrony ubezpieczeniowej. Na rynkach europejskich, w duchu wspierania proekologicznych zachowań, można spotkać się z praktyką udzielania specjalnych zniżek na pakiety ubezpieczeń dla posiadaczy pojazdów elektrycznych. Jak pokazują przytoczone wcześniej badania, 67% respondentów wskazuje, że zachętą do kupna samochodów powinny być niskie stawki ubezpieczeniowe. Być może ubezpieczyciele w Polsce również zastosują takie podejście.

Ograniczone ramy artykułu spowodowały, że podjęty temat został jedynie zasygnalizowany i na pewno nie został wyczerpany. Bardzo ważne wydaje się podjęcie badań nad skutkami kolizji i wypadków z udziałem pojazdów elektrycznych, a także nad kosztami napraw.

Literatura

ACEA (2017a), *Correlation between uptake of electric cars and GDP in the EU*, http://www.acea.be /statistics/article/interactive-map-correlation-between-uptake--of-electric-cars-and-gdp-in-EU.

ACEA (2017b), *New passenger car registrations by market in the EU+EFTA*, http://www.acea.be/statistics/tag/category/passenger-cars-registrations.

ACEA (2016), *New passenger car registrations by market in the EU+EFTA*, http://www.acea.be/statistics/tag/category/passenger-cars-registrations.

Bezprzewodowe ładowanie akumulatorów (2010), „Autonaprawa", nr 11.

Burns L.D. (red.) (2013), *Transforming Personal Mobility*, red. L.D., The Earth Institute, Columbia University.

Czarnigowski J. (2009), *Przyszłość samochodów jest elektryczna*, „AutoElektro", nr 2.

Deloitte Insights (2018), *Great expectations. Insights exploring new automotive business models and consumer preferences.*

DELPHI – Bezprzewodowe ładowanie samochodów elektrycznych stało się rzeczywistością (2010), „AutoElektro", nr 12 (122).

Elektromobilność (2011), Narodowe Centrum Badań i Rozwoju.

Fice M., Setlak R. (2013), *Samochód elektryczny jako zasobnik dla instalacji prosumenckiej*, „Elektryka", z. 4.

Gallardo-Lozano J. i inni (2011), *Non-disturbing bidirectional charger for PHEVs and EVs*, „Przegląd Elektrotechniczny", nr 12.

Gis W. i inni (2013), *Cars with electric drive and external costs of road transport*, „Przegląd Elektrotechniczny", nr 10.

Jankowski M. (2012), *Serwisowanie napędów elektrycznych i hybrydowych*, „AutoElektro", nr 3 (137).

KPMG (2018), *KPMG's 19th consecutive Global Automotive Executive Survey*.

Kotowski W. (2011), *Rozwój akumulatorów dla elektrycznego napędu samochodów*, „Czysta Energia", nr 7–8.

Kotowski W., Konopka E. (2013), *Nowe akumulatory*, „Nowa Energia", nr 11.

Kozłowski M. i inni (2011), *Methodology of determining basic technical parameters of electric-drive car*, „Przegląd Elektrotechniczny", nr 10.

Kto w Polsce chce kupić samochód elektryczny? (2017), http://swiatoze.pl/polsce-chce-kupic-samochod-elektryczny.

Majchrzak H. (2011), *Samochód elektryczny – szanse i zagrożenia*, „Czysta Energia", nr 5.

Makolska-Tenold M. (2015), *Samochód elektryczny ekologicznym rozwiązaniem transportowym zielonego miasta*, „Logistyka", nr 3.

Miller M. (2016), *Internet rzeczy. Jak inteligentne telewizory, samochody, domy i miasta zmieniają świat*, Wydawnictwo Naukowe PWN, Warszawa.

Narożny J. (2016), *Pojazdy samoprowadzące – nowa era przyszłości motoryzacji czy kolejny gadżet?* (2016), „Magazyn Autostrady", nr 6.

Nowoczesne samochody a zabezpieczenia przeciwkradzieżowe firmy DEST (2002), „Ochrona Mienia" nr 6.

Norma PN-HD 60364-7-722 Instalacje elektryczne niskiego napięcia – Część 7-722: Wymagania dotyczące specjalnych instalacji lub lokalizacji. Zasilanie pojazdów elektrycznych.

Norma PN-EN 62196-1 Wtyczki, gniazda wtykowe, złącza pojazdowe i wtyki pojazdowe. Przewodowe ładowanie pojazdów elektrycznych, Część 1, Wymagania ogólne.

Norma PN-EN 62262;2005 Stopnie ochrony przed zewnętrznymi uderzeniami mechanicznymi zapewnianej przez obudowy urządzeń elektrycznych.

Plan rozwoju elektromobilności w Polsce (2016), Ministerstwo Infrastruktury.

Polityka Insight (2017), *Cicha rewolucja w energetyce*, https://www.politykainsight.pl/_resource/multimedium/20106685.

Pojazdy elektryczne dostępne w Polsce (2017), www.orpa.pl/elektromobilnipl.

PWC (2016), *Re-inventing the wheel Scenarios for the transformation of the automotive industry*.

PWC (2017), *Five trends transforming the Automotive Industry*.

Staniszewski P. (2015), *Z napędem elektrycznym*, „Tygodnik Poradnik Rolniczy", nr 31–32.

Świętosławski M. (2016), *Trends of development of rechargeable lithium batteries*, „Przemysł Chemiczny", nr 1.

Teodorczuk K. (2011), *Samochód elektryczny – norma kultura mobilności*, „Czysta Energia", nr 11.

Tomczuk K. (2009), *Akumulatorowe zasobniki energii dla pojazdów elektrycznych*, „Nowa Elektrotechnika", nr 6.

Trawiński G., Dyga G. (2015), *Schematy instalacji elektrycznej pojazdów – zasady „czytania"*, „AutoElektro", nr 12 (182).

Tronowicz M. (2016), *Samochody na prąd coraz popularniejsze oraz doskonalsze*, „Elektronik", nr 7.

Tymosiak P., Sułkowski M.A. (2013), *Stacje ładowania źródeł energii pojazdów elektrycznych. Wymagania w zakresie instalacji elektrycznych niskiego napięcia*, „Elektroinfo", nr 6.

Ustawa z dnia 11 stycznia 2018 r. o elektromobilności i paliwach alternatywnych (Dz.U. poz. 317).

Waśkiewicz J. (2014), *Rozwój samochodów elektrycznych w Polsce według opinii przyszłych użytkowników*, „Logistyka", nr 6.

Streszczenie

Już wiele lat temu dostrzeżono, że rozwój motoryzacji na bazie silników benzynowych i oleju napędowego przyczynia się do emisji wielu szkodliwych dla środowiska substancji. Emisja dwutlenku węgla powoduje stopniowe ocieplanie się klimatu wraz ze wszystkimi negatywnymi konsekwencjami tego zjawiska. Aby powstrzymać szkodliwą emisję, dąży się do ograniczenia substancji produkowanych przez pojazdy samochodowe. Jednym ze sposobów jest rozwój produkcji pojazdów mechanicznych o napędzie elektrycznym. Celem artykułu jest identyfikacja pola ubezpieczeniowego i ocena wybranych aspektów ryzyka związanego z eksploatacją pojazdów mechanicznych o napędzie elektrycznym w kontekście ubezpieczenia autocasco. W związku z różnicami związanymi ze specyfiką eksploatacji tego typu pojazdów proces dostosowania będzie potrzebny.

Słowa kluczowe

ubezpieczenie autocasco, pojazdy elektryczne

Electric vehicle market characterization and risk related to use of electric vehicles in motor own damage insurance (Summary)

It has been a long time ago since the development of petrol and diesel-powered motor vehicles has been recognized as a source of the toxic emissions. Global warming is primarily caused by the burning of fossil fuels that pump carbon dioxide (CO2). The gradual heating of Earth's surface, oceans and atmosphere requires very deep cuts in toxic emissions. One of the directions is to develop electric vehicle (EV) production. The aim of the article is to present electric vehicles market review, development forecasts and to assess risk related to use of electric vehicles in motor own damage insurance. Due to the EV's specifics of construction and operation, insurers will probably adjust their tariffs and general insurance conditions.

Keywords

motor own damage insurance, electric vehicles

Agnieszka Pobłocka*

Tworzenie rezerwy IBNR metodami deterministycznymi na potrzeby wypłacalności w zakładach ubezpieczeń majątkowych

Wstęp

Zakłady ubezpieczeń prowadzące działalność ubezpieczeniową na terytorium Rzeczypospolitej Polskiej są zobligowane ustawowo do tworzenia rezerw techniczno-ubezpieczeniowych, które są zobowiązaniami o nieznanym terminie zapadalności lub nieustalonej kwocie. Rozporządzenie Ministra Finansów z dnia 12 kwietnia 2016 r. w sprawie szczególnych zasad rachunkowości zakładów ubezpieczeń i zakładów reasekuracji określa, że rezerwy techniczno-ubezpieczeniowe (tworzone do celów rachunkowości) „przeznaczone są na bieżące i przyszłe zobowiązania, jakie mogą wyniknąć z zawartych umów ubezpieczenia i umów reasekuracji czynnej" [rozporządzenie, 2016, art. 30 § 1]. Ustawa z dnia 15 września 2015 r. o działalności ubezpieczeniowej i reasekuracyjnej [ustawa, 2015], zaimplementowana z dniem 1 stycznia 2016 r., wprowadziła m.in. nowe zasady sprawozdawczości do celów wypłacalności; odrębne zasady rachunkowości do celów statutowych i do celów wypłacalności. W efekcie wdrożenia nowych norm ubezpieczeniowych powstał od stycznia 2016 r. „dualizm" zasad wyceny aktywów i zobowiązań [Kędziora, 2016, s. 83; Spigarska, 2014, s. 39]. To znaczy, że rezerwy techniczno-ubezpieczeniowe tworzone są, tak jak to było wcześniej, jednocześnie na potrzeby rachunkowości i sprawozdawczości statutowej (wg wartości księgowej i wytycznych prawa bilansowego, tj. ustawy z dnia 29 września 1994 r. o rachunkowości [ustawa, 1994]) oraz na potrzeby wypłacalności[1] (wg wartości rynkowej[2] i wytycznych art. 224–232 ustawy o działalności ubezpieczeniowej oraz Dyrektywy 2009/138/WE Parlamentu Europejskiego i Rady z dnia 25 listopada w sprawie podejmowania i wykonywania działalności ubezpieczeniowej i reasekuracyjnej – Wypłacalność II [dyrektywa, 2009]).

* Dr, Katedra Statystyki, Wydział Zarządzania, Uniwersytet Gdański, ul. Armii Krajowej 101, 81-824 Sopot, agnieszka.poblocka@ug.edu.pl

[1] To nowe podejście i po raz pierwszy w ten sposób są kalkulowane rezerwy techniczno-ubezpieczeniowe.

[2] Tak zwanej wartości godziwej. Art. 224 ust. 3 ustawy o działalności ubezpieczeniowej określa, że rezerwy techniczno-ubezpieczeniowe wycenia się według wartości godziwej równej bieżącej kwocie, którą zakład ubezpieczeń lub zakład reasekuracji zapłaciłby przy natychmiastowym przeniesieniu zobowiązań z tytułu zawartych umów ubezpieczenia na inny zakład ubezpieczeń lub reasekuracji.

Wycena aktywów i zobowiązań według wartości godziwej jest nową ideą w rachunkowości zakładów ubezpieczeń w zakresie oceny wypłacalności[3], gdyż do końca 2015 r. takiej wyceny dokonywano tylko na podstawie wartości księgowej zdefiniowanej w ustawie o rachunkowości [Kędziora, 2016, s. 83]. Oznacza to, że tworzenie rezerw techniczno-ubezpieczeniowych do celów wypłacalności jest nowym wyzwaniem dla zakładów ubezpieczeń, które wiąże się jednocześnie z nowymi możliwościami oraz zagrożeniami [Spigarska, 2014, s. 39].

Głównymi celami artykułu są przedstawienie zasad wyceny rezerw techniczno-ubezpieczeniowych według Dyrektywy Wypłacalność II oraz wskazanie możliwości i zagrożeń dla zakładów ubezpieczeń wynikających z braku precyzyjnego określenia zasad wyceny rezerw techniczno-ubezpieczeniowych według wartości rynkowych. W tym celu zaprezentowano i zaimplementowano trzy deterministyczne metody szacowania rezerwy z tytułu zaistniałych szkód niezgłoszonych zakładowi ubezpieczeń do dnia tworzenia rezerw techniczno-ubezpieczeniowych, które spełniają wymogi wypłacalności w zakładach ubezpieczeń majątkowych według wytycznych Dyrektywy Wypłacalność II. Metody badawcze użyte w pracy to krytyczna analiza literatury przedmiotu oraz metody opisu statystycznego. Dane statystyczne użyte w badaniu to materiał wtórny publikowany przez Komisję Nadzoru Finansowego (zagregowane dane finansowe dla rynku ubezpieczeń) i Główny Urząd Statystyczny (roczne wskaźniki makroekonomiczne) oraz przez Powszechny Zakład Ubezpieczeń S.A. (*Sprawozdanie o wypłacalności i kondycji finansowej na dzień 31 grudnia 2016 r. i za rok obrotowy zakończony tego dnia*[4]).

1. Rezerwy techniczno-ubezpieczeniowe na potrzeby rachunkowości

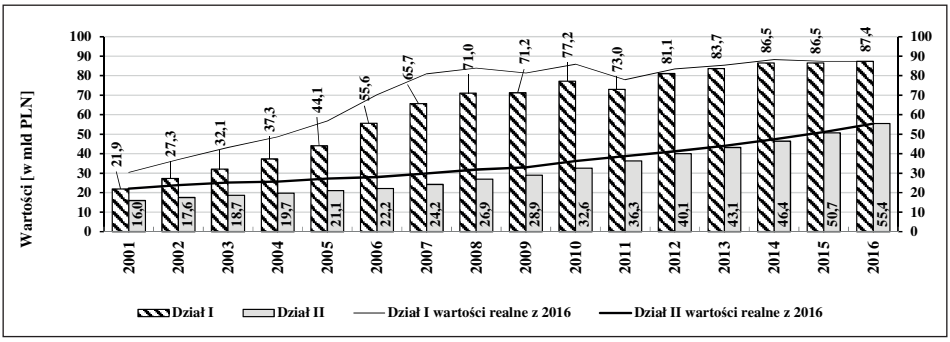
Rezerwy techniczno-ubezpieczeniowe (rezerwy TU) zaliczane są wg Polskich Standardów Rachunkowości (PSR) do kosztów uzyskania przychodów. Wpływają one na wyniki finansowe zakładów ubezpieczeń i zakładów reasekuracji (zmniejszają wynik techniczny, gdy zmiany stanu rezerw są dodatnie, lub zwiększają wynik techniczny, gdy zmiany stanu rezerw są ujemne). Ich poziom jest znaczący i wpływa na gospodarkę finansową zakładów ubezpieczeń i zakładów reasekuracji oraz gospodarkę całego kraju.

[3] Od 2001 r. zakłady ubezpieczeń i zakłady reasekuracji, na skutek implementacji Międzynarodowych Standardów Rachunkowości, mogły stosować wyceny określonych składników majątku według wartości godziwej tylko w zakresie działalności lokacyjnej (MSR 32 i 39).

[4] Sprawozdanie to jest pierwszym (publikowanym w 2017 r. według wytycznych Dyrektywy Wypłacalność II i ustawy o działalności ubezpieczeniowej) i jedynym ogólnodostępnym sprawozdaniem finansowym zakładu ubezpieczeń.

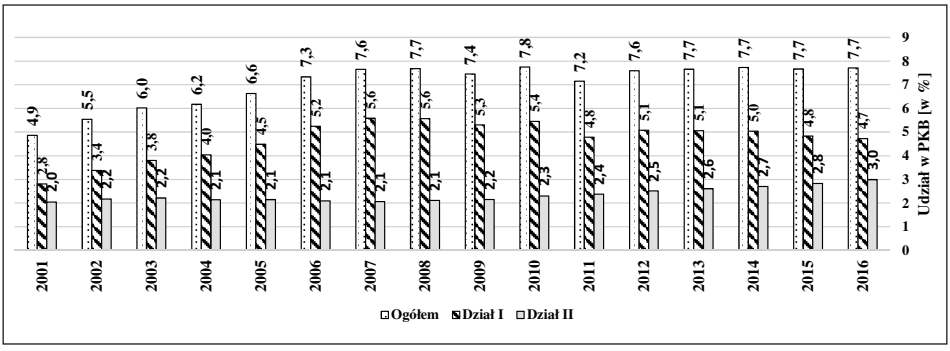
W latach 2001–2016 wartość rezerw techniczno-ubezpieczeniowych brutto tworzonych w zakładach ubezpieczeń ogółem prowadzących działalność ubezpieczeniową na terytorium Polski zmieniała się w czasie (rys. 1) i miała istotny udział w Produkcie Krajowym Brutto (rys. 2). Przykładowo na dzień 31.12.2016 r. wartość rezerw TU kształtowała się na poziomie 142,7 mld PLN (87,4 mld PLN dział I ubezpieczeń; 55,3 mld PLN dział II ubezpieczeń), co stanowiło 7,7% PKB (4,7% PKB dział I ubezpieczeń; 3,0% PKB dział II ubezpieczeń).

Rysunek 1. Rezerwy techniczno-ubezpieczeniowe brutto w Polsce według działów ubezpieczeń, w wartościach nominalnych i realnych z 2016 r. w latach 2001–2016 [w mld PLN]



Dział I to ubezpieczenia na życie, Dział II to pozostałe ubezpieczenia osobowe oraz ubezpieczenia majątkowe.
Źródło: Opracowanie własne na podstawie danych KNF [*Zagregowane dane finansowe. Rocznik rynku ubezpieczeń w latach 2000–2016*] oraz GUS [*Roczne wskaźniki makroekonomiczne w latach 2000–2016*], dostęp: 30.09.2017.

Rysunek 2. Relacja rezerw techniczno-ubezpieczeniowych brutto w PKB w latach 2001–2016 [udział w %]



Źródło: Opracowanie własne na podstawie [*Zagregowane dane finansowe. Rocznik rynku ubezpieczeń 2000–2016*; *Roczne wskaźniki makroekonomiczne*], dostęp: 30.09.2017.

Od wielkości utworzonych rezerw techniczno-ubezpieczeniowych zależy wielkość odprowadzanego do Skarbu Państwa podatku dochodowego. Kontrolę nad prawidłową wielkością tworzonych rezerw TU sprawuje w Polsce [ustawa, 2015, art. 329] centralny organ nadzoru ubezpieczeń – Komisja Nadzoru Finansowego (KNF). Jeżeli w danym zakładzie ubezpieczeń rezerwy TU zostaną utworzone na zbyt niskim poziomie i zabraknie środków na wypłatę zobowiązań z tytułu zawartych umów ubezpieczeń, to zakład ten może otrzymać karę od KNF za brak wywiązania się z zobowiązań ubezpieczeniowych. Kary takie są niemałe, np. z komunikatu z 367 posiedzenia KNF w dniu 29.08.2017 r. [komunikaty, 2017] wynika, że za naruszenie art. 18 ust. 2 ww. ustawy ubezpieczeniowej (poprzez ustalenie składki ubezpieczeniowej w wysokości niezapewniającej wykonania co najmniej wszystkich zobowiązań z umów ubezpieczenia i pokrycia kosztów działalności w ubezpieczeniach grupy 3 (ubezpieczeniach casco pojazdów lądowych, z wyjątkiem pojazdów szynowych, obejmujące szkody w pojazdach samochodowych) oraz grupy 10 (ubezpieczenia odpowiedzialności cywilnej posiadaczy pojazdów mechanicznych z wyłączeniem odpowiedzialności przewoźnika)) nałożono na zakład ubezpieczeń X karę w wysokości 260 tys. PLN, a na zakład ubezpieczeń Y karę w wysokości 600 tys. PLN.

Rezerwy techniczno-ubezpieczeniowe, z wyjątkiem rezerw na wyrównanie szkodowości, ustala się metodą aktuarialną lub inną niż aktuarialna (tj. indywidualną lub ryczałtową) nie później niż na dzień bilansowy [rozporządzenie, 2016, art. 33 § 1][5]. Stopy techniczne stosowane przez zakład ubezpieczeń przy ustalaniu rezerw TU spełniają dwa warunki (art. 46 § 1). Pierwszy, to stopy techniczne nie mogą być wyższe niż 80% średniej ważonej stopy zwrotu z lokat stanowiących pokrycie rezerw TU w ostatnich 3 latach obrotowych, obliczonej metodą do rozporządzenia. W przypadku zobowiązań z umów ubezpieczenia ustalanych w walucie obcej uwzględnia się stopę zwrotu z lokat dokonanych w tej walucie. Wysokość średniej ważonej stopy zwrotu z lokat stanowiących pokrycie rezerw techniczno-ubezpieczeniowych w ostatnich 3 latach obrotowych poprzedzających dany rok obrotowy (S) oblicza się (wg załącznika do rozporządzenia MF z 2016 r.) jako wypadkową z ostatnich trzech lat stóp zwrotu stanowiących pokrycie rezerw techniczno-ubezpieczeniowych według wzoru:

$$K = 50\% \times K1 + 30\% \times K2 + 20\% \times K3 \quad (1)$$

gdzie: K1, K2, K3 to stopy zwrotu z lokat stanowiących pokrycie rezerw: K1 odnosi się do roku obrotowego S–1, K2 do roku S–2, K3 do roku S–3.

[5] Rezerwę na wyrównanie szkodowości ustala się nie później niż na dzień kończący rok obrotowy.

Jeżeli zakład ubezpieczeń wykonuje działalność krócej niż 3 lata, to dla okresu dwuletniego wagi wynoszą 65% dla roku ostatniego i 35% dla przedostatniego.

Drugi warunek (art. 46 § 2–7) dla stóp technicznych jest taki, że mogą być one niższe niż maksymalna stopa techniczna obowiązująca w dniu zawierania umowy ogłaszana przez organ nadzoru. W Polsce Urząd KNF publikuje obowiązującą w danym roku stopę techniczną (tab. 1), której wysokość (J) w danym roku obrotowym (R) oblicza się (wg załącznika do rozporządzenia MF z 2016 r.) według wzoru:

$$J = 60\% \times (50\% \times I1 + 30\% \times I2 + 20\% \times I3) \tag{2}$$

gdzie: I1, I2, I3 to średnie ważone roczne rentowności obligacji Skarbu Państwa o stałym oprocentowaniu, z okresem wykupu nie krótszym niż 8 lat, uzyskane na rynku pierwotnym, w ostatnich 3 latach (I1 oznacza rentowność w roku R–1, I2 rentowność w roku R–2, a I3 rentowność w roku R–3).

Tablica 1. Maksymalna stopa techniczna publikowana przez Urząd Komisji Nadzoru Finansowego w Polsce w latach 2007–2017

Okres obowiązywania danej maksymalnej stopy technicznej	Wartość
Od 1 maja 2007 r. do 30 kwietnia 2008 r.	3,33%
Od 1 maja 2008 r. do 30 kwietnia 2009 r.	3,21%
Od 1 maja 2009 r. do 30 kwietnia 2010 r.	3,38%
Od 1 maja 2010 r. do 30 kwietnia 2011 r.	3,54%
Od 1 maja 2011 r. do 30 kwietnia 2012 r.	3,52%
Od 1 maja 2012 r. do 30 kwietnia 2013 r.	3,54%
Od 1 maja 2013 r. do 30 kwietnia 2014 r.	3,28%
Od 1 maja 2014 r. do 30 kwietnia 2015 r.	2,80%
Od 1 maja 2015 r. do 30 kwietnia 2016 r.	2,42%
Od 1 maja 2016 r. do 30 kwietnia 2017 r.	1,94%
Od 1 maja 2017 r. do 31 kwietnia 2018 r.	1,83%

Źródło: Opracowanie własne na podstawie dzienników Urzędu KNF z lat 2007–2017.

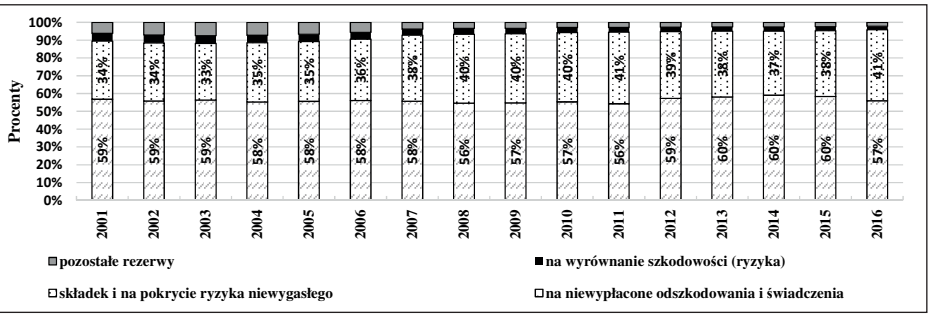
Reasumując, stopy techniczne stosowane przez zakłady ubezpieczeń nie mogą być wyższe niż 80% średniej ważonej stopy zwrotu uzyskanej z wzoru (1) i nie mogą być wyższe niż maksymalna stopa techniczna oszacowana według wzoru (2). W przypadku gdy 80% stopy technicznej wyznaczonej z wzoru (1) jest wyższa niż maksymalna stopa techniczna wyznaczona z wzoru (2), to przyjmuje się stopy techniczne na poziomie maksymalnej stopy technicznej publikowanej przez organ nadzoru. Jednak czy takie rozwiązanie wystarczy, aby zakłady ubezpieczeń i reasekuracji na

„bezpiecznym poziomie" szacowały rezerwy TU oraz składki ubezpieczeniowe? To znaczy tak, aby utworzone rezerwy pokryły bieżące i przyszłe zobowiązania wynikające z zawartych umów ubezpieczenia oraz koszty działalności ubezpieczeniowej?

2. Rezerwa IBNR i metody jej szacowania

W latach 2001–2016 największy udział w rezerwach techniczno-ubezpieczeniowych działu II ubezpieczeń miała rezerwa na niewypłacone odszkodowania i świadczenia (rys. 3). Obejmuje ona trzy grupy rezerw z tytułu szkód: 1) zgłoszonych i oszacowanych, gdy posiadane informacje pozwalają na ocenę wysokości odszkodowań i świadczeń; 2) zgłoszonych, lecz posiadane informacje nie pozwalają na ocenę wysokości odszkodowań i świadczeń; 3) zaistniałych, lecz niezgłoszonych do dnia tworzenia rezerw [rozporządzenie, 2016, art. 36 § 1]. Najważniejszą z nich (i najważniejszą według autora w ubezpieczeniach majątkowych) jest rezerwa z tytułu szkód zaistniałych i niezgłoszonych do dnia tworzenia rezerw, tzw. rezerwa IBNR (*Incurred But Not Reported*)[6]. Jej wartość wpływa znacząco na cały poziom rezerw TU, a jej niedoszacowanie może przyczynić się do niewypłacalności zakładów ubezpieczeń. Z tego powodu wycena rezerwy IBNR na potrzeby rachunkowości (wg prawa bilansowego) i na potrzeby wypłacalności (wg Dyrektywy Wypłacalność II) jest tematem rozważań w niniejszej pracy.

Rysunek 3. Struktura rezerw techniczno-ubezpieczeniowych brutto w dziale II ubezpieczeń w Polsce według rodzajów rezerw w latach 2001–2016 [w %]



Pozostałe rezerwy to suma rezerwy na premie i rabaty dla ubezpieczonych, rezerwy na zwrot składek dla członków oraz pozostałych rezerw techniczno-ubezpieczeniowych określonych w statucie przedsiębiorstwa.

Źródło: Opracowanie własne na podstawie [*Zagregowane dane finansowe. Rocznik rynku ubezpieczeń 2001–2016*], dostęp: 30.09.2017.

[6] Moment wystąpienia, liczba, a także wielkość szkód z tytułu rezerwy IBNR są nieznane zakładom ubezpieczeń w chwili tworzenia rezerw ubezpieczeniowych. Z tego powodu rezerwa IBNR jest utożsamiana z największym ryzykiem niedoszacowania rezerw techniczno-ubezpieczeniowych i niewypłacalnością zakładów ubezpieczeń majątkowych.

Rezerwa IBNR może być szacowana dla różnych grup ubezpieczeniowych (lub linii biznesu) różnymi metodami, jednak „w wysokości odpowiadającej ustalonej lub przewidywanej ostatecznej wartości przyszłych wypłat odszkodowań i świadczeń związanych z zaistniałymi, do dnia, na który ustala się rezerwę, szkodami, powiększonej o koszty likwidacji szkód" [rozporządzenie, 2016, art. 36 § 1). Wiele metod kalkulacji rezerwy IBNR opiera się na historycznych danych szkodowych prezentowanych w macierzach szkód według okresów wypadkowych (wystąpienia lub zgłoszenia szkody) i rozliczeniowych (pomiędzy wystąpieniem bądź zgłoszeniem szkody, a jej zlikwidowaniem, tzw. okresów rozwoju szkody lub opóźnień w wypłacie odszkodowania) za pomocą tzw. trójkątów rozliczania szkód (*the run of triangle*). Rzeczywisty trójkąt szkód wypłaconych prezentuje tablica 2.

Tablica 2. Wypłacone odszkodowania i świadczenia z tytułu pozostałych ubezpieczeń osobowych i majątkowych w latach 2007–2016 [w tys. PLN]

Okres wypadkowy (i)	Okres rozliczeniowy szkody (okres opóźnienia w wypłacie odszkodowania) (j)									
Lata	0	1	2	3	4	5	6	7	8	9
2007	2 849 575	812 777	181 813	73 875	48 451	36 935	30 758	34 850	39 293	32 510
2008	2 962 018	921 431	145 938	75 528	43 091	39 546	35 705	41 779	33 184	
2009	3 187 617	777 862	105 337	70 484	58 357	41 046	39 102	29 116		
2010	3 656 532	1 082 632	167 835	124 897	61 176	49 732	58 005			
2011	2 954 518	679 211	137 124	86 133	58 728	52 411				
2012	2 798 833	669 697	123 736	84 337	68 451					
2013	2 514 308	627 689	130 608	88 366						
2014	2 591 302	673 671	152 472							
2015	2 919 700	1 017 734								
2016	3 592 983									

Źródło: [Sprawozdanie…, 2017].

2.1. Metoda średnich wartości współczynników rozliczania szkód

Metoda średnich wartości (SW) zakłada, że przyszłe szkody zależą od danych historycznych, są proporcjonalne do szkód z okresów opóźnienia w rozliczaniu szkód. Została ona opisana m.in. przez [Wieteska, 2004, s. 94–95; Pobłocka, 2012, s. 66].

Metoda SW bazuje na skumulowanych danych szkodowych $\lfloor C_{i,j} \rfloor = \{C_{i,j}: i = 0, 1,\ldots, n–1;\ j = 0, 1,\ldots, n – i – 1\}$, z których wyznacza się współczynniki rozwoju szkód (*claims development factors*) wg wzoru:

$$f_{i,j} = C_{i,j} \div C_i \text{ dla } i, j = 0, 1, 2,\ldots, n–1 \tag{3}$$

i średnie współczynniki rozliczania szkód wyznacza się wg wzoru:

$$\tilde{f}_j = \left(\sum_{i=0}^{n-j-1} f_{i,j} \right) \div \left(n-j-1 \right) \text{ dla } i, j = 0,1,2,...,n-1 \tag{4}$$

gdzie: $C_{i,j}$ to skumulowane płatności (wypłacone odszkodowania i świadczenia) z tytułu szkód zgłoszonych w *i*-tym okresie wypadkowym i wypłaconych z opóźnieniem o *j*-okresów; C_i to skumulowane płatności z tytułu wszystkich szkód zgłoszonych w *i*-tym okresie wypadkowym (wartość ta jest nieznana i musi być szacowana, np. $\hat{C}_i = C_{i,\,n-i-1}$); n to liczba okresów badanych.

Skumulowane przyszłe płatności oblicza się wg wzoru:

$$\hat{C}_{i,j} = C_{i,\,n-i-1} \div \tilde{f}_{n-i-1} \cdot \tilde{f}_j \tag{5}$$

gdzie: $C_{i,\,n-i-1}$ to skumulowane płatności wypłacone do dnia tworzenia rezerwy (dane z przekątnej w macierzy szkód); $\tilde{f}_{n-i-1}$, $\tilde{f}_j$ to średnie współczynniki rozliczania szkód.

Łączną rezerwę IBNR wyznacza się według wzoru:

$$\hat{R} = \sum_{i=0}^{n-1} \hat{R}_i = \sum_{i=0}^{n-1} \left(\hat{C}_{i,n-1} - C_{i,n-i-1} \right) \tag{6}$$

gdzie: $\hat{C}_{i,n-1}$ to oszacowane skumulowane płatności z tytułu szkód zgłoszonych w *i*-tym okresie wypadkowym i wypłaconych z opóźnieniem o „*n*–1" okresów, $C_{i,\,n-i-1}$ to skumulowane płatności z *i*-go okresu wypadkowego wypłacone z opóźnieniem o „*n*–*i*–1" okresów (dane na przekątnej w trójkącie szkód, czyli dane wypłacone do dnia tworzenia rezerwy).

2.2. Metoda chain ladder

Metoda chain ladder (CL) to jedna z najstarszych i najbardziej popularnych na świecie metod kalkulacji rezerwy IBNR. Zakłada ona, że przyszłe szkody zależą od danych historycznych i są proporcjonalne do szkód z okresów opóźnienia w rozliczaniu szkód. Została ona opisana m.in. przez [Hossak i inni, 1983, s. 208–221; Wieteska, 2004 s. 95–98; Pobłocka, 2012, s. 69–70; Wuthrich, Merz, 2008, s. 15–20].

Metoda CL bazuje na skumulowanych danych szkodowych $\lfloor C_{i,j} \rfloor = \{C_{i,j}: i = 0, 1,\ldots, n–1;\ j = 0, 1,\ldots, n-i-1\}$, z których wyznacza się współczynniki przejścia wg wzoru:

$$f_j = \sum_{i=1}^{n-j+1} C_{i,j+1} \div \sum_{i=1}^{n-j+1} C_{i,j}, \text{ dla } j = 1,2,...,n-2 \tag{7}$$

gdzie: $C_{i,j}$ to skumulowane płatności z *i*-go okresu wypadkowego wypłacone do *j*-go okresu rozliczeniowego; n to liczba okresów badanych.

Skumulowane przyszłe płatności oblicza się według wzoru:

$$\hat{C}_{i,j} = C_{i,j-k} \cdot \tilde{f}_k \text{ dla i} = 2,\ldots,n-1;\ j = n-i+1,\ n-i,\ldots,n-1 \tag{8}$$

gdzie: $C_{i,j-k}$ to skumulowane płatności z *i*-go okresu wypadkowego wypłacone do „*j–k*"-go okresu rozliczeniowego;

$$\tilde{f}_k = \prod_{j=1}^{k} f_j \tag{9}$$

2.3. Metoda grossing up

Metoda grossing up (GU) zakłada, że istnieje zamknięty okres wypadkowy (tzw. okres zerowy, ozn. „0"), po którym nie wystąpi już żadna płatność, a przyszłe szkody zależą od danych historycznych i są proporcjonalne do szkód z okresów opóźnienia w rozliczaniu szkód. Została ona opisana m.in. przez [Wolny, 2000, s. 17–19; Wolny, 2005, s. 16–51; Wieteska, 2004, s. 100–101; Pobłocka, 2012, s. 68].

Metoda GU bazuje na skumulowanych danych szkodowych $\lfloor C_{i,j} \rfloor = \{C_{i,j}: i = 0, 1,\ldots, n–1;\ j = 0, 1,\ldots, n-i-1\}$, z których wyznacza się współczynniki całkowitej szkody (opisujące zależność pomiędzy wypłaconymi odszkodowaniami i świadczeniami a łączną wartością wypłaconych już świadczeń) wg wzoru:

$$f_{0,j} = \frac{C_{0,j}}{C_0} \quad \text{dla} \quad j = 0,1,2,\ldots,n-1 \tag{10}$$

gdzie: $C_{0,j}$ to skumulowane płatności z tytułu szkód zgłoszonych w „0"-ym okresie wypadkowym i wypłaconych z opóźnieniem o *j*-okresów; C_0 to skumulowane płatności z tytułu szkód zgłoszonych w „0"-ym okresie wypadkowym; *n* to liczba okresów badanych.

Skumulowane przyszłe płatności oblicza się wg wzoru:

$$\hat{C}_i = \frac{C_{i,n-1-i}}{f_{0,n-1-i}} \tag{11}$$

gdzie: $C_{i,n-i-1}$ to skumulowane płatności wypłacone do dnia tworzenia rezerwy (dane z przekątnej w macierzy szkód); $f_{0,n-1-i}$ to współczynniki całkowitej szkody.

3. Rezerwy techniczno-ubezpieczeniowe na potrzeby wypłacalności

Ustawę o działalności ubezpieczeniowej z 2015 r. wprowadzono w celu przystosowania polskiego prawa do norm UE (Dyrektywy Wypłacalność II) oraz regulacji międzynarodowych w zakresie rachunkowości i nadzoru nad prowadzeniem działalności ubezpieczeniowej w UE. Zgodnie z nią zakład ubezpieczeń „tworzy rezerwy techniczno-ubezpieczeniowe dla

celów wypłacalności, przeznaczone na pokrycie bieżących i przyszłych zobowiązań, jakie mogą wyniknąć z zawartych umów ubezpieczenia, umów gwarancji ubezpieczeniowych lub umów reasekuracji" [ustawa, 2015, art. 224 § 1]. Wartość tych rezerw „jest równa bieżącej kwocie, którą zakład ubezpieczeń lub zakład reasekuracji zapłaciłby przy natychmiastowym przeniesieniu zobowiązań z tytułu zawartych umów ubezpieczenia lub umów reasekuracji na inny zakład ubezpieczeń lub reasekuracji" (art. 224 § 3 ww. ustawy). Wartość tę ustala się „w sposób ostrożny, wiarygodny i obiektywny", zgodnie z „danymi dostępnymi na rynkach finansowych oraz ogólnodostępnymi danymi na temat różnych rodzajów ryzyka aktuarialnego" (art. 224 § 4 ww. ustawy). Wartość liczbowa ww. rezerw według Dyrektywy Wypłacalność II i (art. 225 § 1 ww. ustawy) to wartość rynkowa będąca sumą najlepszego oszacowania oraz marginesu ryzyka; a według Polskich Standardów Rachunkowości to wartość księgowa otrzymana z metod statystycznych uwzględniających narzuty bezpieczeństwa. Najlepsze oszacowanie (*best estimate*) to „oczekiwana wartość obecna przyszłych przepływów pieniężnych ważonych prawdopodobieństwem ich realizacji, obliczonej przy zastosowaniu odpowiedniej struktury terminowej stopy procentowej wolnej od ryzyka"[7] (art. 225 § 2 ww. ustawy). Oblicza się ją, „na podstawie aktualnych i wiarygodnych informacji i realistycznych założeń", w wysokości brutto, bez pomniejszania o kwoty należne z umów reasekuracji i od spółek celowych. „Prognozy przepływów pieniężnych stosowane w obliczeniu najlepszego oszacowania uwzględniają wpływy i wypływy pieniężne wymagane do rozliczenia zobowiązań z tytułu zawartych umów ubezpieczenia lub umów reasekuracji w okresie trwania tych zobowiązań" (art. 225 § 2 ww. ustawy). Margines ryzyka (*margin risk*) to narzut bezpieczeństwa zapewniający, aby wartość rezerw TU do celów wypłacalności „odpowiadała kwocie, której zakład ubezpieczeń lub zakład reasekuracji zażądałyby za przejęcie zobowiązań z tytułu zawartych umów ubezpieczenia lub umów reasekuracji i wywiązanie się z zobowiązań" (art. 225 § 3 ww. ustawy). Przy kalkulacji rezerw należy uwzględnić m.in.: 1) koszty, jakie zostaną poniesione w związku z obsługą zobowiązań; 2) oczekiwany wzrost kosztów i roszczeń ubezpieczeniowych; 3) płatności na rzecz ubezpieczających, ubezpieczonych lub uprawnionych z umów ubezpieczenia

[7] Oczekiwana wartość obecna przyszłych przepływów pieniężnych była do tej pory rozumiana przez autora (statystyka) jako wartość oczekiwana z wartości aktuarialnej (bieżącej) z przyszłych przepływów pieniężnych. Z tego powodu autor uważał, że rezerwy TU tworzone na potrzeby wypłacalności będą mogły być szacowane tylko metodami stochastycznymi. Jednakże po implementacji Dyrektywy Wypłacalność II pewne zakłady ubezpieczeń nadal tworzą rezerwy TU do celów wypłacalności metodami deterministycznymi. Z tego powodu autor postanowił w niniejszym artykule zaimplementować proste metody deterministyczne do celów wypłacalności.

oraz cedentów z umów reasekuracji; 4. wartości gwarancji finansowych i opcji umownych (art. 231 par. 1 ww. ustawy).

Podstawy, metody i założenia wyceny rezerw TU na potrzeby wypłacalności według Dyrektywy Wypłacalność II nie różnią się istotnie od zasad wyceny stosowanych na potrzeby rachunkowości według PSR. Różnice w wycenie wynikają głównie z innych zasad wyceny poszczególnych rodzajów rezerw oraz innej klasyfikacji rezerw[8]. Istotne różnice przy szacowaniu rezerw TU [*Sprawozdanie*, 2017, s. 46] dla linii biznesu 1–12 na potrzeby wypłacalności wymienia następująco: rezerwy oblicza się w wartości zdyskontowanej; w rezerwach uwzględnia się przyszłe składki, koszty ogólne, margines ryzyka; w rezerwach nie uwzględnia się regresów i odzysków; rezerwa składek pomniejszana jest o przyszły zysk. A na potrzeby rachunkowości: rezerwy oblicza się w wartości niezdyskontowanej; w rezerwach uwzględnia się przyszłe regresy i odzyski, a nie uwzględnia się przyszłych składek, kosztów ogólnych i marginesu ryzyka; rezerwa składek zawiera przyszły zysk. Dla linii biznesu 34 na potrzeby wypłacalności rezerwy oblicza się dla zmiennej stopy dyskontowej i zmiennej stopy wzrostu, z uwzględnieniem marginesu ryzyka. A na potrzeby rachunkowości: rezerwy oblicza się dla stałej stopy dyskontowej i stałej stopy wzrostu, bez uwzględnienia marginesu ryzyka.

Przy obliczaniu rezerw do celów wypłacalności stosuje się odpowiednie struktury terminowe stopy procentowej wolnej od ryzyka, bez korekty dopasowującej, oraz korekty z tytułu zmienności (tab. 3).

Tablica 3. Struktura terminowej stopy procentowej wolnej od ryzyka dla waluty PLN, EUR, USD według terminu zapadalności w latach

Termin zapadalności (w latach)	Stopa procentowa wolna od ryzyka dla waluty		
	PLN	EUR	USD
1	1,41%	–0,16%	0,77%
2	1,55%	–0,13%	1,08%
3	1,76%	–0,04%	1,33%
4	1,97%	0,10%	1,50%
5	2,18%	0,23%	1,66%
6	2,40%	0,38%	1,77%
7	2,64%	0,53%	1,88%
8	2,78%	0,67%	1,98%
9	2,85%	0,80%	2,06%
10	2,99%	0,92%	2,15%

Źródło: Opracowanie własne na podstawie [rozporządzenie UE, 2016].

[8] Dyrektywa Wypłacalność II dzieli rezerwy według linii biznesu, a ustawa o rachunkowości i ustawa ubezpieczeniowa dzieli rezerwy według rodzajów rezerw.

Niestety ww. wytyczne w zakresie metod tworzenia rezerw techniczno-ubezpieczeniowych są bardzo ogólne, co daje możliwość uzyskania różnych wyników dla tych samych danych źródłowych (problem zostanie rozwinięty w kolejnym rozdziale). Odpowiedzialność za tworzenie rezerw techniczno-ubezpieczeniowych „spada" na zarządy zakładów ubezpieczeń i zakładów reasekuracji.

4. Wyniki badania

W badaniu na wstępie sprawdzono, czy istnieje różnica w szacowaniu wartości księgowej rezerwy IBNR w ubezpieczeniach majątkowych na podstawie danych szkodowych wyrażonych w wartościach nominalnych i realnych. W tym celu, na podstawie danych szkodowych z tablicy 2 (wielkości wypłaconych odszkodowań i świadczeń wyrażonych w wartościach nominalnych w tys. PLN), oszacowano rezerwy IBNR na potrzeby rachunkowości i sprawozdawczości statutowej według PSR za pomocą trzech badanych metod: średnich wartości, chain ladder i grossing up (kolumny 2–4 w tablicy 4). Następnie na podstawie odpowiednich współczynników inflacji CPI publikowanych przez GUS obliczono wartości realne danych szkodowych z tablicy 2 (w cenach z 2016 r.) i w kolejnym kroku badania oszacowano rezerwy IBNR w wartościach realnych (kolumny 5–7 w tablicy 4).

W dalszym ciągu badania, w poszukiwaniu prawidłowości w tworzeniu rezerw TU, obliczono wartość rezerwy IBNR na potrzeby wypłacalności według trzech badanych metod. W tym celu dla struktury stopy procentowej wolnej od ryzyka wyrażonej w walucie PLN (z tablicy 3 wg EIOPA) oszacowano wartość rynkową rezerwy IBNR (tzn. zdyskontowano przyszłe płatności dla stopy procentowej wolnej od ryzyka na moment tworzenia rezerwy[9]), osobno dla nominalnych i realnych wartości danych szkodowych (kolumny 8–13 w tablicy 4). Z badania wynika, że wartości rynkowe rezerwy IBNR, obliczone na podstawie wartości nominalnych wypłaconych odszkodowań i świadczeń brutto, są wyższe we wszystkich badanych metodach, niż odpowiadające im wartości rynkowe rezerwy IBNR obliczone na podstawie wartości realnych z 2016 r. (to potwierdza wyniki wcześniejszej analizy).

[9] Zdyskontowane płatności z k-go roku obliczono jako iloczyn płatności z k-tego roku i czynnika dyskontującego (dla stopy procentowej wolnej od ryzyka równej i) podniesionego do potęgi k-tej równej terminowi zapadalności. Czynnik dyskontujący to v=1/(1+i).

Tablica 4. Rezerwa IBNR w latach 2008–2016 oraz łącznie: wartość księgowa wg Polskich Standardów Rachunkowości i wartość rynkowa wg Dyrektywy Wypłacalność II (wartości nominalne i realne z 2016 r.) [w mln PLN]

Lata	Rezerwy IBNR – wartość księgowa wg Polskich Standardów Rachunkowości [w mln PLN]						Rezerwy IBNR – wartość rynkowa wg Dyrektywy Wypłacalność II [w mln PLN]					
	wartości nominalne w metodzie			wartości realne z 2016 r. w metodzie			wartości nominalne w metodzie			wartości realne z 2016 r. w metodzie		
	SW1	CL1	GU1	SW2	CL2	GU2	SW1	CL1	GU1	SW2	CL2	GU2
2008	17	34	34	16	33	33	17	34	34	16	32	32
2009	36	72	76	34	68	71	36	71	75	34	67	70
2010	66	131	137	61	120	123	65	129	135	60	119	121
2011	69	138	136	43	99	100	68	136	134	42	97	99
2012	87	170	165	62	126	122	86	167	162	61	124	120
2013	107	201	191	85	156	146	105	197	188	83	153	143
2014	162	284	264	138	232	211	159	279	259	136	227	207
2015	306	491	514	277	425	438	300	482	505	272	417	430
2016	1160	1558	1628	1117	1462	1495	1142	1533	1602	1100	1439	1472
Razem	2010	3079	3146	1834	2719	2739	1978	3028	3093	1805	2675	2694

SW – metoda średnich wartości, CL – metoda chain-ladder, GU – metoda grossing up. Indeks „1" – obliczenia na podstawie trójkąta szkód z tablicy 2 w wartościach nominalnych. Indeks „2" – obliczenia na podstawie trójkąta szkód z tablicy 2 w wartościach realnych z 2016 r.
Źródło: Opracowanie własne, na podstawie na podstawie danych z tablicy 2 oraz danych z GUS (wskaźników inflacji CPI).

Wartości rynkowe rezerwy IBNR (zaprezentowane w tablicy 4), obliczone na potrzeby wypłacalności, dla wszystkich badanych metod mają niższe wartości niż odpowiadające im wartości rezerw na potrzeby sprawozdawczości statutowej i to bez względu, czy dane szkodowe były wyrażone w wartościach nominalnych, czy realnych[10]. To skłania do zastanowienia, kiedy wartości rynkowe rezerw TU mogą być wyższe od wartości księgowych? Odpowiedź na to pytanie to sytuacje, kiedy stopy procentowe wolne od ryzyka byłyby ujemne. Taki przypadek dla waluty PLN nie wystąpi, gdyż prezentowane w tablicy 3 stopy procentowe wolne od ryzyka są zawsze dodatnie. Jednakże dla waluty EUR, przy terminie zapadalności od 1 do 3 lat, stopy procentowe wolne od ryzyka są ujemne i rozważany przypadek jest możliwy.

[10] Podobne wyniki [*Sprawozdanie*, 2017, s. 48] PZU opublikował dla danych z 2016 r., gdzie wartość rezerw techniczno-ubezpieczeniowych innych niż ubezpieczeniach na życie: wg sprawozdania finansowego była równa 13,4 mld zł, a według Dyrektywy Wypłacalność II 8,6 mld zł; a wartość rezerw TU w ubezpieczeniach na życie: wg sprawozdania finansowego była równa 5,5 mld zł, a według Dyrektywy Wypłacalność II: 4,8 mld zł.

Kolejnym rozważanym problemem jest inna struktura stóp procentowych wolnych od ryzyka dla różnych walut (PLN, EUR, USD) prezentowana w tablicy 3. Jaki to może mieć wpływ na wartość tworzonych rezerw, tzn., jak kurs walut i odpowiadająca mu struktura procentowa stóp wolnych od ryzyka wpływa na poziom tworzonych rezerw? Poszukując odpowiedzi na to pytanie, w kolejnym kroku badania próbowano zweryfikować, czy publikowane struktury stóp procentowych wolnych od ryzyka są odpowiednio bezpieczne w poszczególnych walutach, tzn. jak poziom tworzonych rezerw TU zależy od rodzaju waluty i odpowiadającej jej struktury stóp procentowych wolnych od ryzyka?

W tym celu, na podstawie kursów walut publikowanych przez Narodowy Bank Polski w ostatnich dniach roku w badanych latach, dane szkodowe z tablicy 2 wyrażono w cenach euro oraz dolara amerykańskiego (tab. 5 i 6). Następnie, na podstawie wyznaczonych trójkątów rozliczenia szkód wyrażonych w euro i dolarze amerykańskim, obliczono wartości rynkowe przyszłych zobowiązań ubezpieczeniowych oraz rezerwy IBNR dla potrzeb wypłacalności, wg odpowiedniej struktury stopy procentowej wolnej od ryzyka danej waluty. Na koniec, dla porównania wyników z wcześniejszych analiz, uzyskane wyniki wyrażono w jednej walucie, tj. w PLN (tab. 7). Na podstawie wyznaczonych trójkątów rozliczenia szkód w euro i dolarze amerykańskim obliczono wartości rynkowe przyszłych zobowiązań ubezpieczeniowych oraz rezerwy IBNR dla potrzeb wypłacalności, wg odpowiedniej struktury stopy procentowej wolnej od ryzyka danej waluty. Na koniec, dla porównania wyników z wcześniejszych analiz, uzyskane wyniki wyrażono w jednej walucie – złoty (tab. 7).

Tablica 5. Wypłacone odszkodowania i świadczenia z tytułu pozostałych ubezpieczeń osobowych i majątkowych w latach 2007–2016 [w mln EUR]

Okres wypadkowy (i)	Okres opóźnienia w wypłacie odszkodowania (j)										Kurs 1 EUR	
	0	1	2	3	4	5	6	7	8	9	z dnia	wartość w PLN
2007	796	195	44	19	11	9	7	8	9	7	31.12.2007	3,5820
2008	710	224	37	17	11	10	8	10	8		31.12.2008	4,1724
2009	776	196	24	17	14	10	9	7			31.12.2009	4,1082
2010	923	245	41	30	14	12	13				31.12.2010	3,9603
2011	669	166	33	20	14	12					30.12.2011	4,4168
2012	685	161	29	20	15						31.12.2012	4,0882
2013	606	147	31	20							31.12.2013	4,1472
2014	608	158	34								31.12.2014	4,2623
2015	685	230									31.12.2015	4,2615
2016	812										30.12.2016	4,4240

Źródło: Opracowanie własne, na podstawie kursu euro publikowanego przez NBP oraz danych szkodowych PZU S.A. z 2017 r. [*Sprawozdanie*, 2017].

Tablica 6. Wypłacone odszkodowania i świadczenia z tytułu pozostałych ubezpieczeń osobowych i majątkowych w latach 2007–2016 [w mln USD]

Okres wypadkowy (i)	Okres opóźnienia w wypłacie odszkodowania (j)										Kurs 1 USD	
	0	1	2	3	4	5	6	7	8	9	z dnia	wartość w PLN
2007	1170	274	64	25	14	12	10	10	10	8	31.12.2007	2,4350
2008	1000	323	49	22	14	13	10	11	8		31.12.2008	2,9618
2009	1118	262	31	23	19	12	10	7			31.12.2009	2,8503
2010	1234	317	54	41	17	13	14				31.12.2010	2,9641
2011	865	219	46	25	15	13					30.12.2011	3,4174
2012	903	222	35	22	16						31.12.2012	3,0996
2013	835	179	33	21							31.12.2013	3,0120
2014	739	173	36								31.12.2014	3,5072
2015	748	244									31.12.2015	3,9011
2016	860										30.12.2016	4,1793

Źródło: Opracowanie własne, na podstawie kursu USD publikowanego przez NBP oraz danych szkodowych PZU S.A. z 2017 r. [*Sprawozdanie*…, 2017].

Tablica 7. Rezerwy IBNR oszacowane metodą chain lader, grossing up i średnich wartości w wartościach księgowych (w cenach nominalnych i realnych) oraz w wartościach rynkowych (dla waluty PLN, EUR, USD i odpowiadającej im struktury stóp procentowych wolnych od ryzyka)

Lata	Wartość rezerwy IBNR [w mln PLN] w metodzie														
	chain ladder					grossing up					średnich wartości				
	rynkowa			księgowa		rynkowa			księgowa		rynkowa			księgowa	
	PLN	EUR	USD	N	R	PLN	EUR	USD	N	R	PLN	EUR	USD	N	R
2008	34	31	29	34	33	34	31	29	34	33	17	15	15	17	16
2009	71	68	67	72	68	75	71	69	76	71	36	35	34	36	34
2010	129	128	120	131	120	135	130	124	137	123	65	65	62	66	60
2011	136	127	119	138	99	134	121	119	136	100	68	65	63	69	42
2012	167	168	165	170	126	162	156	160	165	122	86	88	89	87	61
2013	197	196	199	201	156	188	176	184	191	146	105	107	113	107	83
2014	279	275	256	284	232	259	242	231	264	211	159	160	154	162	136
2015	482	478	423	491	425	505	470	433	514	438	300	303	276	306	272
2016	1533	1497	1376	1558	1462	1602	1400	1299	1628	1495	1142	1129	1059	1160	1100
Suma	3028	2968	2753	3079	2719	3093	2797	2648	3146	2739	1978	1967	1864	2010	1805

N – wartość nominalna w cenach bieżących, R – wartość realna w cenach z 2016 r.
Źródło: Opracowanie własne, na podstawie danych z tablicy 6 i kursu USD wg NBP.

Z tablicy 7 wynika, że dla danych szkodowych (z tablicy 2), wyrażonych w różnej walucie (w PLN z tablicy 2, EUR z tablicy 5 i USD z tablicy 6), otrzymano różne wartości rynkowe rezerwy IBNR szacowanej dla potrzeb wypłacalności. Najwyższe wartości rynkowe rezerwy IBNR uzyskano dla danych szkodowych wyrażonych w walucie PLN, a najniższe wartości rezerwy uzyskano dla danych szkodowych wyrażonych w walucie USD. Najwyższe wartości rynkowe rezerwy IBNR oszacowanej na potrzeby wypłacalności dla danych szkodowych wyrażonych w walucie PLN uzyskano metodą grossing up, a najniższe metodą średnich wartości. Dla danych szkodowych wyrażonych w walucie EUR i USD najwyższe wartości rynkowe rezerwy uzyskano metodą chain lader, a najniższe metodą średnich wartości. Najwyższe wartości księgowe rezerwy IBNR oszacowanej na potrzeby rachunkowości otrzymano metodą grossing up, a najniższe metodą średnich wartości. Reasumując, można stwierdzić, że w zależności od rodzaju waluty użytej do prezentowania danych szkodowych uzyskuje się różne wartości rynkowe rezerwy IBNR.

Zakończenie

Badanie wykazało, ze istnieje dualizm w wycenie aktywów i zobowiązań zakładów ubezpieczeń i zakładów reasekuracji, w szczególności rozbieżność z zakresie tworzenia rezerw techniczno-ubezpieczeniowych do celów statutowych i na potrzeby rachunkowości. Badanie potwierdziło wcześniejsze analizy, że wartość rezerwy IBNR zależy od używanej do kalkulacji metody [Pobłocka, 2010; Wolny-Dominiak, 2013]. Badanie pokazało, że wartość rezerwy IBNR tworzona na potrzeby sprawozdawczości wg wartości księgowej różni się od wartości rezerwy tworzonej na potrzeby wypłacalności wg wartości rynkowej. Różne wartości rezerw mogą wpływać m.in. na ustalenie odrębnej masy upadłościowej zakładów ubezpieczeń i zakładów reasekuracji na pokrycie zobowiązań w przypadku upadłości. W badaniu wartości rezerwy IBNR tworzonej na potrzeby wypłacalności są niższe niż wartości na potrzeby rachunkowości (dla dodatniej struktury stopy procentowej wolnej od ryzyka). Jeżeli wartość rynkowa rezerw TU byłaby używana do celów rachunkowości, to zakłady ubezpieczeń wykazałyby mniejszą zmianę stanu rezerw i uzyskałyby większą podstawę opodatkowania, a w konsekwencji zapłaciłyby większy podatek dochodowy. Reasumując, wartość rezerwy zależy m.in.: od rodzaju danych z tytułu wypłaconych odszkodowań i świadczeń w trójkątach rozliczania szkód (wartości nominalne szkód generują, dla dodatnich stóp procentowych wolnych od ryzyka, wyższe wartości rezerw, niż wartości realne); od rodzaju waluty danych szkodowych; od rodzaju waluty struktury stopy procentowej wolnej od ryzyka. W efekcie badanie wykazało, że brak

wytycznych w zakresie tworzenia rezerwy IBNR powoduje rozbieżności w oszacowanych wartościach bieżących i przyszłych zobowiązań z tytułu zawartych umów ubezpieczenia i umów gwarancji ubezpieczeniowych. To stwarza zagrożenie, że tworzone przez zakłady ubezpieczeń i zakłady reasekuracji rezerwy techniczno-ubezpieczeniowe mogą okazać się niewystarczające w celu zapewnienia ochrony ubezpieczeniowej.

Literatura

Biuletyny Roczne Komisji Nadzoru Finansowego, Rynek ubezpieczeń. Część III Zagregowane dane finansowe (2016), Komisja Nadzoru Finansowego, Warszawa.

Dyrektywa 2009/138/WE Parlamentu Europejskiego i Rady z dnia 25 listopada w sprawie podejmowania i wykonywania działalności ubezpieczeniowej i reasekuracyjnej – Wypłacalność II, Dz.U. UE L 335 z 17.12.2009 r. , z późn. zm.

Hossak I.B., Pollard J. H., Zehnwirth B. (1983), *Introductory statistics with applications in general insurance*, Great Britain at the University Press, Cambridge.

Kędziora K. (2016), *Wpływ ustawy o działalności ubezpieczeniowej i reasekuracyjnej na rachunkowość i sprawozdawczość zakładów ubezpieczeń – wybrane zagadnienia*, „Prawo Asekuracyjne", nr 1.

Komunikat z 367 posiedzenia Komisji Nadzoru Finansowego w dniu 29.08.2017 roku, Komisja Nadzoru Finansowego, https://www.knf.gov.pl/o_nas/komunikaty?articleId=58474&p_id=18, dostęp: 1.09.2017.

Pobłocka A. (2012), *Aktuarialne modele szacowania rezerwy na zaistniałe niezgłoszone szkody w ubezpieczeniach majątkowych*, rozprawa doktorska obroniona dnia 26.09.2012 r. na Wydziale Zarządzania Uniwersytetu Gdańskiego, Sopot.

Pobłocka A. (2010), *Klasyczne metody szacowania rezerwy IBNR w ubezpieczeniach majątkowych – jaką technikę wybrać w praktyce?*, „Zeszyty Naukowe Uniwersytetu Ekonomicznego w Poznaniu", nr 182.

Rozporządzenie Ministra Finansów z dnia 12 kwietnia 2016 r. w sprawie szczególnych zasad rachunkowości zakładów ubezpieczeń i zakładów reasekuracji, Dz.U. poz. 562.

Rozporządzenie delegowane Komisji (UE) 2015/35UE z dnia 10 października 2014 r. uzupełniające ww. dyrektywę Parlamentu Europejskiego i Rady 2009/138/WE w sprawie podejmowania i prowadzenia działalności ubezpieczeniowej i reasekuracyjnej (Wypłacalność II), Dz.U. UE L 12 z 17.01.2015 r.

Spigarska E. (2014), *Polityka rachunkowości zakładów ubezpieczeń wobec dyrektywy Wypłacalność II*, w: J. Bieliński, E. Wycinka (red.), *Problemy współczesnego zarządzania*, Wydział Zarządzania Uniwersytetu Gdańskiego.

Sprawozdanie o wypłacalności i kondycji finansowej na dzień 31 grudnia 2016 roku i za rok obrotowy zakończony tego dnia, Powszechny Zakład Ubezpieczeń S.A., https://www.pzu.pl/c/document_library/get_file?uuid=0c0f3e8c-07bf-45bc-a68d-83c3a81b5825&groupId=10172, dostęp: 30.09.2017.

Ustawa z dnia 29 września 1994 r. o rachunkowości, t.j. Dz.U. z 2018 r. poz. 395.

Ustawa z dnia 22 maja 2003 r. o działalności ubezpieczeniowej, Dz.U. Nr 124, poz. 1151 ze zm.

Ustawa z dnia 11 września 2015 r. o działalności ubezpieczeniowej i reasekuracyjnej, Dz.U. poz. 1844.
Wieteska S. (2004), *Rezerwy techniczno-ubezpieczeniowe zakładów ubezpieczeń majątkowo-osobowych*, Branta, Bydgoszcz– Łódź.
Wolny-Dominiak A. (2013), *Możliwości programu R w szacowaniu rezerwy IBNR*, „Wiadomości ubezpieczeniowe", nr 2.
Wolny A., *Kalkulacja rezerwy szkodowej. Metoda grossing up*, w: W. Szkutnik (red.), *Seria: Statystyka ubezpieczeniowa*, Wydawnictwo Akademii Ekonomicznej im. Karola Adamieckiego w Katowicach, Katowice.
Wuthrich M.V., Merz M. (2008), *Stochastic claims reserving methods in insurance*, John Wiley & Sons, Hoboken.

Streszczenie

Zakłady ubezpieczeń i zakłady reasekuracji są zobligowane ustawowo do tworzenia rezerw techniczno-ubezpieczeniowych. Rezerwy te do końca 2015 r. tworzono tylko dla potrzeb sprawozdawczości statutowej (wg wartości księgowej), a od stycznia 2016 r. tworzy się dodatkowo dla potrzeb wypłacalności (wg wartości rynkowej). W artykule przedstawiono zasady wyceny rezerw techniczno-ubezpieczeniowych według Dyrektywy Wypłacalność II oraz wskazano możliwości i zagrożenia dla zakładów ubezpieczeń wynikające z braku precyzyjnego określenia zasad wyceny rezerw na potrzeby wypłacalności według Dyrektywy Rady Unii Europejskiej Wypłacalność II, która weszła w życie dnia 1.01.2016 r. W tym celu zaimplementowano trzy deterministyczne metody szacowania rezerwy z tytułu szkód zaistniałych i niezgłoszonych w zakładzie ubezpieczeń do dnia tworzenia rezerw (tzw. rezerwy IBNR) w ubezpieczeniach majątkowych do rzeczywistych danych szkodowych publikowanych przez Powszechny Zakład Ubezpieczeń S.A. W badaniu użyto rezerwy IBNR, która jest jedną z najważniejszych rezerw szkodowych w ubezpieczeniach majątkowych. Jej wartość istotnie wpływa na poziom tworzonych rezerw techniczno-ubezpieczeniowych, a jej niedoszacowanie może spowodować niewypłacalność zakładów ubezpieczeń i zakładów reasekuracji.

Słowa kluczowe

rezerwy techniczno-ubezpieczeniowe – wartość księgowa i rynkowa, rezerwa IBNR, metoda średnich wartości, metoda Bornhuettera-Fergusona, metoda chain-ladder, metoda grossing-up, dyrektywa Wypłacalność II

Creating IBNR provision using deterministic methods for solvency in non-life insurance companies (Summary)

Insurance and reinsurance undertakings are legally obliged to establish technical provisions. These reserves were created for the purposes of statutory reporting only (at book value) until the end of 2015, and since January 2016 they are created additionally for solvency purposes (at market value). The article attempts to identify opportunities and threats for insurance and reinsurance undertakings resulting from the lack of precise determination of the principles of valuation

of technical provisions for solvency purposes under the Directive of the Council of the European Union Solvency II, which entered into force on January 1st 2016. For this purpose, three deterministic methods of estimating the provision for damage occurred and not reported in the insurance company until the date of creation of technical provisions (so-called reserve IBNR) in property insurance to actual damage data published by public insurance company (PZU SA.). Reserve IBNR is one of the most important losses in property insurance. Its value has a significant impact on the level of technical provisions created and its underestimation may result in the insolvency of insurance and reinsurance undertakings.

Keywords
technical provisions – book value and market value of reserves, IBNR provision; methods of the average value of development indicators losses, Bornhuetter-Ferguson method, the chain ladder method, grossing up method

Ewa Wycinka*

Tablice wieloprzyczynowe w ocenie ryzyka niewypłacalności. Podejście kohortowe

Wstęp

Ryzyko kredytowe to możliwość niewywiązania się dłużnika z zobowiązań wynikających z zawartej z bankiem umowy kredytowej w terminie określonym warunkami kredytu [Jajuga, 2004, s. 23]. Zaprzestanie spłat kredytu określane jest jako niewypłacalność kredytobiorcy. Teoria w zakresie modeli ryzyka kredytowego, w tym prawdopodobieństwa niewypłacalności, jest bardzo dobrze rozwinięta [Jajuga, 2007, s. 148–161]. Modele te charakteryzują się bardzo dużą różnorodnością. Rozwój wiedzy z zakresu zarządzania ryzykiem był przesłanką do powstania i wdrożenia w życie Nowej Umowy Kapitałowej (*Basel II*), która umożliwiła bankom rozwój wewnętrznych metod oceny ryzyka. W ramach tzw. ratingów wewnętrznych bank wyznacza między innymi prawdopodobieństwo utraty zdolności kredytowej (niewypłacalności) (*probability of default*, PD) [tamże, s. 207–213]. Zgodnie z regulacjami wprowadzonymi przez Nową Umowę Kapitałową oraz Dyrektywę Europejską (2006/48/WE) niewypłacalność definiowana jest w horyzoncie jednego roku, natomiast nowe międzynarodowe standardy rachunkowości IFRS 9, które weszły w życie 1 stycznia 2018 r. wprowadzają wymóg szacowania niewypłacalności dla całego okresu trwania umowy w celu wyliczenia oczekiwanej straty kredytowej (*lifetime expected credit losses)*. Narzędziami najbardziej odpowiednimi do badania prawdopodobieństwa niewypłacalności są w tej sytuacji metody analizy przeżycia (zarówno parametryczne, jak i nieparametryczne) [Dirick i inni, 2017, s. 652].

W niniejszym artykule uwagę poświęcono możliwości zastosowania metod nieparametrycznych do oceny ryzyka niewypłacalności. Celem badania empirycznego jest określenie rozkładu ryzyka niewypłacalności dla kohorty kredytów detalicznych za pomocą tablic wieloprzyczynowych. Kohortę kredytów stanowi jednorodna grupa kredytów udzielonych przez bank w określonym przedziale czasu i obserwowana przez okres równy harmonogramowi spłat kredytów. Według wiedzy autorki niniejsza praca jest pierwszą, w której w sposób kompleksowy przedstawiono propozycję

* Dr, Katedra Statystyki, Wydział Zarządzania, Uniwersytet Gdański, ul. Armii Krajowej 101, 81-824 Sopot, ewa.wycinka@ug.edu.pl

zastosowania tablic wieloprzyczynowych do oceny ryzyka niewypłacalności. Uzyskane wyniki zostaną porównane z oszacowaniami metodą wskaźników umieralności[1] zaproponowaną przez Altmana [1989].

1. Metoda wskaźników umieralności Altmana

W terminach płatności kolejnych rat kredytu kohorta kredytów zmniejszać się może nie tylko z powodu wystąpienia we wcześniejszych okresach niewypłacalności, ale również wcześniejszego zakończenia kredytu z innych przyczyn, np. wcześniejszej spłaty kredytu, zmiany warunków umowy (np. przewalutowanie kredytu) itp. Zdarzenia te wpływają na liczbę kredytów analizowanych w kolejnych okresach płatności (wielkość kohorty), a przez to na oszacowane prawdopodobieństwa niewypłacalności w badanej kohorcie kredytów.

Modelem, który uwzględnia w badaniu ryzyka niewypłacalności zmniejszanie się portfela kredytów z różnych przyczyn, jest metoda wskaźników umieralności (*mortality rates*) zaproponowana przez Altmana [1989] do badania natężenia niewypłacalności, w kolejnych latach od udzielenia pożyczki lub emisji obligacji, dla kohort pożyczek i obligacji emitowanych przez przedsiębiorstwa.

Krańcowe wskaźniki umieralności (*marginal mortality rate*) dla *i*-tej kohorty w okresie *t* według czasu własnego kohorty obliczane są jako:

$$MMR_{t,i}=\frac{D_{t,i}^{(1)}}{N_{t,i}} \tag{1}$$

gdzie $D_{t,i}^{(1)}$ jest wartością obligacji/pożyczek w kohorcie *i*, dla których odnotowano niewypłacalność w roku *t* od udzielenia, natomiast $N_{t,i}$ jest wartością obligacji/pożyczek w portfelu kohorty *i* na początku roku *t*. Wartość $N_{t,i}$ obliczana jest jako różnica między początkową wartością portfela (dla $t = 0$) i wartością wszystkich pożyczek/obligacji, które z różnych przyczyn przestały być obsługiwane do roku $t - 1$ włącznie [Saunders, Allen, 2002, s. 121–123]. Aby wyeliminować wahania wskaźnika wynikające z niskiej wartości pożyczek/obligacji udzielonych w danym roku kalendarzowym, Altman zaproponował uśrednianie tak wyznaczonych wskaźników za pomocą średnich ważonych z *k* kohort:

$$\overline{MMR}_t = \Sigma_{i=1}^{k} MMR_{t,i} \cdot w_i \tag{2}$$

[1] Właściwszą nazwą byłyby „wskaźniki odejść", jednak zarówno w źródłowej pracy Altmana [1989], jak i później w pracach, które do niego nawiązywały [Altman i Hotchkiss 2006; Saunders i Allen 2002 oraz 2010], używane były nazwy *mortality rates* oraz *cumulative mortality rates*. W literaturze polskiej, zarówno w tłumaczeniu pracy Saundersa [2001], jak i w pracy Jajugi [2004], również użyto nazw „współczynnik umieralności", dlatego w niniejszym artykule zachowano oryginalne nazewnictwo.

gdzie wagi w_i są wyznaczane na podstawie wielkości portfela udzielonych pożyczek w kolejnych latach

$$w_i = \frac{N_{oi}}{\sum_{i=1}^{k} N_{oi}} \tag{3}$$

Wskaźnik przeżywalności jest dopełnieniem do jedności wskaźnika umieralności

$$\overline{SR}_t = 1 - \overline{MMR}_t \tag{4}$$

Ostatecznie skumulowany wskaźnik umieralności (*cumulative mortality rate*), wyrażający prawdopodobieństwo, że niewypłacalność wystąpi w okresie t pierwszych lat od udzielenia pożyczki, wyznaczany jest jako

$$CMR_t = 1 - \Pi_{j=1}^{k} \overline{SR}_j \tag{5}$$

Powyższe estymatory zostały skonstruowane dla kohort tworzonych z pożyczek udzielanych w ciągu jednego roku, mogą być jednak zastosowane do kohort składających się z jednostek, które dołączały do kohorty w dłuższym lub krótszym przedziale czasu. Należy wziąć pod uwagę jedynie to, aby liczebność kohorty była wystarczająca do celów analizy, ale jednocześnie warunki wewnętrzne i zewnętrzne mające wpływ na analizowane ryzyko nie uległy zmianie.

Metoda zaproponowana przez Altmana jest wykorzystywana do wyznaczania oczekiwanej stopy niewypłacalności portfela w całym okresie jego życia (*lifetime expected credit losses*). Altman i Hotchkiss [2006] za pomocą tej metody oszacowali krańcowe i skumulowane wskaźniki umieralności dla obligacji przedsiębiorstw do 10 lat od emisji na podstawie danych za lata 1971–2004. Saunders i Allen mimo prezentowania szeregu metod wyceny ryzyka niewypłacalności, w tym wielu nowszych niż propozycja Altmana, podkreślają, że metoda wskaźników umieralności udowodniła swoją przydatność i jest wciąż aktualna [Saunders, Allen, 2010, s. 117].

2. Istota tablic trwania (wymieralności)

Tablice trwania są formą prezentacji statystycznego modelu umieralności badanej populacji [Jackowska, 2013, s. 46]. Jest to model nieparametryczny, w którym za pomocą zestawu estymatorów szacuje się podstawowe funkcje czasu trwania w ustalonych jego przedziałach.

Tablice kohortowe budowane są na podstawie obserwacji wymierania jednostek należących do obserwowanej rzeczywistej kohorty (populacji). Tablice standardowe, inaczej jednoprzyczynowe (*standard life tables, single decrement life tables*), uwzględniają tylko jedną przyczynę opuszczenia populacji, natomiast tablice wieloprzyczynowe (*multiple decrement life tables*) biorą pod uwagę więcej niż jedną przyczynę [tamże, s. 49].

Tablice trwania wykorzystywane są przede wszystkim w demografii, jednak tak jak pozostałe metody analizy przeżycia znalazły zastosowanie w innych dziedzinach nauki, takich jak: nauki techniczne (analiza niezawodności), nauki społeczne czy ekonomiczne [szerzej: Balicki, 2006, s. 20–25]. W tej ostatniej dziedzinie tablice trwania wykorzystywane są między innymi do badania czasu istnienia przedsiębiorstw [m.in. Ptak-Chmielewska, 2010; Markowicz, 2012; Jackowska, 2015; Gurgul, Zając, 2016]. W przytoczonych publikacjach wykorzystywano tablice standardowe ze wskaźnikami umieralności liczonymi liczbą likwidowanych firm.

Tablice wieloprzyczynowe są natomiast szeroko wykorzystywane w demografii (tablice według przyczyn zgonów) oraz w metodach aktuarialnych (tablice wielostanowe). Flores i inni [2010] wykorzystali tablice wieloprzyczynowe do wyznaczenia składek i rezerw w ubezpieczeniach kredytów hipotecznych, uwzględniając ryzyko niewypłacalności i wcześniejszej spłaty.

3. Budowa wieloprzyczynowych tablic trwania

W celu zachowania spójności z oznaczeniami stosowanymi w metodzie Altmana w niniejszym artykule, do budowy tablic trwania przyjęto następujące oznaczenia: wielkość N_t dla kolejnych momentów $t = 0,\ldots,\omega$ oznacza liczbę jednostek w kohorcie, gdzie N_0 jest początkową wielkością kohorty, D_t jest liczbą zdarzeń (bez względu na rodzaj zdarzenia), a $D_t^{(i)}$ jest natomiast liczbą zdarzeń typu i w okresie $[t, t+1)$. Zachodzi przy tym równość $\Sigma_{i=1}^{p} D_t^{(i)} = D_t$.

W części empirycznej pracy rozważane są trzy przyczyny zdarzenia: niewypłacalność, wcześniejsza spłata oraz terminowe zakończenie umowy. Z tego powodu wystarczy rozważyć estymatory tablicowe dla danych pełnych (bez uwzględniania cenzurowania).

3.1. Tablica standardowa (jednoprzyczynowa)

Pierwszym krokiem budowy tablicy według przyczyn jest budowa tablicy standardowej dla wszystkich zdarzeń analizowanych łącznie [Siegel, Swanson, 2004, s. 332]. Podstawą tablicy jest arbitralnie ustalana wielkość l_O, najczęściej $l_O = 100\ 000$ wyznaczająca umowną początkową liczebność populacji (w momencie $t = 0$). Wielkości l_t dla kolejnych okresów są liczbą jednostek w populacji, które dotrwały do momentu t i mogą być wyznaczone z zależności $l_t = l_{t-1}(1 - q_{t-1})$.

Prawdopodobieństwo łącznej realizacji zdarzenia z jakiejkolwiek przyczyny w okresie $[t, t+1)$ dla jednostki, która dotrwała do momentu t, wynosi

$$q_t = \frac{d_t}{l_t} \tag{6}$$

Liczba zdarzeń w okresie $[t, t+1)$ jest oznaczana jako $d_t = l_t - l_{t+1}$.

Prawdopodobieństwo łącznej realizacji zdarzeń może być oszacowane jako

$$\hat{q}_t = \frac{D_t}{N_t} \tag{7}$$

Następnie, na podstawie przedstawionych zależności między funkcjami, mogą być szacowane pozostałe funkcje tablicowe. Funkcja L_t nazywana funduszem czasu trwania w przedziale $[t, t+1)$, przedstawia łączną liczbę okresów w tym przedziale, które przetrwały te z jednostek, które dotrwały w kohorcie do momentu t. Przy założeniu, że w jednostkowym przedziale czasu $[t, t+1)$ zdarzenia są równomiernie rozłożone, funkcja ta jest wyznaczana jako

$$L_t = l_{t+1} + \frac{1}{2} d_t \tag{8}$$

Łączny fundusz dalszego trwania T_t jest zsumowaną liczbą lat, które jeszcze przetrwają w kohorcie wszystkie te jednostki, które dotrwały do momentu t

$$T_t = \Sigma_{y=t}^{\omega} L_y, \tag{9}$$

gdzie ω jest maksymalnym czasem trwania.

Przeciętne dalsze trwanie dla jednostek, które dotrwały do momentu t, wyznacza się jako

$$e_t = \frac{T_t}{l_t} \tag{10}$$

3.2. Tablice wieloprzyczynowe

W tablicach wieloprzyczynowych liczba zdarzeń z przyczyny i w przedziale czasu $[t, t+1)$ szacowana jest jako

$$d_t^{(i)} = d_t \frac{D_t^{(i)}}{D_t} \tag{11}$$

Podstawy tablic wieloprzyczynowych dla każdej przyczyny i osobno wyznacza się jako

$$l_0^{(i)} = \sum_{t=0}^{\omega-1} d_t^{(i)} \tag{12}$$

zachodzi przy tym równość

$$\sum_{i=1}^{p} l_0^{(i)} = l_0 \tag{13}$$

Wartości l_t^i dla kolejnych przedziałów wyznacza się z zależności

$$l_{t+1}^{(i)} = l_t^{(i)} - d_t^{(i)} \tag{14}$$

Prawdopodobieństwo brutto tego, że jednostka, która dotrwała do czasu t, dozna zdarzenia i w okresie $[t, t+1)$, w sytuacji gdy istnieją wszystkie pozostałe rodzaje ryzyka, jest równe

$$q_t^{(i)} = \frac{d_t^{(i)}}{l_t} \tag{15}$$

Krańcowe prawdopodobieństwo zdarzenia z przyczyny i (*eventual probability*)[2] wyznacza prawdopodobieństwo, że jednostka, która nie doznała żadnego zdarzenia do czasu t, kiedykolwiek dozna zdarzenia i, oblicza się jako:

$$\psi_t^{(i)} = \frac{\sum_{k=t}^{\omega-1} d_k^{(i)}}{l_t} \tag{16}$$

3.3. Tablice z eliminacją wybranej przyczyny

Tablice z eliminacją wybranej przyczyny są konstruowane przy założeniu, że określona przyczyna zdarzeń zostaje wyeliminowana, co oznacza, że jednostki nie mogą opuścić populacji z tej przyczyny, ale są narażone na pozostałe rodzaje przyczyn odejść. Tablice z eliminacją wybranej przyczyny są więc tablicami dla hipotetycznej kohorty, dla której przyjmuje się założenie, że wyeliminowanie jednej przyczyny zdarzeń nie wpływa na natężenie zdarzeń z powodu pozostałych przyczyn.

Podstawowymi wielkościami szacowanymi w tablicach z eliminacją wybranej przyczyny są prawdopodobieństwa netto.

Prawdopodobieństwo netto w przedziale czasu $(t, t+1)$ zdarzenia z jakiejkolwiek innej przyczyny, gdy przyczyna i jest wyeliminowana, można oszacować ze wzoru

$$q_t^{(\cdot i)} = 1-(1-q_t)^{1-\frac{q_t^i}{q_t}} \tag{17}$$

Natomiast prawdopodobieństwo zdarzenia z powodu przyczyny i, w przedziale czasu $(t, t+1)$, gdy i jest jedyną przyczyną zdarzeń wynosi

$$q_t^{(i\cdot)} = 1-(1-q_t)^{\frac{q_t^i}{q_t}} \tag{18}$$

W tablicach wieloprzyczynowych z więcej niż dwiema przyczynami wyznaczyć można również cząstkowe prawdopodobieństwo brutto

[2] W pracy [Balicki, 2006] użyta została nazwa „prawdopodobieństwo ewentualnego zgonu", takie tłumaczenie zwrotu *eventual probability* sugeruje jednak możliwość uniknięcia zdarzenia i, co nie jest przedmiotem badania powyższego estymatora. Dlatego w niniejszym opracowaniu zaproponowano określenie „krańcowe prawdopodobieństwo zdarzenia z przyczyny i".

zdarzenia z przyczyny *i* przy założeniu, że przyczyna *m* została wyeliminowana [Bolesławski, 1973, s. 51–57]:

$$q_t^{(i)(.m)} = \frac{q_t^{(i)}}{q_t - q_t^{(m)}} \left(1 - (1 - q_t)^{1 - \frac{q_t^i}{q_t}} \right) \tag{19}$$

Podstawę tablicy z eliminacją wybranej przyczyny zgonów ustala się taką samą jak dla tablicy standardowej, tj. $l_0^i = 100\ 000$. Liczba jednostek w kohorcie w okresie *t*, przy założeniu, że przyczyna *i* została wyeliminowana, jest równa

$$l_t^{(.i)} = (1 - q_{t-1}^{(.i)}) \cdot l_{t-1}^{(.i)} \tag{20}$$

Fundusz czasu trwania w przedziale $[t, t+1)$ dla jednostek, które dotrwały w kohorcie do momentu *t*, przy założeniu, że przyczyna *i* została wyeliminowana, wyznacza się jako

$$L_t^{(.i)} = (1 - \frac{l_t - L_t}{l_t - l_{t+1}}) l_t^{(.i)} + \frac{l_t - L_t}{l_t - l_{t+1}} l_{t+1}^{(.i)} \tag{21}$$

Łączny fundusz dalszego trwania $T_t^{(.i)}$, przy założeniu, że przyczyna *i* została wyeliminowana, jest równy

$$T_t^{(.i)} = \sum_{y=t}^{\omega} L_y^{(.i)} \tag{22}$$

Przeciętne dalsze trwanie dla jednostek, które dotrwały do momentu *t*, przy założeniu, że przyczyna *i* została wyeliminowana, wynosi

$$e_t^{(.i)} = \frac{T_t^{(.i)}}{l_t^{(.i)}} \tag{23}$$

Na podstawie tablic z eliminacją wybranej przyczyny można ponadto wyznaczyć zysk w przeciętnym dalszym trwaniu dla kohorty z tytułu wyeliminowania przyczyny *i*

$$g_t^{(.i)} = e_t^{(.i)} - e_t \tag{24}$$

oraz zysk w przeciętnym trwaniu dla jednostek, które ostatecznie odeszłyby z kohorty z przyczyny *i*, gdyby ta nie została wyeliminowana

$$\gamma_t^{(.i)} = \frac{e_t^{(.i)} - e_t}{\psi_t^i} \tag{25}$$

[Anderson, 1999, s. 2–7; Siegel, Swanson, 2004, s. 324–331; Balicki, 2006, s. 259–276].

4. Wyniki badania empirycznego i dyskusja

Badanie empiryczne przeprowadzono na próbie 8070 24-miesięcznych kredytów detalicznych, udzielonych w okresie 6 kolejnych miesięcy kalendarzowych przez jedną z instytucji finansowych w Polsce. Jako wystąpienie niewypłacalności traktowane jest 90-dniowe opóźnienie w płatności należnej raty kredytowej[3].

Rozkład w czasie niewypłacalności dla kredytów detalicznych zależy od długości umowy kredytowej [Stepanova, Thomas, 2002, s. 281], dlatego kohortę utworzono jedynie z kredytów o takim samym okresie umowy. Kredyty o innych długościach powinny być analizowane jako oddzielne kohorty.

Oszacowania przeprowadzono na dwa sposoby:

1. Podejście I – analizując liczbę niewypłacalności do liczby kredytów w kohorcie (podejście według ilości).
2. Podejście II – analizując pozostałą do spłaty wartość kredytów, które stały się niewypłacalne, do pozostałej do spłaty wartości wszystkich kredytów, które były w kohorcie na początku danego okresu (podejście według wartości).

Podejście I. Łącznie w sześciu miesięcznych kohortach było 794 przypadków niewypłacalności (9,9%) oraz 5669 wcześniejszych spłat kredytów (70,4%). W celu wyznaczenia wskaźników Altmana dla każdej z sześciu kohort wyznaczono wskaźniki $MMR_{t,i}$, a następnie oszacowania zostały uśrednione $\overline{MMR_t}$ i na ich podstawie wyznaczony został wskaźnik CMR. Wyniki przedstawia tablica 1[4] oraz rysunek 1.

Ze względu na przyjętą definicję niewypłacalności zdarzenie to może wystąpić dopiero od trzeciego miesiąca trwania kredytu. Najwyższe prawdopodobieństwo wystąpienia niewypłacalności (0,0098) występuje w piątym miesiącu trwania kredytu[5] (pod warunkiem że kredyt pozostał w kohorcie do piątego miesiąca), natomiast prawdopodobieństwo, że niewypłacalność wystąpi w ciągu pierwszych pięciu miesięcy trwania kredytu wynosi 0,0243. W analogiczny sposób można interpretować wszystkie pozostałe wartości wyliczonych wskaźników.

---

[3] Szerzej o definicji niewypłacalności w przypadku umowy kredytowej [zob. Jajuga, 2007, s. 222–223].

[4] Ze względu na obszerność tablic wynikowych w pracy zawarto tylko ich fragment, a wybrane prawdopodobieństwa dla wszystkich okresów są prezentowane dodatkowo na rysunku 1.

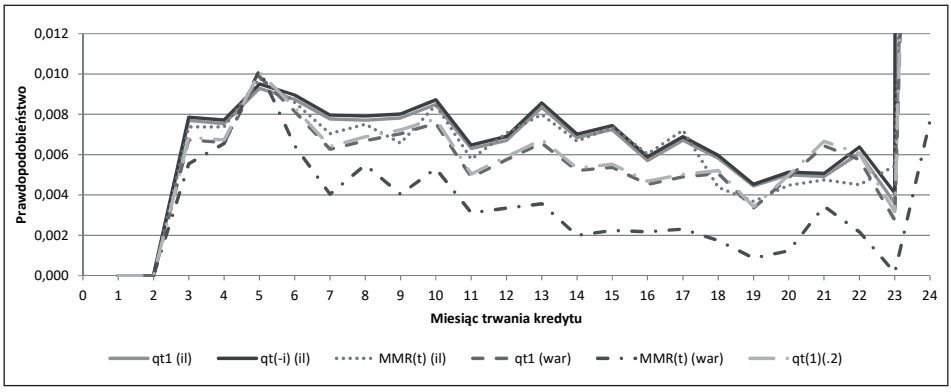
[5] Zaobserwowane w 24 miesiącu wysokie prawdopodobieństwo jest efektem zgrupowania w 24 miesiącu liczby kredytów z niewypłacalnością z tytułu opóźnień płatności rat należnych za miesiące od 21 do 24 włącznie, dlatego wartości oszacowań dla tego miesiąca nie będą w pracy interpretowane.

Tablica 1. Wskaźniki umieralności według Altmana

Czas trwania kredytu (t)	Podejście według ilości		Podejście według wartości	
	$\overline{MMR}(t)$	CMR(t)	$\overline{MMR}(t)$	CMR(t)
1	0,0000	0,0000	0,0000	0,0000
2	0,0000	0,0000	0,0000	0,0000
3	0,0074	0,0074	0,0055	0,0055
4	0,0074	0,0147	0,0066	0,0121
5	0,0098	0,0243	0,0102	0,0221
6	0,0086	0,0327	0,0064	0,0284
…	…	…	…	…
21	0,0047	0,1207	0,0034	0,0712
22	0,0045	0,1246	0,0022	0,0732
23	0,0054	0,1294	0,0001	0,0734
24	0,0664	0,1872	0,0077	0,0805

Źródło: Opracowanie własne.

Rysunek 1. Prawdopodobieństwa wystąpienia niewypłacalności oszacowane metodą tablic wieloprzyczynowych i wskaźnikiem umieralności Altmana



(il) – podejście według ilości, (war) – podejście według wartości
Źródło: Opracowanie własne.

Pierwszym krokiem w budowie tablic trwania było połączenie miesięcznych kohort w jedną, w celu wyeliminowania wahań losowych wynikających z małej liczebności kohort miesięcznych. Następnie zbudowano tablicę standardową dla odejść z kohorty bez względu na przyczynę (tab. 2). Oszacowane wielkości q_t pokazują, jakie jest prawdopodobieństwo, że kredyt przestanie być obsługiwany w miesiącu t (niezależnie od przyczyny), pod warunkiem że kredyt pozostał w kohorcie do tego miesiąca. Największe prawdopodobieństwo wyjścia z kohorty jest więc w 23 miesiącu. Średni czas trwania w kohorcie dla nowych kredytów wynosi 13,37 miesięcy. Średni dalszy czas trwania zależy od czasu, do którego dotrwały kredyty.

Przykładowo, kredyty, które pozostały w kohorcie do 5 miesiąca, będą jeszcze średnio w kohorcie przez 11,45 miesiąca.

Tablica 2. Standardowa tablica trwania dla ryzyka przerwania umowy kredytowej niezależnie od przyczyny (podejście według ilości)

T	l_t	q_t	d_t	L_t	T_t	e_t
1	100 000	0,029	2 919	98 540	1 337 043	13,37
2	97 081	0,036	3 453	95 354	1 238 503	12,76
3	93 627	0,050	4 646	91 304	1 143 149	12,21
4	88 981	0,053	4 683	86 640	1 051 845	11,82
5	84 298	0,056	4 708	81 944	965 205	11,45
6	79 590	0,056	4 460	77 360	883 261	11,10
…		…	…	…	…	…
21	32 845	0,065	2 137	31 776	95 839	2,92
22	30 708	0,102	3 143	29 137	64 062	2,09
23	27 565	0,233	6 422	24 354	34 925	1,27
24	21 143	1,000	21 143	10 571	10 571	0,50

Źródło: Opracowanie własne.

Tablice wieloprzyczynowe pokazują, jaki jest udział każdej z przyczyn w prawdopodobieństwie opuszczenia kohorty. W podejściu według ilości rozważane są dwie przyczyny: niewypłacalność oraz wcześniejsza spłata. W piątym miesiącu trwania kredytu prawdopodobieństwo wystąpienia niewypłacalności dla kredytów, które pozostały w kohorcie i są również narażone na ryzyko wcześniejszej spłaty, wynosi 0,009 (tab. 3), natomiast prawdopodobieństwo, że ostateczną przyczyną opuszczenia przez nie kohorty będzie niewypłacalność (w dowolnym miesiącu od piątego począwszy) wynosi 0,101. Ryzyko wcześniejszej spłaty w piątym miesiącu dla kredytów, które pozostały w kohorcie, wynosi 0,047, a prawdopodobieństwo krańcowe 0,666.

Tablica 3. Tablice wieloprzyczynowe dla ryzyka niewypłacalności i ryzyka wcześniejszej spłaty (podejście według ilości)

T	Tablica wieloprzyczynowa (ryzyko niewypłacalności)				Tablica wieloprzyczynowa (ryzyko wcześniejszej spłaty)			
	$l_t^{(1)}$	$q_t^{(1)}$	$d_t^{(1)}$	$\psi_t^{(1)}$	$l_t^{(2)}$	$q_t^{(2)}$	$d_t^{(2)}$	$\psi_t^{(2)}$
1	9 863	0,000	0	0,099	70 422	0,029	2 919	0,704
2	9 863	0,000	0	0,102	67 503	0,036	3 453	0,695
3	9 863	0,008	720	0,105	64 050	0,042	3 925	0,684
4	9 143	0,008	671	0,103	60 124	0,045	4 012	0,676
5	8 472	0,009	783	0,101	56 112	0,047	3 925	0,666
6	7 689	0,009	696	0,097	52 186	0,047	3 764	0,656

T	Tablica wieloprzyczynowa (ryzyko niewypłacalności)				Tablica wieloprzyczynowa (ryzyko wcześniejszej spłaty)			
	$l_t^{(1)}$	$q_t^{(1)}$	$d_t^{(1)}$	$\psi_t^{(1)}$	$l_t^{(2)}$	$q_t^{(2)}$	$d_t^{(2)}$	$\psi_t^{(2)}$
…	…	…	…	…	…	…	…	…
21	1 876	0,005	161	0,057	11 255	0,060	1 975	0,343
22	1 714	0,006	186	0,056	9 280	0,096	2 957	0,302
23	1 528	0,004	99	0,055	6 323	0,229	6 323	0,229
24	1 429	0,068	1 429	0,068	0	0,000	0	0,000

Źródło: Opracowanie własne.

Tablica z eliminacją wybranej przyczyny została zbudowana dla przypadku eliminacji wcześniejszej spłaty. Tablica ta przedstawia więc hipotetyczny rozkład niewypłacalności dla kohorty kredytów, w których nie ma możliwości wcześniejszej spłaty (np. poprzez ograniczenia zawarte w umowie) (tab. 4).

Tablica 4. Tablica z eliminacją wcześniejszej spłaty (podejście według ilości)

T	$l_t^{(-2)}$	$q_t^{(-2)}$	$L_t^{(-2)}$	$T_t^{(-2)}$	$e_t^{(-2)}$	$g_t^{(-2)}$	$\gamma_t^{(-2)}$
1	100 000	0,000	100 000	2 181 321	21,81	8,44	11,99
2	100 000	0,000	100 000	2 081 321	20,81	8,06	11,59
3	100 000	0,008	99 607	1 981 321	19,81	7,60	11,11
4	99 214	0,008	98 831	1 881 715	18,97	7,15	10,57
5	98 448	0,010	97 980	1 782 883	18,11	6,66	10,01
6	97 512	0,009	97 076	1 684 903	17,28	6,18	9,43
…	…	…	…	…	…	…	…
21	87 633	0,005	87 411	303 916	3,47	0,55	1,61
22	87 189	0,006	86 911	216 505	2,48	0,40	1,31
23	86 633	0,004	86 455	129 594	1,50	0,23	1,00
24	86 278	1,000	43 139	43 139	0,50	0,00	–

(–2) – eliminacja wcześniejszej spłaty jako przyczyny odejść z kohorty
Źródło: Opracowanie własne.

Gdyby wcześniejsze spłaty kredytów nie były możliwe i przy założeniu, że nie wpłynęłoby to na prawdopodobieństwo niewypłacalności, to w piątym miesiącu prawdopodobieństwo niewypłacalności wynosiłoby 0,01, a średni dalszy czas trwania kredytu w kohorcie wynosiłby 18,11 miesięcy i byłby dłuższy od aktualnego o 6,66 miesięcy. Natomiast dla kredytów, które są w kohorcie w piątym miesiącu i z rzeczywistej kohorty odeszłyby z powodu wcześniejszej spłaty, średni dalszy czas trwania wydłużyłby się o 10,01 miesięcy.

Podejście II. Tablice zbudowane w niniejszym badaniu mają za zadanie jedynie zobrazować przydatność stosowania tablic wieloprzyczynowych do badania rozkładu niewypłacalności w czasie, a nie służyć monitorowaniu niewypłacalności i poziomu strat rzeczywistego portfela kredytowego. Z tego powodu przyjęto uproszczenia odnośnie do wyceny analizowanych wartości. Jako wartość kredytu w momencie ($t = 0$) przyjęto sumę rat z 24 miesięcy. Suma wartości wszystkich udzielonych kredytów stanowi początkową wartość kohorty. Wartość ta ulega zmniejszeniu z trzech powodów: niewypłacalności, wcześniejszej spłaty oraz terminowej spłaty (wynikającej z umowy). W przypadku wystąpienia niewypłacalności, wartość straty liczona jest jako suma rat, które pozostały jeszcze do spłaty począwszy od raty, dla której odnotowano opóźnienie[6]. Tak samo obliczane są wartości wcześniejszej spłaty[7]. Wartość spłaty terminowej jest równa sumie rat, które zostały zapłacone w danym okresie, zgodnie z warunkami umów, z pominięciem wcześniejszych spłat. Wartość kohorty na początku kolejnych okresów jest różnicą między wartością portfela na początku poprzedniego okresu i sumą wartości spłat terminowych, spłat wcześniejszych i wartości straty z tytułu niewypłacalności.

W proponowanym podejściu według wartości, w przeciwieństwie do podejścia według ilości, jednostki nie mają takiej samej wagi. Wagą jest wartość pozostałego do spłaty kredytu. W rezultacie prawdopodobieństwa uzyskane zarówno przy wykorzystaniu metody wskaźników Altmana, jak i tablic trwania w podejściu według wartości i według ilości mogą się różnić.

W podejściu według wartości największe prawdopodobieństwo niewypłacalności, według wskaźnika Altmana, jest w piątym miesiącu (tak samo wskazywało podejście według ilości – tab. 1). Prawdopodobieństwo, że do niewypłacalności dojdzie do piątego miesiąca włącznie wynosi 0,022. W każdym miesiącu (poza piątym) oszacowania prawdopodobieństwa niewypłacalności metodą według wartości są niższe od oszacowań metodą według ilości, co prowadzi do niższego wskaźnika CMR(t) w całym okresie. Prawdopodobieństwo niewypłacalności w okresie 24 miesięcy obliczone metodą według ilości wynosi 0,1872, natomiast metodą według wartości 0,0805 (tab. 1). Nie jest to jednak prawidłowość, której należy zawsze oczekiwać, stosując podejście według ilości i według wartości. Wartość

[6] Nie uwzględniono stopy odzysku (*Loss Given Default* – LGD), która w praktyce ma kluczowe znaczenie w szacowaniu wielkości straty

[7] Innym sposobem, który można zastosować, jest doliczenie części kapitałowej spłacanych wcześniej rat do wartości spłat terminowych. Wówczas wartość wcześniejszej spłaty obejmowałaby tylko część odsetkową raty, której w wyniku wcześniejszej spłaty bank nie zarobi. Jest to więc faktyczna strata banku (utrata przewidywanego zysku). Należy również analizowane wartości zdyskontować na ustalony moment czasu.

wskaźnika obliczanego metodą według wartości będzie się różnie kształtować w zależności od zróżnicowania wielkości kredytów w kohorcie i od tego, w jakich miesiącach trwania kredytów dojdzie do niewypłacalności dużych kredytów. Zestawienie wskaźników obliczanych obiema metodami jest więc źródłem dodatkowej informacji dla kredytodawcy.

Tablica 5 przedstawia standardową tablicę trwania dla zdarzeń bez względu na przyczynę dla której funkcje wyliczono metodą według wartości. Przeciętny czas trwania umowy kredytowej wynosi 8,6 miesiąca.

Tablica 5. Standardowa tablica umieralności łącznie dla ryzyka niewypłacalności, wcześniejszej spłaty i regularnej spłaty (podejście według wartości)

T	l_t	q_t	d_t	L_t	T_t	e_t
1	**100 000**	0,065	6 455	96 772	859 558	8,60
2	93 545	0,071	6 623	90 233	762 786	8,15
3	86 921	0,086	7 512	83 165	672 553	7,74
4	79 409	0,089	7 047	75 885	589 388	7,42
5	72 362	0,099	7 132	68 796	513 502	7,10
6	65 230	0,099	6 479	61 990	444 707	6,82
…	…	…	…	…	…	…
21	5 790	0,301	1 746	4 917	10 350	1,79
22	4 045	0,394	1 594	3 247	5 433	1,34
23	2 450	0,608	1 490	1 705	2 185	0,89
24	960	1,000	960	480	480	0,50

Źródło: Opracowanie własne.

Na warunkowe prawdopodobieństwo zmniejszenia się kohorty w danym okresie wpływa prawdopodobieństwo wszystkich trzech analizowanych przyczyn. Przykładowo w piątym miesiącu prawdopodobieństwo zmniejszenia się wartości kohorty wynosi 0,099 (tab. 5), na które składają się przedstawione w tablicy 6 prawdopodobieństwa: niewypłacalności (0,01), wcześniejszej spłaty (0,041) oraz terminowej spłaty (0,047).

Przy założeniu wyeliminowania wcześniejszych spłat jako przyczyny zmniejszenia wartości kohorty prawdopodobieństwo, że kredyt opuści kohortę w piątym miesiącu (z powodu niewypłacalności lub terminowej spłaty), pod warunkiem że kredyt pozostał w kohorcie do piątego miesiąca, wynosi 0,059 (tab. 7). Natomiast prawdopodobieństwo, że kredyt opuści kohortę z powodu niewypłacalności wynosi 0,01. Przeciętny dalszy czas trwania kredytów w kohorcie, przy wyeliminowaniu wcześniejszych spłat, wzrósłby o 2,57 miesiąca, a dla jednostek, które z rzeczywistej kohorty odeszły z powodu wcześniejszej spłaty, o 6,82 miesiąca.

Tablica 6. Tablice wieloprzyczynowe dla ryzyka niewypłacalności, wcześniejszej spłaty oraz terminowej spłaty (podejście według wartości)

T	Tablica wieloprzyczynowa (ryzyko niewypłacalności)				Tablica wieloprzyczynowa (ryzyko wcześniejszej spłaty)				Tablica wieloprzyczynowa (terminowa spłata)			
	$l_o^{(1)}$	$d_t^{(1)}$	$q_t^{(1)}$	$\psi_t^{(1)}$	$l_0^{(2)}$	$d_t^{(2)}$	$q_t^{(2)}$	$\psi_t^{(2)}$	$l_o^{(3)}$	$d_t^{(3)}$	$q_t^{(3)}$	$\psi_t^{(3)}$
1	4 891	0	0,000	0,049	38 371	2 388	0,024	0,384	56 738	4 067	0,041	0,567
2	4 891	0	0,000	0,052	35 982	2 672	0,029	0,385	52 671	3 951	0,042	0,563
3	4 891	585	0,007	0,056	33 310	3 146	0,036	0,383	48 720	3 781	0,044	0,561
4	4 306	524	0,007	0,054	30 164	2 904	0,037	0,380	44 939	3 618	0,046	0,566
5	3 782	714	0,010	0,052	27 260	2 984	0,041	0,377	41 320	3 433	0,047	0,571
6	3 067	531	0,008	0,047	24 275	2 684	0,041	0,372	37 887	3 264	0,050	0,581
…	…	…	…	…	…	…	…	…	…	…	…	…
21	128	37	0,006	0,022	1 229	360	0,062	0,212	4 433	1 348	0,233	0,766
22	91	23	0,006	0,022	869	346	0,086	0,215	3 085	1 225	0,303	0,763
23	67	7	0,003	0,027	523	523	0,213	0,213	1 860	960	0,392	0,759
24	61	61	0,063	0,063	0	0	0,000	0,000	900	900	0,937	0,937

Źródło: Opracowanie własne.

Tablica 7. Tablica z eliminacją wcześniejszej spłaty (podejście według wartości)

T	$l_t^{(-2)}$	$q_t^{(-2)}$	$q_t^{(1)(-2)}$	$L_t^{(-2)}$	$T_t^{(-2)}$	$e_t^{(-2)}$	$g_t^{(-2)}$	$\gamma_t^{(-2)}$
1	100 000	0,041	0,000	97 941	1 162 922	11,63	3,03	7,91
2	95 883	0,043	0,000	93 828	1 064 980	11,11	2,95	7,68
3	91 773	0,051	0,007	89 425	971 152	10,58	2,84	7,42
4	87 076	0,053	0,007	84 761	881 728	10,13	2,70	7,12
5	82 447	0,059	0,010	80 033	796 966	9,67	2,57	6,82
6	77 619	0,059	0,008	75 312	716 933	9,24	2,42	6,50
…	…	…	…	…	…	…	…	…
21	15 567	0,248	0,007	13 638	31 707	2,04	0,25	1,17
22	11 709	0,325	0,006	9 809	18 069	1,54	0,20	0,93
23	7 908	0,456	0,003	6 107	8 260	1,04	0,15	0,71
24	4 306	1,000	0,063	2 153	2 153	0,50	0,00	–

(–2) – eliminacja wcześniejszej spłaty jako przyczyny odejść z kohorty
Źródło: Opracowanie własne.

Wszystkie zbudowane powyżej tablice trwania to tzw. tablice surowe. Obserwowane wahania oszacowań prawdopodobieństw w kolejnych okresach są w dużej mierze efektem losowej realizacji zdarzeń. W mniejszych kohortach należy spodziewać się większych wahań. W celu wyeliminowania nieregularności będących efektem błędów losowych można zastosować metody wyrównywania funkcji tablicowych [Jackowska, 2013, s. 213–244].

Zakończenie

Jak przedstawiono na rysunku 1, zastosowane metody szacowania prawdopodobieństwa niewypłacalności dały różniące się wyniki. Jest to efektem ujmowania przez powyższe miary różnych aspektów ryzyka. Zestawienie wyników uzyskanych różnymi metodami jest więc źródłem dodatkowej informacji o naturze ryzyka. Niewypłacalność w ostatnich miesiącach trwania kredytu, gdy większość kapitału i odsetek jest już spłacona, jest znacznie mniej dotkliwym finansowo ryzykiem dla banku niż wystąpienie niewypłacalności w pierwszych miesiącach trwania kredytów. Różnice te uwidacznia podejście według wartości.

Na wielkość różnic w oszacowaniach metodą według ilości oraz metodą według wartości ma wpływ rozkład w czasie niewypłacalności dużych kredytów. W badanej kohorcie największe niewypłacalności występowały w okresach, w których również liczba niewypłacalności była największa, co przełożyło się na spójną ocenę tego, w których okresach są najwyższe prawdopodobieństwa niewypłacalności. Natomiast gdyby bardzo duże niewypłacalności wystąpiły w okresach o niskiej częstotliwości tego zdarzenia, to istotnej zmianie uległby kształt krzywej prawdopodobieństwa. Podejście według wartości daje więc obraz niewypłacalności w danej kohorcie, ale nie powinno ono służyć prognozom niewypłacalności w kolejnych kohortach. Do celów prognostycznych lepsze jest podejście według ilości.

Oba prezentowane podejścia do badania niewypłacalności mogą służyć porównywaniu ryzyka niewypłacalności w różnych kohortach kredytów (tzw. analiza vintage).

Tablice trwania oprócz prawdopodobieństwa niewypłacalności, które szacowane jest również w metodzie wskaźników Altmana, zawierają dodatkowo szereg funkcji opisujących populację. Pewnym mankamentem metody tablic trwania jest brak możliwości wyznaczenia niektórych funkcji w ostatnim przedziale.

Przedstawione w artykule metody mogą być nie tylko użytecznym narzędziem w ocenie niewypłacalności rozumianej jako realizacja ryzyka kredytowego. Posłużyć mogą również do badania rozkładów w czasie przyczyn likwidacji firm czy innych badań, w których oprócz badanego zdarzenia występują zdarzenia konkurujące, które również mogą być przyczyną opuszczenia kohorty przez jednostkę.

Literatura

Altman E. (1989), *Measuring Corporate Bond Mortality and Performance*, „The Journal of Finance", Vol. 44, No. 4.

Altman E., Hotchkiss E. (2006), *Corporate Financial Distress and Bancruptcy*, 3rd ed., Wiley, New Jersey.

Anderson R. (1999), *U.S. decennial life tables for 1989–91*, Vol. 1, No. 4, United States life tables eliminating certain causes of death. National Center for Health Statistics, Hyattsville, Meryland, https://www.cdc.gov/nchs/data/lifetables/life89_1_4.pdf, dostęp: 20.10.2017.

Balicki A. (2006), *Analiza przeżycia i tablice wymieralności*, PWE, Warszawa.

Bolesławski L. (1973), *Budowa tablic trwania życia. Teoria i praktyka*, GUS, Warszawa.

Dirick L., Claeskens G., Baesens B. (2017), *Time to default in credit scoring using survival analysis: a benchmark study*, „Journal of Operational Research Society", Vol. 68, No. 6.

Flores J., Navarro V., Cedillo I. (2010), *An actuarial approach to pricing Mortgage Insurance considering simultaneously mortgage default and prepayment*, International Congress of Actuaries, Cape Town, http://www.actuaries.org/EVENTS/Cogresses/Cape_Town/Papers/Financial%20Risks%20(AFIR)/49_final%20paper_Elizondo,%20et%20al.pdf, dostęp: 20.10.2017.

Gurgul H., Zając P. (2016), *Modelowanie nieparametryczne czasu przeżycia przedsiębiorstw Małopolsce*, „Wiadomości Statystyczne", nr 12 (667).

Jackowska B. (2013), *Modele dalszego trwania życia oraz ich zastosowania w przypadku osób starszych*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.

Jackowska B. (2015), *Analiza kohortowa czasu istnienia mikroprzedsiębiorstw w Gdańsku*, „Zarządzanie i Finanse", nr 4, cz. 2.

Jajuga K. (2004), *Modele ryzyka kredytowego a kredyty hipoteczne*, w: K. Jajuga i Z. Krysiak (red.), *Zarządzanie ryzykiem kredytowym wierzytelności hipotecznych*, Związek Banków Polskich, Warszawa.

Jajuga K. (red.) (2007), *Zarządzanie ryzykiem*, Wydawnictwo Naukowe PWN, Warszawa.

Markowicz I. (2012), *Statystyczna analiza żywotności firm*, Wyd. Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Ptak-Chmielewska A. (2010), *Analiza przeżycia przedsiębiorstw w Polsce na przykładzie wybranego województwa*, w: P. Dittmann (red.), *Prognozowanie w zarządzaniu firmą*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", nr 103.

Saunders A., Allen L. (2002), *Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms*, 2nd ed., Wiley, New York.

Saunders A., Allen L. (2010), *Credit risk management in and out of the financial crisis: new approaches to value at risk and other paradigms*, 3rd ed., Wiley, New York.

Saunders A. (2001), *Metody pomiaru ryzyka kredytowego*, Wydawnictwo Oficyna Ekonomiczna, Kraków.

Siegel J., Swanson D. (2004), *The methods and materials of demography*, Elsevier Academic Press, Amsterdam.

Stepanova, M., Thomas L. (2002), *Survival Analysis Methods for Personal Loan Data*, „Operations Research", Vol. 50, No. 2.

Streszczenie

Szacowanie prawdopodobieństwa niewypłacalności rozumianej jako trwałe zaprzestanie spłat rat zobowiązań należy do jednych z najważniejszych zadań banków w zakresie badania ryzyka kredytowego. W obliczu ostatnich zmian

prawnych banki zostały zobowiązane do szacowania prawdopodobieństwa niewypłacalności w całym okresie trwania umowy kredytowej, tak aby mogło ono posłużyć do wyznaczenia wartości oczekiwanej straty w całym okresie (*lifetime expected credit losses*). Niewypłacalność jest jedną z przyczyn zakończenia umowy kredytowej, do pozostałych przyczyn należą: wcześniejsza całkowita spłata oraz upływ terminu, na który zawarta była umowa. Modelując prawdopodobieństwo niewypłacalności w kolejnych okresach trwania umowy, należy uwzględnić możliwość zakończenia umowy z innych powodów. W artykule przedstawiono dwie metody, które uwzględniają powyższy warunek: metodę wskaźników umieralności Altmana oraz propozycję zaczerpniętą z demografii – wieloprzyczynowych tablic trwania. Zaproponowano dwa podejścia do badania wielkości kohorty kredytów: według ilości – oparte na liczbie umów – oraz według wartości – biorące pod uwagę wartość udzielonych i pozostających do spłaty zobowiązań.

W części empirycznej pracy powyższe metody i podejścia zastosowano do badania rozkładu w czasie niewypłacalności kohorty kredytów detalicznych udzielonych przez jedną z instytucji finansowych w Polsce. Omówiono przyczyny różnic w uzyskanych wynikach i wskazano obszary zastosowań proponowanych metod.

Słowa kluczowe
ryzyko kredytowe, wskaźniki umieralności, prawdopodobieństwo niewypłacalności, zdarzenia konkurujące, kredyty detaliczne

Multi-decrement tables in the assessment of the probability of default. Cohort approach (Summary)

Default is defined as the sustainable cease of the payment of instalments. The estimation of the probability of default is one of the banks' most important tasks in the credit risk assessment. Due to the last changes in legal requirements banks are obliged to evaluate the probability of lifetime default in order to estimate lifetime expected credit losses. Default is one of the reasons for credit termination, others are: prepayment and maturity. Evaluation of the probability of default in following periods requires taking into account a loan termination due to other reasons. In the paper two methods were presented which deal with this condition: first, Altman's mortality rates and multi-decrement life tables, and the second one drawn from the demography. Two approaches were proposed to measure the size of the cohort: the quantitative approach – based on the number of units, and amount approach based on the value of credits and the value of the repayment.

In the empirical part of the study, the above methods and approaches were applied to model the time distribution of the probability of default in the cohort of retail credits granted by one of the financial institutions in Poland. The reasons for the differences in the obtained results have been discussed as well as the potential application areas of the proposed methods.

Keywords
credit risk, mortality rates, probability of default, competing risks, retail credits

Tomasz Jastrzębski*

Zastosowanie symulacji Monte Carlo z korektą Imana-Conovera w pomiarze ryzyka systemowego

Wstęp

Ryzyko systemowe może być rozumiane jako możliwość wystąpienia wielu jednoczesnych nieprawidłowości w funkcjonowaniu dużych instytucji finansowych [Huang i inni, 2009]. Kwestia jego definiowania oraz pomiaru nabrała szczególnego znaczenia w okresie ostatniego kryzysu finansowego w latach 2007–2009. Jako główne źródło tego kryzysu wskazuje się między innymi na rosnący poziom zadłużenia państw, zarówno na poziomie makro-, jak i mikroekonomicznym. W związku z powyższym pierwszą reakcją Parlamentu Europejskiego było rozpoczęcie prac nad zmianą uregulowań związanych z funkcjonowaniem sektora finansowego [Basel Committee on Banking Supervision, 2011]. Efektem tych prac jest wprowadzony 26 czerwca 2013 r. (a obowiązujący od 1 stycznia 2014 r.) pakiet CDR IV/CRR, składający się z dyrektywy Capital Requirements Directive IV oraz rozporządzenia Capital Requirements Regulation, mający za zadanie zwiększenie bezpieczeństwa systemu finansowego, zarówno na poziomie całego systemu, jak i indywidualnego podmiotu. W związku z wprowadzeniem nowych regulacji ostrożnościowych nasuwa się więc pytanie: jak zmiana tych uwarunkowań prawnych wpłynie na poziom wrażliwości ekspozycji na ryzyko podmiotów z sektora bankowego?

Złożoność i zmienność ram teoretycznych i sposobu pomiaru ryzyka systemowego powoduje, że niezwykle trudne jest określenie i kwantyfikowanie czynników ryzyka systemowego w sektorze bankowym[1]. Dodatkowo w swej istocie ryzyko to ma naturę zdarzenia rzadkiego i tym samym nie mają zastosowania miary klasyczne. Uzupełniając to spostrzeżenie o fakt braku danych oraz ich niską częstotliwość, można dojść do wniosku o konieczności zastosowania metod symulacyjnych, których zaletą jest możliwość projektowania różnych scenariuszy wydarzeń.

* Mgr, Katedra Ekonometrii, Wydział Zarządzania, Uniwersytet Gdański, ul. Armii Krajowej 101, 81-824 Sopot, tomasz.jastrzebski@ug.edu.pl

[1] Ryzyko systemowe – mimo że może być spowodowane przez inne rodzaje ryzyka, np. kredytowe lub płynności – różni się od innych rodzajów ryzyka, ponieważ identyfikowane jest poprzez skutki, a nie źródło [Kabza, 2012, s. 3]. Jest dynamiczne i ewoluuje, co z kolei wynikać może ze zmieniających się regulacji prawnych sektora bankowego.

Celem artykułu jest przeprowadzenie symulacji mierników ryzyka systemowego zdefiniowanych jako odległość między różnymi stanami wartości wskaźników a ich wartościami wzorcowymi i antywzorcowymi. W związku z tak sformułowanym celem postawiono hipotezę badawczą, iż wraz ze wzrostem poziomu ryzyka systemowego obserwowana jest większa wrażliwość wartości oczekiwanych wskaźników niż ich wariancji. Jako narzędzie w procesie weryfikacji hipotezy badawczej w zaprojektowanym eksperymencie użyta została symulacja Monte Carlo z korektą Imana-Conovera.

1. Problemy pomiaru ryzyka systemowego

System bankowy od zawsze był ważnym sektorem uzupełniającym gospodarkę, gdyż główną rolą systemu bankowego jest pośrednictwo finansowe między podmiotami gospodarczymi. Umożliwiając alokację środków podmiotom posiadającym nadwyżki finansowe przy jednoczesnym udostępnianiu tych środków podmiotom mającym problemy płynnościowe, system bankowy przyczynia się do znacznego zwiększenia efektywności wykorzystania zasobów gospodarki jako całości. Jednak system bankowy nie jest tworem idealnym i co pewien czas można zaobserwować efekty nasileń nieprawidłowości w jego działaniu, które nazywamy mianem kryzysów finansowych. Te anomalie w ciągu ostatnich kilkudziesięciu lat przybierały coraz większe rozmiary, doprowadzając do tak zwanego efektu przelewania (*spillover effect*), czyli transmisji kryzysu z sektora finansowego na pozostałe sektory gospodarki, a w dobie globalizacji również na gospodarki innych krajów [Hellwig, 2009]. Mając na uwadze powyższe przesłanki, organy nadzoru finansowego i banki centralne zaczęły skupiać swoją uwagę na monitorowaniu i regulowaniu sektora bankowego jako całości. Ich rozporządzenia były w przeszłości skupione na zapewnieniu stabilności poszczególnych banków, natomiast po kryzysie na rynku dot-com pojawiły się tendencje do koncentrowania się na zapewnieniu stabilności całego systemu bankowego, które w literaturze przedmiotu są znane jako perspektywy makroostrożnościowe regulacji bankowych [Borio, 2003]. Wskaźniki dedykowane pomiarowi ryzyka systemowego zaczęły być konstruowane zarówno na poziomie indywidualnym, jak i zagregowanym, przy czym uwzględnia się również korelacje pomiędzy portfelami aktywów banków [Lehar, 2005]. Spotykane są też konstrukcje wskaźników opartych nie o dane z bilansów banków, ale o dane z rynku, takie jak *credit default swaps* (CDS) i ceny akcji banków [Huang i inni, 2009]. Prowadzone są również badania nad wpływem niepokojów na rynkach finansowych na sferę gospodarki realnej [Acharya i inni, 2010]. Konstruowane są wskaźniki ryzyka systemowego, które agregują informacje dotyczące recesji spośród

indywidualnych wskaźników pochodzących z rynku w oparciu o regresję kwantylową [Giglio i inni, 2016]. W literaturze przedmiotu można znaleźć przykłady badań poświęconych ryzyku systemowemu, w których wykorzystano metody symulacyjne. Przykładowo w opracowaniu [Brämer i inni, 2014] przedstawiono zastosowanie modelu sieciowego, w którym zależności między instytucjami finansowymi wyrażone są przez wielkość wzajemnych zobowiązań i kolejno przeprowadzono symulacje polegające na eksperymentalnym wprowadzeniu szoku obejmującego jedną instytucję i badaniu, jak wpływa to na poziom wrażliwości całego systemu w postaci bankructwa pozostałych instytucji w następnych okresach.

2. Symulacje

Metody symulacyjne są szeroko stosowaną grupą narzędzi służących do badania złożonych związków między zmiennymi, także w naukach ekonomicznych. Ich niewątpliwą zaletą jest bezkosztowa możliwość odtworzenia warunków charakterystycznych dla danego zjawiska lub obiektu przy użyciu modelu matematycznego z wykorzystaniem programu komputerowego. Można wskazać jedną z definicji, w której eksponowana jest jej rola w procesie modelowania: „podstawą jej (metody symulacyjnej – przyp. autora) jest modelowanie statystycznego eksperymentu za pomocą środków techniki obliczeniowej i rejestracja charakterystyk liczbowych otrzymanych z tego eksperymentu" [Buslenko i inni, 1967, s. 5]. Symulacje dają zatem nie tylko możliwość analizy badanego zjawiska bez ingerencji w jego stan, ale również wirtualnego projektowania skutków odejścia od typowego przebiegu danego zjawiska i analizy różnych wariantów jego przebiegu.

Podejmując się próby klasyfikacji metod symulacyjnych, można wskazać, iż najważniejszym kryterium ich podziału jest występowanie elementu losowości w modelu przedstawiającym dane zjawisko. W związku z powyższym można wyróżnić:

1. Metody deterministyczne, w których stan symulowanego zjawiska jest w sposób jednoznaczny i powtarzalny przypisany do wartości zmiennych danych na wejściu, pozyskanych empirycznie lub podlegających regułom założonych teoretycznych scenariuszy, co oznacza że nie zawiera elementu losowości. W naukach ekonomicznych są one rzadko wykorzystywane ze względu na dużą złożoność systemów ekonomicznych i występujących w nich elementów losowych.
2. Metody stochastyczne, w których uwzględnia się składnik losowy wraz z właściwościami jego rozkładu, bądź jakiś inny element losowości (metody repróbkowania). Są to metody najczęściej

wykorzystywane, głównie ze względu na fakt, że wiele zjawisk jest niezwykle złożona. Do stochastycznych metod symulacyjnych zaliczamy:

– metody nieparametryczne (np. *jackknife*, *bootstrap*), używane gdy nie mamy pewności co do rozkładu prawdopodobieństwa badanych zmiennych, na przykład z powodu dysponowania małą próbą. Wnioskowanie statystyczne jest wtedy przeprowadzane nie na próbie, ale na sztucznie utworzonej populacji generalnej, której elementy są w określony sposób wylosowywane ze znanej próby;
– metody parametryczne, które określa się mianem metod Monte Carlo [Jermakow, 1976]. Można je określić jako metody polegające na generowaniu zmiennych losowych o założonym rozkładzie prawdopodobieństwa w celu oszacowania parametrów rozkładów uzależnionych od nich zmiennych losowych. Użycie tej metody polega zazwyczaj na konstrukcji takiego procesu losowego, którego parametry rozkładu równe są szukanym wielkościom dla danego zagadnienia. Metoda ta, rozwijana od 1949 r.[2], znalazła wiele zastosowań, takich jak: numeryczne obliczanie całek (w tym wielokrotnych), czy rozwiązywanie zadań masowej obsługi i wiele innych. Istnieją różnorodne rozwinięcia tej metody poprawiające jej właściwości w określonych przypadkach. Są to między innymi: próbkowanie warstwowe stosowane w przypadku populacji charakteryzującej się dużym zróżnicowaniem oraz jej rozwinięcie na przypadek wielowymiarowy, czyli próbkowanie hipersześcianu łacińskiego, korekta Imana-Conovera [Iman, Conover, 1982] pozwalająca wprowadzić do wygenerowanego wektora losowego pożądaną macierz korelacji przy zachowaniu rozkładów brzegowych zmiennych losowych będących jego składowymi, czy metody wykorzystujące teorię kopuł [Haas, 1999].

3. Metodyka badania

3.1. Definicja zmiennych kontrolnych, miary ryzyka

Niech $X = (X_1, X_2, X_3, X_4)$ będzie wektorem losowym o składowych będących zmiennymi kontrolnymi procesu ekspozycji na ryzyko systemowe. Składowymi tego wektora są zmienne dobrane w taki sposób, aby możliwe było przy ich użyciu monitorowanie ekspozycji danego obiektu na ryzyko systemowe.

Miarami ryzyka w proponowanym podejściu do pomiaru ryzyka systemowego będą: wartość oczekiwana wektora losowego, którego składowymi

[2] Metoda Monte Carlo została opracowana w latach 1943–1945 przez Stanisława Ulama i zastosowana do modelowania wyniku zderzenia cząstki o wysokiej energii z jądrem złożonym, przyczyniając się tym samym do konstrukcji bomby atomowej.

są zmienne losowe reprezentowane przez wybrane wskaźniki, czyli $E(X) = (E(X_1), E(X_2),\ldots, E(X_N))$, gdzie $E(X_i)$ są całkami jednowymiarowego rozkładu brzegowego w granicach $\langle d_i, g_i\rangle$, $i = 1, 2,\ldots, N$, oraz wariancja wektora losowego X, będąca wektorem wariancji każdego z tych wskaźników czyli $Var(X) = (Var(X_1), Var(X_2),\ldots, Var(X_N))$, gdzie $Var(X_i)$ są całkami jednowymiarowego rozkładu brzegowego w granicach $\langle d_i, g_i\rangle$, $i = 1, 2,\ldots, N$ [Zemke, 2009].

3.2. Stany ryzyka, unormowane stany ryzyka, konstrukcja wektora ryzyka

Niech dla każdego $\langle d_i, g_i\rangle$, $i = 1, 2,\ldots, N$ przedział $\langle d_i^w, g_i^w\rangle$ oznacza wzorcowy, czyli pożądany przez decydenta przedział zmienności i-tej zmiennej kontrolnej. Natomiast $\langle d_i^{aw}, g_i^{aw}\rangle$ niech oznacza dla każdego $i = 1, 2,\ldots, N$ antywzorcowy, czyli uznany przez decydenta za zagrażający realizacji celów strategicznych przedział zmienności i-tej zmiennej kontrolnej. Wtedy wzorzec wektora wartości oczekiwanych $E^w(X)$ jest wektorem składowych $E^w(X_i)$, będących całkami jednowymiarowego rozkładu brzegowego w granicach $\langle d_i^w, g_i^w\rangle$ dla $i = 1, 2,\ldots, N$. Podobnie definiowany jest wzorzec wektora wariancji $Var^w(X)$ jako wektor składowych $Var^w(X_i)$. Analogicznie skonstruowane są $E^{aw}(X)$, $Var^{aw}(X)$, czyli odpowiednio antywzorzec wektora wartości oczekiwanych oraz antywzorzec wektora wariancji.

Aby zbudować wektor ryzyka, należy skonstruować miarę informującą, w jakim stopniu realizowane są założenia strategiczne w danej fazie kontroli. Oznaczmy zatem $E^{(t)}(X)$, $Var^{(t)}(X)$ jako odpowiednio wartość oczekiwaną i wariancję w t-tej fazie kontroli. Wtedy odległość $d^wE^{(t)} = d_M(E^{(t)}(X), E^w(X))$[3] mierzy stan poziomu wartości oczekiwanych zmiennych kontrolnych w t-tej fazie kontroli względem ich wzorca. Przez analogię odległość $d^{aw}E^{(t)} = d_M(E^{(t)}(X), E^{aw}(X))$ mierzy stan poziomu wartości oczekiwanych zmiennych kontrolnych w t-tej fazie kontroli względem ich antywzorca.

Podobnie odległości $d^wVar^{(t)} = d_M(Var^{(t)}(X), Var^w(X))$ oraz $d^{aw}Var^{(t)} = d_M(Var^{(t)}(X), Var^{aw}(X))$ mierzą stan poziomu wariancji zmiennych kontrolnych w t-tej fazie kontroli względem ich wzorca i antywzorca.

Następnie zdefiniowane zostają unormowane miary ryzyka. Niech:

- $R_E^{(t)} = \dfrac{d^{aw}E^{(t)}}{d^wE^{(t)} + d^{aw}E^{(t)}}$ oznacza stopień realizacji założeń wzorcowych wartości oczekiwanej w t-tej fazie kontroli.
- $R_{Var}^{(t)} = \dfrac{d^{aw}Var^{(t)}}{d^wVar^{(t)} + d^{aw}Var^{(t)}}$ oznacza stopień realizacji założeń wzorcowych wariancji w t-tej fazie kontroli [Jastrzębski, 2015, s. 53–54].

[3] Odległością w przestrzeni probabilistycznej (Ω, F, P) dwóch wektorów $X = (X_1, X_2,\ldots, X_N)$ oraz $Y = (Y_1, Y_2,\ldots, Y_N)$ w sensie Mahalanobisa nazywamy $d_M(X,Y) = \sqrt{(X-Y)\cdot C^{-1}\cdot(X-Y)^T}$, gdzie C jest macierzą kowariancji wektorów losowych X i Y.

Należy zauważyć, że miary $R_E^{(t)}$ oraz $R_{Var}^{(t)}$ przyjmują wartości z przedziału $\langle 0,1\rangle$. Wektor $[R_E^{(t)}, R_{Var}^{(t)}]$ jest unormowanym wektorem ryzyka w t-tej fazie kontroli, który znajduje się w kwadracie $\langle 0,1\rangle \times \langle 0,1\rangle$. Wartości kolejnych współrzędnych tego wektora bliskie kresu górnego poszczególnych przedziałów określoności świadczą o dobrej sytuacji badanego podmiotu w kontekście ekspozycji na ryzyko systemowe, wartości bliskie dolnego kresu świadczą o wysokim poziomie ryzyka systemowego.

3.3. Symulacje Monte Carlo, korekta Imana-Conovera

Dobrze znane metody symulacji Monte Carlo, wraz z ich rozszerzeniami, takimi jak próbkowanie warstwowe czy próbkowanie metodą hipersześcianu łacińskiego, przy generowaniu wektora losowego nie uwzględniają zależności pomiędzy zmiennymi losowymi. Procedurą eliminującą tę niedoskonałość jest podejście zaproponowane przez Imana i Conovera [1982], będące szeregiem przekształceń algebraicznych, stanowiących uzupełnienie dowolnego ze schematów próbkowania o algorytm wprowadzania do wygenerowanego uprzednio wektora losowego zadanych korelacji między zmiennymi losowymi, będącymi składowymi tego wektora, przy zachowaniu ich rozkładów brzegowych.

Algorytm postępowania można przedstawić następująco [Iman, Conover, 1982]:

1) generowanie wektora losowego X dowolną metodą generowania danych losowych taką, że zmienne losowe tego wektora są nieskorelowane i cechują się pożądanymi rozkładami brzegowymi;
2) dobór macierzy C będącej pożądaną macierzą korelacji, którą powinna cechować się transformacja macierzy X;
3) wyznaczenie macierzy P, takiej że $C = P \cdot P^T$[4];
4) wykonanie przekształcenia $X \cdot P^T$, w wyniku którego otrzymywana jest poszukiwana transformacja macierzy X, mająca oczekiwaną macierz korelacji C.

4. Symulacje modelu ryzyka systemowego – badanie empiryczne

4.1. Materiał badania

Przeprowadzone badanie objęło trzy obiekty: Alior Bank, BZ WBK oraz PKO BP. Są to banki z pierwszej dziesiątki polskich banków ze względu na wartość aktywów. Obiekty badania są zróżnicowane pod względem:

– aktywów (największy z nich, a zarazem największy bank działający na polskim rynku, PKO BP ma aktywa około dwa razy większe niż BZ WBK i ponad czterokrotnie większe niż Alior Bank),

[4] Ponieważ macierz C jest dodatnio określona i symetryczna, zatem istnieje taka dolna trójkątna macierz P, że $C = P \cdot P^T$. Macierz P można uzyskać na przykład przy użyciu rozkładu Choleskiego.

– doświadczenia w działalności na rynku polskim (najstarszym jest PKO BP, którego korzenie sięgają 1919 r., najmłodszym Alior Bank działający od końca 2008 r.),
– struktury własnościowej (PKO BP i Alior są w ok. 30% w posiadaniu Skarbu Państwa, BZ WBK należy do hiszpańskiej grupy Santander).

Szeregi czasowe mają częstotliwość kwartalną i obejmują okres 2014Q1 – 2016Q3. Dane pochodzą ze sprawozdań kwartalnych banków oraz raportów ich adekwatności kapitałowej.

4.2. Założenia modelu

Prezentowany model ryzyka wymaga przyjęcia następujących założeń:

1. Matematycznym modelem ryzyka jest wektor losowy $X = (X_1, X_2, X_3, X_4)$ zmiennych kontrolnych procesu ekspozycji na ryzyko systemowe. Składowymi tego wektora są:
   – X_1 – współczynnik wypłacalności CET1, zdefiniowany w pakiecie CDR IV / CRR [Basel Committee on Banking Supervision, 2011]. Jest on jednym z podstawowych wskaźników sygnalizujących złą sytuację instytucji finansowej z punktu widzenia ekspozycji na ryzyko. Komisja ustaliła poziom 4,5% jako wartość graniczną, poniżej której stan kapitałów najwyższej jakości postrzegany jest jako zagrażający możliwości danego podmiotu do sprostania warunkom kryzysu na rynku finansowym i w konsekwencji umożliwia efekt przelewania. Dlatego w badaniu przedział $\langle 0 ; 0{,}045 \rangle$ uznano za przedział antywzorca. Zgodnie z zaleceniami Komisji Nadzoru Finansowego w sprawie polityki dywidendowej [Komisja Nadzoru Finansowego, 2014] banki ze znaczącym udziałem w rynku depozytów sektora niefinansowego mogą wypłacać dywidendę do 100% zysku wypracowanego, gdy poziom współczynnika CET1 jest wyższy od 12%, tj. 9% wymaganego przez KNF minimalnego poziomu powiększonego o 3% buforu ryzyka systemowego. Można by zatem uznać, że poziom zwiększony o kolejne 3 p.p. byłby poziomem bardzo bezpiecznym, który w tym badaniu jest definiowany jako wzorcowy. Z tej przyczyny w badaniu przyjęto przedział $\langle 0{,}15 ; 0{,}2 \rangle$ jako przedział wzorca;
   – X_2 – relacja kredytów do depozytów jest jednym z najprostszych wskaźników określających płynność banku i prowadzoną przez bank politykę kredytową[5]. Wartość 100% i więcej jest osiągana podczas prowadzenia agresywnej polityki kredytowej, która obarczona jest większym ryzykiem niewypłacalności banku

[5] Zdecydowanie lepszymi miarami płynności byłyby wskaźniki narzucone przez Bazyleę III, tzn. LCR i NSFR. Wskaźniki te nie zostały wykorzystane w tym badaniu, ze względu na brak dostępu do tych danych.

związanym ze zwiększoną wrażliwością danej instytucji na nieprzewidziane zjawiska zewnętrzne (np. masowe wycofywanie depozytów przez Greków podczas kryzysu zadłużenia w Grecji w 2009 r.). W badaniu przedział $\langle 1 ; 1{,}1\rangle$ uznano za przedział antywzorca, gdyż z punktu widzenia ekspozycji na ryzyko osiąganie takich wartości przez ten wskaźnik jest niekorzystne. Przyjmuje się, że poziom wskaźnika poniżej 80% należy uznać za bezpieczny z punktu widzenia ryzyka systemowego. KNF w Raporcie o Sytuacji Banków za I kwartał 2017 r. [Komisja Nadzoru Finansowego, 2017] poziom 96,7% tego wskaźnika dla sektora bankowego ocenia jako zadowalający w zakresie bieżącej płynności i dlatego w badaniu przyjęto przedział $\langle 0{,}7 ; 0{,}8\rangle$ jako przedział wzorca;

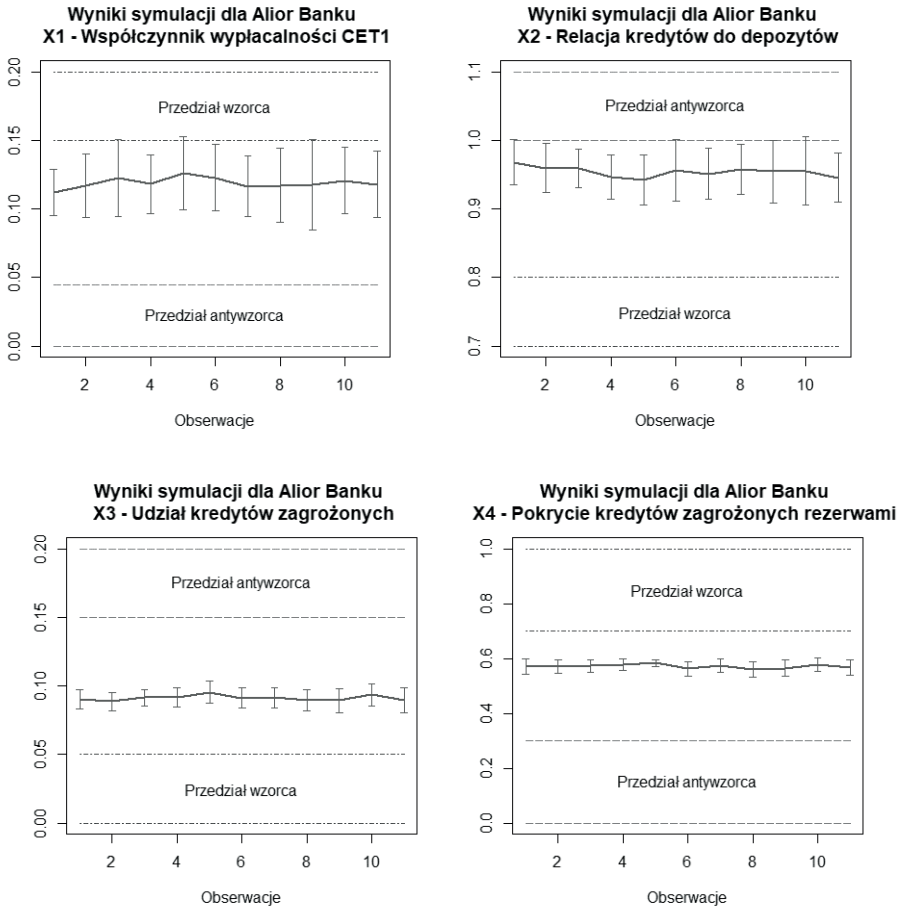
– X_3 – udział kredytów zagrożonych, który jest miernikiem jakości portfela kredytowego i pozwala na ocenę ryzyka znacznej części prowadzonej przez bank działalności. Nie ma sformułowanej teorii, z której wynikałoby, jakie wartości należy uznać za oznaczające wysoki poziom ryzyka systemowego, a jakie oznaczają jego brak. Z części Raportu o Sytuacji Banków za I kwartał 2017 r. [Komisja Nadzoru Finansowego, 2017] dotyczącej jakości portfela kredytowego wynika, że poziom tego wskaźnika w wysokości 6% dla całego sektora bankowego Komisja Nadzoru Finansowego uznaje za sytuację zadowalającą. W związku z tym w badaniu przyjęto przedział $\langle 0 ; 0{,}05\rangle$ za przedział wzorca, natomiast $\langle 0{,}15 ; 0{,}2\rangle$ za przedział antywzorca;
– X_4 – pokrycie kredytów zagrożonych rezerwami, które jest miarą określającą, w jakim stopniu dokonane odpisy pokrywają kredyty zagrożone. Im wskaźnik jest wyższy, tym mniej dotkliwe jest urzeczywistnienie ryzyka kredytowego na sytuację finansową banku. W literaturze nie występuje pojęcie poziomu bezpiecznego dla tego wskaźnika [Rogowska, 2009]. Na potrzeby tego badania przyjęto przedział $\langle 0{,}7 ; 1\rangle$ za wzorcowy, a $\langle 0 ; 0{,}3\rangle$ za antywzorcowy.

2. Funkcją gęstości jest funkcja gęstości 4-wymiarowego rozkładu normalnego wektora losowego X, tzn. $f(X)=\frac{1}{\sqrt{(2\pi)^4|\Sigma|}}e^{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)}$, gdzie μ jest wektorem wartości oczekiwanych, natomiast Σ jest macierzą wariancji i kowariancji.

4.3. Wyniki empiryczne

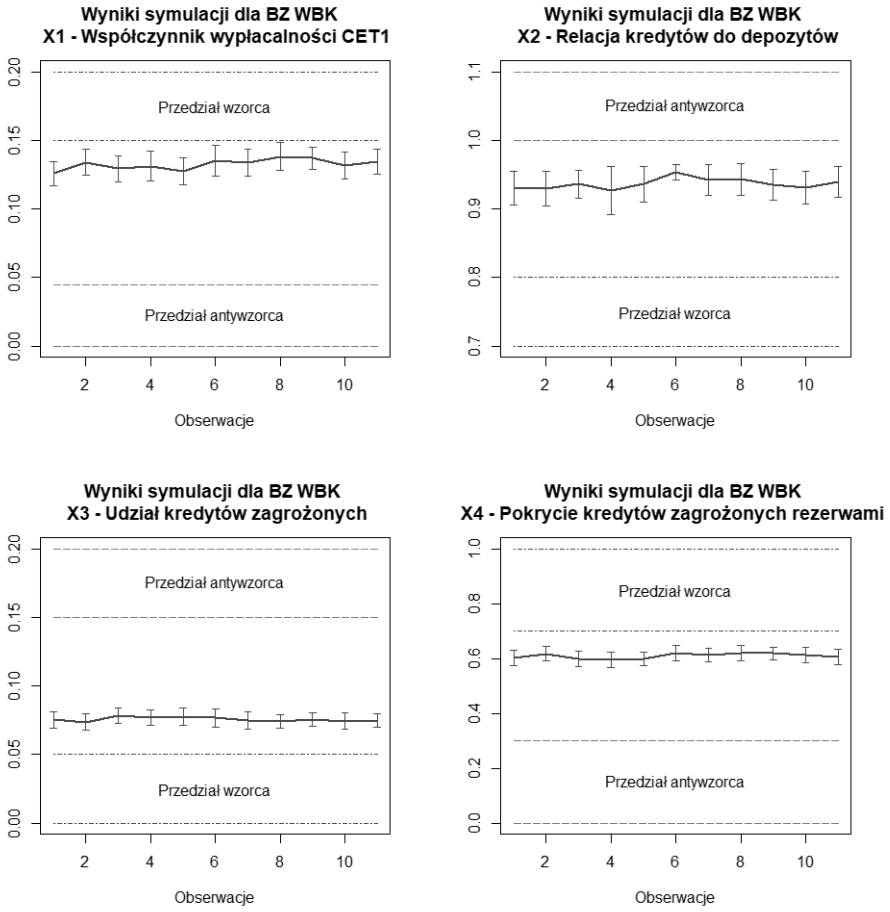
Ze względu na krótkie szeregi czasowe oraz zamiar zbadania różnych scenariuszy przebiegu badanego zjawiska dokonano symulacji. Jak już wspomniano w poprzednich częściach, w przypadku jednoczesnego symulowania większej ilości zmiennych metodą Monte Carlo pojawia się problem ich wzajemnych powiązań, których ta metoda nie uwzględnia. Dlatego otrzymaną metodą Monte Carlo realizację wektora losowego poddano transformacji metodą Imana-Conovera, dzięki której otrzymano zmienne losowe skorelowane ze sobą w ustalony przez badacza sposób. Ze względu na krótką próbę bezzasadna jest weryfikacja typu rozkładu zmiennych. Będzie to możliwe wraz z upływem czasu, kiedy będą napływać nowe obserwacje. Przyjęto założenie, że zmienne mają rozkład normalny. Następnie dokonana została symulacja Monte Carlo zgodnie z otrzymanymi uprzednio rozkładami, a na koniec przeprowadzona została transformacja według metody Imana-Conovera, gdzie za macierz korelacji przyjęto macierz korelacji z próby. Wyniki symulacji przedstawione są na rysunkach 1, 2 i 3, gdzie linie ciągłe reprezentują wysymulowaną wartość zmiennej, końce słupków natomiast wartość powiększoną oraz pomniejszoną o odchylenie standardowe. W ten sposób dla każdej ze zmiennych i dla każdego obiektu uzyskano trzy scenariusze: neutralny, optymistyczny i pesymistyczny. Na rysunkach zaznaczone zostały również przedziały uznane za wzorcowe i antywzorcowe dla kolejnych zmiennych.

Rysunek 1. Wyniki symulacji dla poszczególnych zmiennych kontrolnych wraz z ich poziomami wzorcowymi i antywzorcowymi – Alior Bank



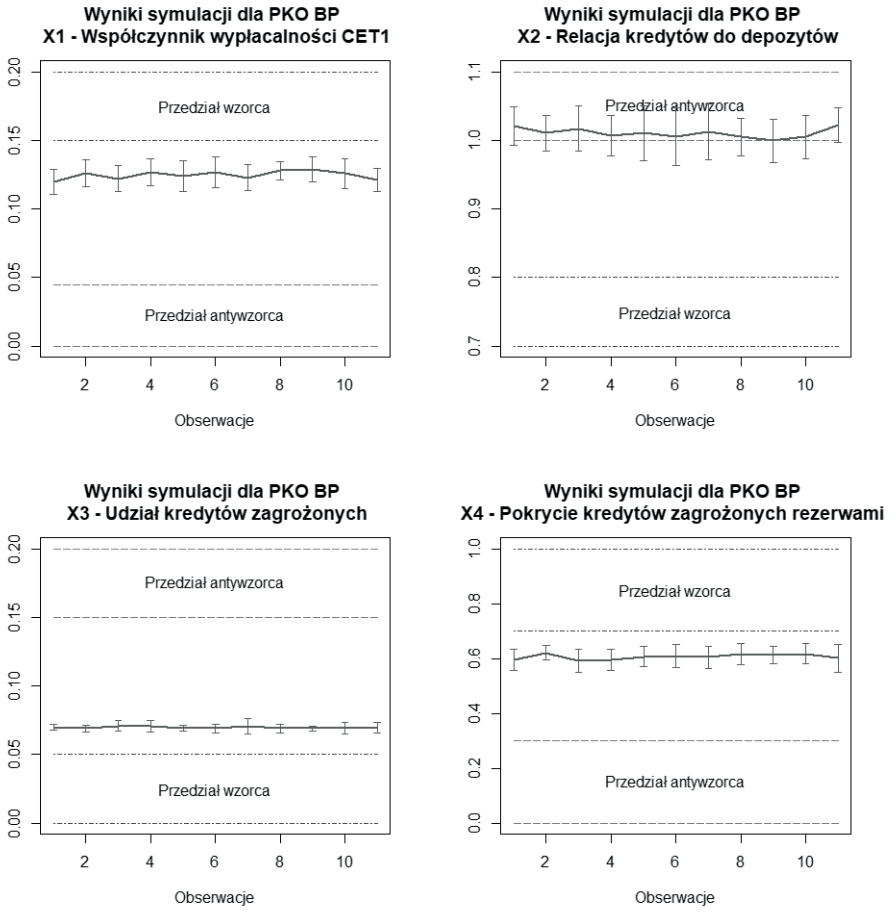
Źródło: Opracowanie własne przy użyciu programu Statistica i środowiska R.

Rysunek 2. Wyniki symulacji dla poszczególnych zmiennych kontrolnych wraz z ich poziomami wzorcowymi i antywzorcowymi – BZ WBK



Źródło: Opracowanie własne przy użyciu programu Statistica i środowiska R.

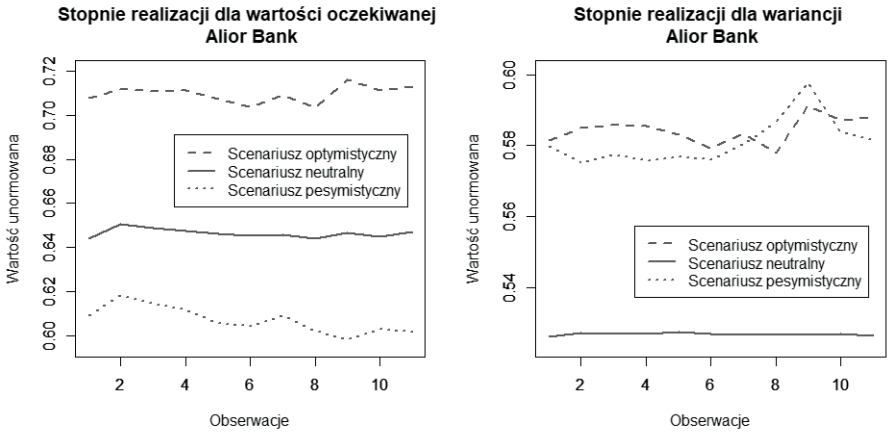
Rysunek 3. Wyniki symulacji dla poszczególnych zmiennych kontrolnych wraz z ich poziomami wzorcowymi i antywzorcowymi – PKO BP



Źródło: Opracowanie własne przy użyciu środowiska programu Statistica i R.

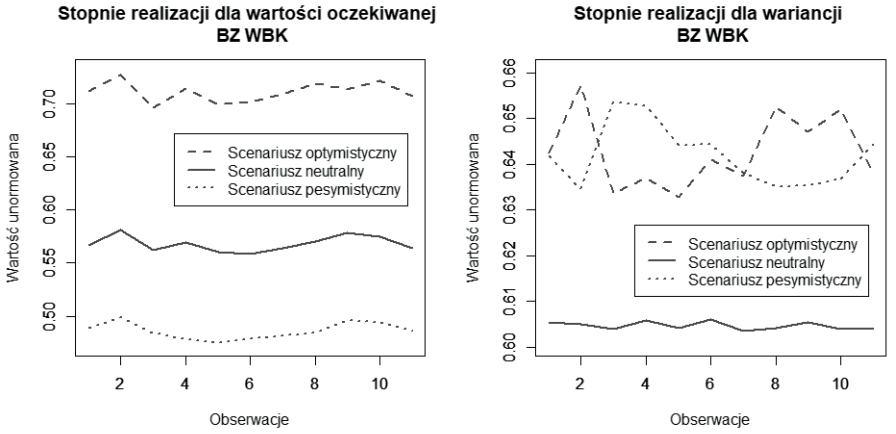
Rysunki 4, 5 i 6 przedstawiają syntetyczny wskaźnik ryzyka, jakim są stopnie realizacji założeń wzorcowych dla wartości oczekiwanej oraz stopnie realizacji założeń wzorcowych dla wariancji dla kolejnych banków.

Rysunek 4. Syntetyczny wskaźnik ryzyka, porównanie scenariuszy – Alior



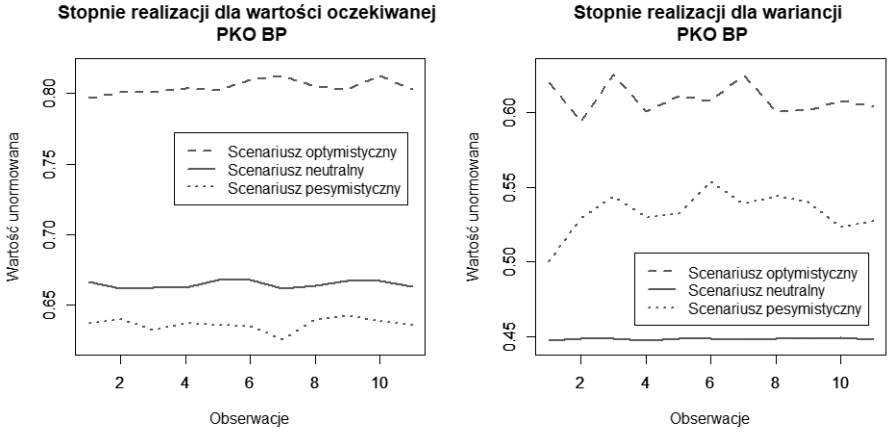
Źródło: Opracowanie własne przy użyciu środowiska R.

Rysunek 5. Syntetyczny wskaźnik ryzyka, porównanie scenariuszy – BZ WBK



Źródło: Opracowanie własne przy użyciu środowiska R.

Rysunek 6. Syntetyczny wskaźnik ryzyka, porównanie scenariuszy – PKO BP



Źródło: Opracowanie własne przy użyciu środowiska R.

5. Wnioski

Analizując otrzymane wartości stopni realizacji dla wartości oczekiwanej przedstawionych na rysunkach 4, 5 i 6, można wnioskować, że w najlepszej sytuacji spośród badanych obiektów jest PKO BP, dla którego w scenariuszu optymistycznym wartość stopni realizacji dla wartości oczekiwanej oscylują wokół wartości 0,81, a w scenariuszu pesymistycznym powyżej 0,62. Alior Bank i BZ WBK również mają dość wysokie wartości analizowanego wskaźnika: Alior od około 0,60 do 0,72, BZ WBK od około 0,48 do 0,73 w zależności od przyjętego scenariusza. Ponieważ omawiany wskaźnik jest unormowany i przyjmuje wartości od 0 (w przypadku sytuacji skrajnie niekorzystnej ze względu na występujące ryzyko) do 1 (w przypadku występowania uwarunkowań wzorcowych), należy uznać powyższe wyniki jako wskazujące na występowanie niewielkiego ryzyka systemowego u badanych obiektów. Należy również zauważyć, że narzędzie jest wrażliwe na zmianę uwarunkowań ryzyka. Dla Alior Banku w scenariuszu optymistycznym wartości wszystkich zmiennych kontrolnych dla obserwacji od 9 do 11 zbliżają się wyraźnie w stronę wzorca ze względu na fakt, że wysymulowana wariancja dla tych obserwacji była relatywnie duża, co znajduje swoje odzwierciedlenie w zauważalnie większych wartościach stopni realizacji dla wartości oczekiwanej. Z analogicznych przyczyn w scenariuszu pesymistycznym dla tych samych obserwacji stopnie realizacji założeń wzorcowych dla wartości oczekiwanej przyjmują wyraźnie mniejsze wartości. Podobnie dla BZ WBK w ścieżce optymistycznej dla obserwacji 2, 4 i 8 kolejne zmienne kontrolne zbliżają się do wzorca, a za nimi rosną zauważalnie wartości stopni realizacji dla wartości oczekiwanej,

natomiast w scenariuszu pesymistycznym dla obserwacji 5, 6 i 7 zmienne kontrolne podążają w kierunku antywzorca (w szczególności zmienne X_1 i X_2), a wartości stopni realizacji dla wartości oczekiwanej wyraźnie osiągają dołek. Dla PKO BP w scenariuszu optymistycznym dla obserwacji 6, 7 i 10 wartości stopni realizacji dla wartości oczekiwanej ulegają wzrostowi, w ślad za zmianą wartości zmiennych kontrolnych, w szczególności zmiennej X_2, której wartości w tym scenariuszu w sposób znaczący przybliżyły się w stronę wzorca. Z kolei w scenariuszu pesymistycznym stopnie realizacji założeń wzorcowych dla wartości oczekiwanej maleją dla 3, 7 i 11 obserwacji, dla których można zaobserwować istotne zmiany wartości zmiennych kontrolnych w stronę ich wartości antywzorcowych (w szczególności zmienne X_2 i X_4).

Analiza stopni realizacji dla wariancji również wskazuje na występowanie niewielkiego ryzyka systemowego u badanych obiektów. W zależności od przyjętego wariantu wskaźnik ten przyjmował: dla Alior Banku wartości od 0,53 do 0,60, dla BZ WBK od 0,60 do 0,66, dla PKO BP od 0,45 do 0,63. Wrażliwość stopni realizacji dla wariancji na zmianę uwarunkowań nie zawsze jest taka sama jak stopni realizacji dla wartości oczekiwanej. Można jednak dostrzec dużo podobieństw, przede wszystkim w ścieżce optymistycznej dla Alior Banku oraz dla BZ WBK, gdzie kształt wykresu obu wskaźników względem kolejnych obserwacji jest niemal identyczny.

Podsumowując analizę wartości unormowanego wektora ryzyka, można wyciągnąć następujące wnioski:

1. Stopnie realizacji założeń wzorcowych dla wartości oczekiwanej okazały się wrażliwe na wahania zmiennych kontrolnych w każdym z wysymulowanych scenariuszy i były one bardziej wrażliwe na zmianę uwarunkowań niż stopnie realizacji założeń wzorcowych dla wariancji.
2. Dla scenariuszy w których zmienne kontrolne przyjmują wartości bliżej wzorca sposoby przebiegu stopni realizacji dla wartości oczekiwanej i stopni realizacji dla wariancji są bardziej zgodne.
3. Koncepcja budowy syntetycznej unormowanej miary ryzyka uwzględniającej korelacje między zmiennymi kontrolnymi umożliwiło sformułowanie problemu mierzenia ryzyka w sposób kompleksowy i upraszcza interpretację uzyskanych wyników.
4. Poziom ryzyka badanych obiektów w wysymulowanych wariantach wydarzeń nie jest wysoki, nawet w scenariuszu pesymistycznym.

Zakończenie

Miara ryzyka została ujęta jako wektor losowy o składowych będących zmiennymi kontrolnymi procesu mierzenia ryzyka systemowego dla

wybranych banków. Wprowadzenie pojęcia wartości wzorca i antywzorca oraz podstawowych charakterystyk statystycznych pozwoliło na zdefiniowanie unormowanych stanów ryzyka w oparciu o odległości wektorów wartości oczekiwanych i wariancji od ich wartości wzorca i antywzorca. Tak skonstruowane miary są łatwe do zrozumienia i interpretacji.

Celem artykułu było przeprowadzenie symulacji mierników ryzyka systemowego zdefiniowanych jako odległość między różnymi stanami wartości wskaźników a ich wartościami wzorcowymi i antywzorcowymi. I tak sformułowany cel zrealizowany został na podstawie metody Monte Carlo z uwzględnieniem korekty Imana-Conovera. Otrzymane wyniki badań potwierdziły stawianą hipotezę badawczą, iż wraz ze wzrostem poziomu ryzyka systemowego obserwowana jest większa wrażliwość wartości oczekiwanych wskaźników niż ich wariancji. Jest to zbieżne z założeniami wprowadzonego pakietu CDR IV/CRR, którego głównym celem jest wprowadzenie takich wskaźników, które zwiększają poziom wrażliwości na zmianę poziomu ekspozycji na ryzyko systemowe.

Kierunkami przyszłych badań jest uzupełnienie proponowanego narzędzia do pomiaru ryzyka systemowego o analizę jego wrażliwości na odstępstwa pojedynczych wskaźników od ścieżki neutralnej, czyli zaplanowanie takich symulacji, których jeden ze wskaźników w danym okresie odchyla się od antywzorca, natomiast pozostałe wskaźniki realizują scenariusz neutralny.

Literatura

Acharya V.V., Pedersen L.H., Philippon T., Richardson M. (2010), *Measuring Systemic Risk*, „Review of Financial Studies", Vol. 30, No. 1.

Basel Committee on Banking Supervision (2011), *Basel III: A global regulatory framework for more resilient banks and banking systems.*

Borio C. (2003), *Towards a macroprudential framework for financial supervision and regulation?*, BIS Working Paper No. 128.

Brämer P., Gischer H., Lücke C. (2014), *A simulation approach to evaluate systemic risk*, „European Journal of Political Economy", Vol. 34.

Buslenko N.P., Golenko D.I., Sobol I.M., Sragowicz W.G., Szrejder J.A. (1967), *Metoda Monte Carlo*, PWN, Warszawa.

Giglio S., Kelly B., Pruitt S. (2016), *Systemic risk and the macroeconomy: An empirical evaluation*, „Journal of Financial Economics", Vol. 119, No. 3.

Haas C. (1999), *On Modeling Correlated Random Variables in Risk Assesment*, „Risk Analysis", Vol. 10, No. 6.

Hellwig M.F. (2009), *Systemic Risk in the Financial Sector: An Analysis of the Subprime-Mortgage Financial Crisis*, „De Economist", Vol. 157, No. 2.

Huang X., Zhou H., Zhu H. (2009), *A Framework for Assessing the Systemic Risk of Major Financial Institutions*, „Journal of Banking & Finance", Vol. 33, No. 11.

Iman R.L., Conover W.J. (1982), *A distribution-free approach to including rank correlation among input variables*, „Communications in Statistics", Vol. 11, No. 3.

Informacja o sytuacji banków w I kwartale 2017 r. (2017), Komisja Nadzoru Finansowego, Warszawa, https://www.knf.gov.pl/knf/pl/komponenty/img/RAPORT_O_SYTUACJI_BANKOW_Ikw_2017_58291.pdf, dostęp: 11.09.2017.

Jastrzębski T. (2015), *Ryzyko obsługi zobowiązań długoterminowych gminy Sopot w świetle perspektywy budżetowej Unii Europejskiej na lata 2007–2013*, „Współczesna Gospodarka", Uniwersytet Gdański, Gdańsk.

Jermakow S.M. (1976), *Metoda Monte Carlo i zagadnienia pokrewne*, PWN, Warszawa.

Kabza M. (2012), *Ryzyko systemowe – cecha współczesnych rynków finansowych*, „Studia Ekonomiczne", nr 3, INE PAN, Warszawa.

Lehar A. (2005), *Measuring systemic risk: A risk management approach*. „Journal of Banking and Finance", Vol. 29, No. 10.

Rogowska E. (2009), *Rezerwy na ryzyko związane z działalnością banków*, „Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania", nr 24, Wydawnictwo Uniwersytetu Szczecińskiego, Szczecin.

Stanowisko KNF w sprawie polityki dywidendowej instytucji finansowych (2014), Komisja Nadzoru Finansowego, Warszawa, https://www.knf.gov.pl/knf/pl/komponenty/img/knf_123947_knf_123947_KNF_polityka_dywidendowa_2_12_2014_39873.pdf, dostęp: 10.09.2017.

Zemke J. (2009), *Ryzyko zarządzania kapitałem organizacji gospodarczej*, „Zeszyty Naukowe", nr 4, t. 4, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.

Streszczenie

Ryzyko systemowe może być rozumiane jako możliwość wystąpienia wielu jednoczesnych nieprawidłowości w funkcjonowaniu dużych instytucji finansowych. Kwestia jego definiowania oraz pomiaru nabrała szczególnego znaczenia w okresie ostatniego kryzysu finansowego lat 2007–2009. Złożoność i zmienność ram teoretycznych i sposobu pomiaru ryzyka systemowego powoduje, że niezwykle trudne jest określenie i kwantyfikowanie czynników ryzyka systemowego w sektorze bankowym . Dodatkowo w swej istocie ryzyko to ma naturę zdarzenia rzadkiego i tym samym nie mają zastosowania miary klasyczne. Uzupełniając to spostrzeżenie o fakt braku danych oraz ich niską częstotliwość, można dojść do wniosku o konieczności zastosowania metod symulacyjnych, których zaletą jest możliwość projektowania różnych scenariuszy wydarzeń.

Celem artykułu jest przeprowadzenie symulacji mierników ryzyka systemowego zdefiniowanych jako odległość między różnymi stanami wartości wskaźników a ich wartościami wzorcowymi i antywzorcowymi. W związku z tak sformułowanym celem postawiono hipotezę badawczą, iż wraz ze wzrostem poziomu ryzyka systemowego obserwowana jest większa wrażliwość wartości oczekiwanych wskaźników niż ich wariancji. Jako narzędzie w procesie weryfikacji hipotezy badawczej w zaprojektowanym eksperymencie użyta została symulacja Monte Carlo z korektą Imana-Conovera.

Słowa kluczowe
ryzyko systemowe, Monte Carlo, korekta Imana-Conovera

Application of Monte Carlo simulations with Iman-Conover correcion in systemic risk measurement (Summary)

Systemic risk can be defined as the possibility of multiple simultaneous irregularities in the functioning of large financial institutions. The issue of its definition and measurement has become particularly important during the last financial crisis of 2007–2009. The complexity and variability of the theoretical framework and the methods of measuring systemic risk made it extremely difficult to identify and quantify systemic risk factors in the banking sector. In addition nature of this kind of risk is a rare occurrence and hence application of classical measures are not effective. In addition data are often of low frequency or there is no data available at all. Therefore one can conclude that simulations should be used to carry out such a research. This approach has the advantage that researcher is able to design different scenarios of events.

The purpose of this article is to simulate systemic risk measures defined as the distance between different quantity of indicator values and their ideal and anti-ideal levels. As a result of this formulation, the research hypothesis has been expressed that with increased systemic risk, the sensitivity of the expected values of indicators to the variance if indicators is higher. As a tool for the verification of the research hypothesis in the designed experiment Monte Carlo simulation with the Iman-Conover correction was used.

Keywords
systemic risk, Monte Carlo, Iman-Conover correction

Adam Marszk*

Structure of the US investment company industry over the period 2000 to 2017: substitution analysis

Introduction

Over the last several years the investment company industry has undergone deep transformation in the majority of the most advanced economies, including the United States. Traditional investment companies, structured as mutual funds, have lost their market share to the companies offering investment solutions within the category of exchange-traded funds (ETFs). Exchange-traded funds may be considered as more innovative in comparison to traditional investment companies (i.e. mutual funds or similar) – their main distinctive features are listing and trading in their units on the stock exchanges as well as unique creation and redemption mechanism. US investment company industry is the largest in the world in terms of the value of key indicator, i.e. managed assets, and may be considered as the reference market, indicating some trends that may in the future take place in other countries. This topic is thus of relevance not only for the researchers who focus on the US financial sector but also for the ones interested in the developments in other regions, e.g. in Europe.

This paper is designed to present the in-depth insights into the changes in the structure of the US investment company industry and to determine whether the traditional investment companies (mutual funds; other similar companies were omitted due to their much lower importance) have lost their market share due to the growth of the innovative companies (ETFs); supplementary aim is the prediction of future changes in the structure of the US industry. Logistic substitution model is applied in order to reach these aims; it is based on the three parameter logistic function, modified to conduct the analysis of the financial sector. Key parameters of the logistic curve are estimated, showing the rate of the substitution. It is the first study to empirically verify the occurrence and intensity of substitution between mutual funds and ETFs in the US investment company industry with the detailed analysis for various groups of funds (classified according to their exposure).

* Ph.D., CFA, Department of Economic Sciences, Faculty of Management and Economics, Gdańsk University of Technology, Gabriela Narutowicza 11/12, 80-233 Gdańsk, amarszk@zie.pg.gda.pl

This paper consists of four sections. First section outlines the key theoretical issues linked and brief literature review. Section two presents the methodology of the research and data sources. Third section is empirical and comprises two parts: preliminary evidence on the US investment company industry, followed by the discussion of the results of the analysis conducted by deploying logistic substitution models. Final section includes the conclusions.

This research was supported by a grant from the CERGE-EI Foundation under a program of the Global Development Network 2016. Title of the research project: "Reshaping financial systems – identifying the role of ICT in diffusion of financial innovations. Recent evidence from European countries". The article is also the result of scientific project no. 2015/19/D/HS4/00399 financed by the National Science Centre of Poland. Supported by the Foundation for Polish Science (FNP).

1. Overview of main categories of investment companies. Literature review

The Investment Company Institute (ICI), the national association of the investment companies domiciled in the United States, distinguishes four groups of investment companies in this country: mutual funds, closed-end funds, exchange-traded funds (ETFs) and unit investment trusts [Investment Company Institute, 2017, p. 9]. Due to their relatively much higher importance (measured e.g. in terms of assets), presented discussion will focus on differences between mutual funds and ETFs which constitute the vast majority of the US investment company industry.

Mutual funds are the dominant category of investment companies in most countries, including the United States (in many, especially less developed economies, they are the only category available). Mutual funds are investment companies in which managers buy a portfolio of assets and manage them according to the stated financial objective. The key attribute of mutual funds is their readiness to buy back previously issued shares at their current net asset value (therefore those shares are labeled 'redeemable securities'). One subcategory of mutual funds are index funds, i.e. funds in which managers are obliged to track the performance of a selected market index (similar to the most popular ETFs).

Exchange-traded funds are the newest category of investment companies – first ETF, tracking the S&P 500 index, was launched on the US market in 1993 [Deville, 2008, p. 4–6]. ETFs are considered to be hybrid investment companies [Deutsche Bank, 2015, p. 100]. ETFs are legally and structurally similar to mutual funds yet their units (shares) are listed and traded on stock exchanges, like shares of closed-end funds [International Monetary

Fund, 2011, p. 68]. However, detailed mechanisms of ETFs differ significantly from those applied by other categories of investment companies.

It should be noted that for many years ETFs were substitutes only for index mutual funds as they were passive investing tools, tracking selected indexes. However, in the last few years the category of ETFs has become much more diversified and includes also semi-passive or active ETFs. Therefore, it is more adequate to analyze the substitution between mutual funds and ETFs by the category of tracked assets (i.e. including all types of mutual funds), and not limit the analysis to index funds.

Table 1. Key attributes of mutual funds and ETFs: comparison

Attribute	Mutual funds	ETFs
Distribution channels	bank offices and similar	mostly brokers
Creation/redemption of units	conducted by the fund or related entities	conducted through transactions between authorized participants and funds
Valuation of units	net asset value determined by the fund	net asset value determined by the fund; market price depends on demand and supply
Costs for investors	depend on, for instance, distribution and management fees	mostly costs of stock transactions – usually lower than in mutual funds

Source: Author's compilation based on [Abner, 2016, p. 282–283; Agapova, 2011, p. 323–330; Deutsche Bank, 2010, p. 20; ICI, 2017, p. 58–63; International Monetary Fund, 2011, p. 68–69; Lechman, Marszk, 2015, p. 355–361].

In contrast with mutual funds ETFs give their users access to innovative features – they were the main cause of their rapidly growing popularity (see table 1). Apart from the different distribution channels (ETFs are available mostly through brokers), the key feature of ETFs is the unique creation and redemption process. In the United States, in almost all cases, shares of ETFs are created and redeemed with the intermediation of authorized participants. In case of creation, authorized participant exchanges a basket of securities (it may also include other assets) for a large number of ETF's units. Redemption occurs in the reverse direction. Both creation and redemption occur on the primary ETF market, whereas secondary ETF market covers the transactions between investors, market participants and other entities (they are conducted through the stock exchanges). Activity of authorized participants ensures that deviations of the market prices of the ETF's units from their net asset values are limited.

One of the main benefits of ETFs *versus* mutual funds is lower costs of ETFs for investors (due to, for example, different distribution channels). Other relative advantages include, among other, higher accessibility (possibility to purchase even one unit in order to gain access to a large number of investment strategies), higher transparency (due to more frequent publication of their portfolio's composition) and possibility to use derivatives (mostly futures) based on their units [Cameron, 2015, p. 7–16; Hill et al., 2015, p. 2–23].

Development of the ETF market may also be considered in the perspective of the substantial change taking place in the investment industry, known as the rise of passive investing (its sustainability remains, though, still to be seen). It may be briefly described as the increased popularity of investment products tracking returns of selected benchmarks (i.e. passive) instead of the ones which attempt to achieve better performance (i.e. active). Theoretical and empirical background for this preference has been described in numerous publications over the last several years (among most prominent [Fama, French, 2010, p. 1915–1947; Samuleson, 1974, p. 17–19; Sharpe, 1991, p. 7–9; Sharpe, 2013, p. 34–41]).

Most publications, which cover the topic of mutual funds and ETFs, focus on the potential benefits for investors, i.e. analyze them using micro approach. The topic of substitution between these investment companies (i.e. with reference to the whole market – macro approach) is described mostly in the theoretical literature (for details see the preceding discussion regarding the comparison of mutual funds *versus* ETFs). Attempts to verify this trend empirically are very scarce, mostly due to rather short history of ETFs and, consequently, lack of data (moreover, for most countries datasets with only annual observations can be constructed; United States is notable exception). According to the results obtained in the previous studies, between 2002 and 2012 ETFs have developed significantly in Mexico while mutual funds lost their market share – the development of the Mexican ETF market can be described with the logistic growth model [Marszk et al., 2017, p. 83–100]; similar results were reached in the study which covered a broader group of countries – Brazil, Mexico, Japan, South Korea and the United States [Lechman, Marszk, 2015, p. 355–376]. In contrast with the current study, annual observations were used for a shorter time period and no analysis of the changes within various types of funds was performed. According to one of the most comprehensive studies on the global ETF industry, the development of ETF markets could be observed in many advanced economies (in North America and Europe but also in Asia-Pacific) and some emerging economies (e.g. Mexico or China) [Hill et al., 2015, p. 160–180]. Apart from the scientific literature,

these results are confirmed in the reports published by Deutsche Bank or ETFGI [Deutsche Bank, 2017; ETFGI, 2017]. Another important topic is the impact of ETFs on the financial system – its evaluation is, though, still difficult due to the short history of the ETF markets yet some attempts have been made and their mixed results were published in several studies (for example [Amenc et al., 2017, p. 59–100; Ben-David et al., 2017, p. 169–189; Madhavan, Sobczyk, 2016, p. 1–17]). Role of ETFs during the 2008 global financial crisis was very limited due to their low assets in comparison to mutual funds at that time (see section 3).

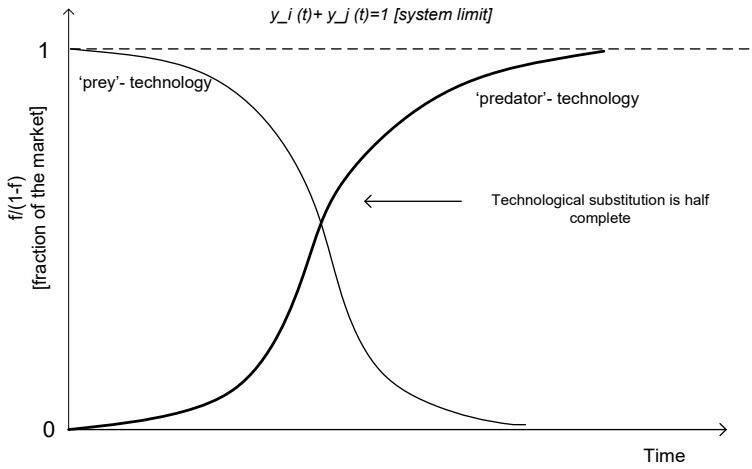
2. Methodological framework

2.1. Research methodology

Analysis starts with descriptive statistics but its main part is conducted by deploying logistic substitution model (LSM). Logistic substitution model (also labeled technological substitution model due to its initial applications) was developed by Fisher and Pry [1972] and since then underwent various modifications. Despite its apparent simplicity, numerous studies have proven its applicability in comparison to other models which have been evaluated as too complex and inefficient in case of technology forecasts (see the discussion of the literature and revised model in Kucharavy and Guio [2011, p. 402–414]). The most significant difficulties and limitations include ambiguities regarding definitions of the key parameters or classifications of competing technologies and lack of formal analysis which would enable the assessment of the obtained fits, for instance with their estimation errors or confidence intervals (analysis of residuals is not possible due to software limitations).

According to the basic assumption of the model, the process of substitution is defined as a gradual replacement of 'old' technologies or products by 'new' ones, which resembles competition between 'old' and 'new' solutions. Another basic element of the model are three characteristic stages of changes in the market share: first, a logistic growth stage, when growth rates are slow; second, an exponential growth stage with a rapid growth of the market share of the 'new' technology or product; third, a saturation stage, when 'new' technology or product reaches its maximum market share. After the end of saturation stage, the 'new' product or technology may be gradually substituted by even more innovative solution [Marchetti, Nakicenovic, 1979, p. 1–8]. For graphical illustration see figure 1.

Figure 1. Theoretical specification of the substitution trajectory according to the logistic substitution model.



Source: [Lechman, 2015].

The logistic substitution model may be applied to analyze the substitution which takes place on financial markets (see for example [Marszk et al., 2017]). When two different products, which replace each other, are considered (in here: units of mutual funds and ETFs; both are labeled as investment products in many publications, e.g., Gastineau [2010] or Hill et al. [2015]), N_i represents the net assets of each product category. Share of product category i at time t can be stated as [Lechman, 2015, p. 47]:

$$f_i(t) = \frac{N_i(t)}{N} \tag{1}$$

The estimate of the time, when the saturation stage of one product category ends and the growth stage of the second category begins, may be determined by checking when [Meyer et al., 1999, p. 247–257]:

$$\frac{y_i''(t)}{y_i'(t)} \rightarrow \min \tag{2}$$

where $y_i(t)$ is the market share of product i according to Fisher-Pry transformation [Fisher, Pry, 1972, p. 75–88]:

$$y_i(t) = \frac{f_i(t)}{1 - f_i(t)} \tag{3}$$

Fisher-Pry transformation is performed in order to normalize the logistic curves, i.e. transform them into straight lines which allows comparisons of various curves and graphical evaluation of deviations of empirical lines from theoretical specifications.

Next step is fitting the curves to the growth or decline stage and estimation. In case of two products (one with growing market share and second with declining), after estimating y_i and y'_i, it is possible to estimate the two key parameters of the logistic curve using following equations [Meyer et al., 1999, p. 247–257]:

$$\Delta t_i = \frac{\ln(81)}{y_i'(t)} \tag{4}$$

and

$$T_{m_i} = \ln\left[\frac{\left(y_i(t) - \frac{\ln(81)}{\Delta t}\right)t}{\frac{\ln(81)}{\Delta t}}\right] \tag{5}$$

Δt_i (labeled 'specific duration') shows the time needed for the product i to increase its share in the combined market from 10% to 90%, and T_{m_i} represents 'mid-point', i.e. point in time when the substitution process is half complete which means that market shares of both products are equal to 50%. For calculations we use IIASA-LSM II statistical software which is the most widely utilized tool for the analysis of substitution (IIASA-LSM II applies non-linear least square method for estimations of the above-mentioned parameters).

2.2. Data sources

To achieve the defined research goal we use the dataset covering the key indicators of the US investment company industry, extracted from the databases and reports delivered by the Investment Company Institute (available commercially on demand). Time scope of the analysis (subject to data availability) is September, 2000 – March, 2017, and monthly observations are used (i.e. the total number of observations in the dataset is 199); projections are presented for periods ending in 2020 and 2022. Key indicator used in the analysis is the value of net assets (defined by ICI as market value of the fund's assets minus its liabilities). Two categories of investment companies considered in this research are mutual funds and ETFs; other types are omitted due to their negligible market share. Consequently, total US investment company industry is understood here as the sum of mutual funds and ETFs, and accordingly market share of mutual funds or ETFs is their share in the combined net assets. Apart from the general analysis of market share of mutual funds and ETFs, their shares in five subcategories, classified by ICI according to the type of managed assets, are also examined: domestic (i.e. US) equity, international equity, bond, hybrid, money market.

3. US investment company industry: structure, substitution, predictions

3.1. Preliminary evidence

US investment company industry has in the recent years changed very significantly, both in terms of structure and size measured in net assets (see table 2). Combined net assets of mutual funds and ETFs have grown between September, 2000 and March, 2017 from ca. 7,3 trillion USD to almost 20 trillion USD (i.e. they have almost tripled); the growth of assets in relation to the GDP has also been considerable yet smaller (by ca. 50% [World Bank, 2017]). The growth of the investment company industry was stable over the whole period and almost unaffected by various events in the US economy. The only significant decrease was observed in the late 2008 and early 2009, during the financial crisis. However, from the mid 2009 assets again grew rapidly.

At the beginning of the analyzed time period ETFs were marginal type of investment company, with less than 1% market share. In March, 2017 their market share increased to ca. 14%, the highest level noted in the history of the US market (see table 2 and figure 2). These results confirm the expansion of ETFs in the US investment company industry, boosted by their innovative features discussed in the previous section. Another reason were changes in the legal and regulatory environment from 2006 onwards, above all permissions granted by the Securities and Exchange Commission, which enabled launch and development of new types of funds – initially, all ETFs listed in the United States were passive funds. Some of the most important regulatory changes involved allowing fund providers to offer funds with modified relation to the benchmark such as inverse or leveraged ETFs (since 2006) as well as active ETFs (first attempts had been made in 2007).

Table 2. Summary statistics on net assets of investment companies in the United States between 2000 and 2017

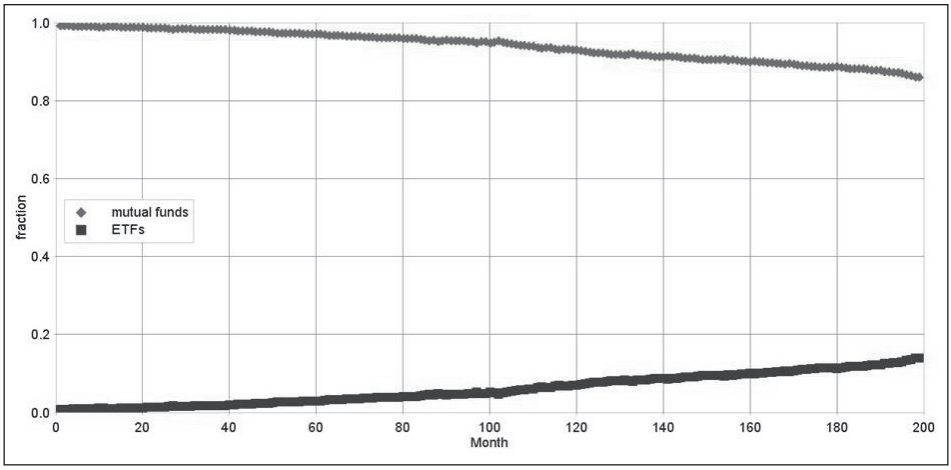
	Exchange-traded funds	Mutual funds	Total market
Value of net assets (in bln USD)			
IX 2000	49,7	7 268	7 318
III 2017	2 774	16 991	19 765
Minimum	49,7	6 057	6 139
Maximum	2 774	16 991	19 765
Total growth (%)	5 482	133,8	170
Market shares (in %)			
IX 2000	0,7	99,3	–
III 2017	14	86	–

	Exchange-traded funds	Mutual funds	Total market
Minimum	0,7	86	–
Maximum	14	99,3	–
Total growth (%)	1 967	–13,4	–
Value of net assets (in bln USD) by asset classes			
domestic equity	1 721	6 760	8 481
international equity	582	2 357	2 939
bond	465	3 776	4 241
hybrid	6	1 433	1 439
money market	–	2 664	2 664

Source: Author's calculations.

As of early 2017 the position of ETFs in the US investment company industry differs in various categories (see table 2). ETFs have reached the highest market share in the 'domestic equity' and 'international equity' category (in both cases share was at ca. 20%). Share of ETFs in the investment companies managing bonds and similar assets was slightly above 10%. ETFs were almost absent in two categories, 'hybrid' and 'money market' – only several funds were launched by the end of the considered time period. The underlying reason are the characteristics and applications of ETFs which are mostly passive funds (whereas hybrid or money-market funds are usually active or semi-active) and the mechanisms of ETFs which in some cases make development of e.g. money-market funds too costly.

Figure 2. Market shares of mutual funds and ETFs in the United States between 2000 and 2017 (shares in total net assets)



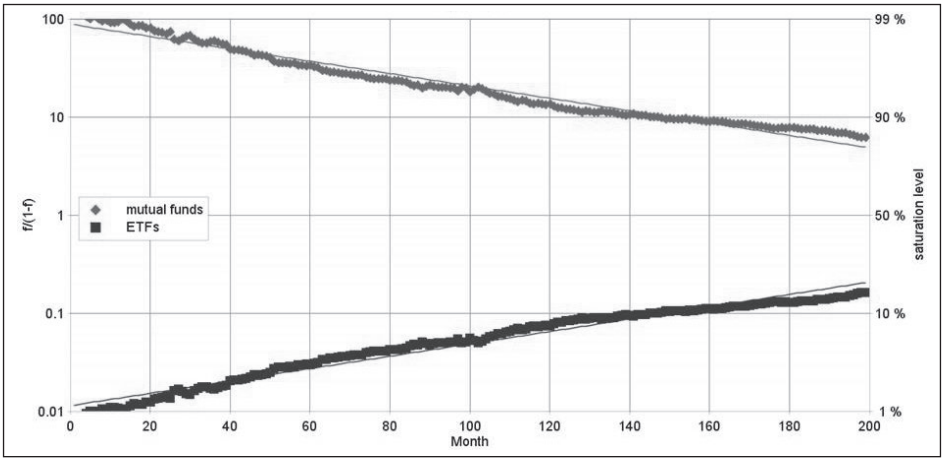
Month 1 = September, 2000. Month 199 = March, 2017.
Source: Author's calculations in IIASA-LSM II software.

3.2. Substitution analysis and predictions

Results of the preliminary analysis outlined in the section 3.1. indicate that over the period 2000–2017 substantial changes took place in the US investment company industry. Mutual funds have lost market shares while ETFs have become increasingly popular. In order to evaluate the exact trends we use the substitution framework discussed in the methodological section. We show the results based on historical market shares in total net assets as well as predictions up to February, 2020 (i.e. three years forward) and, in some cases, up to February, 2022 (i.e. five years forward). Selected lengths of the prediction periods do not seem excessive given the length of the available data set which is approximately 17 years and trajectory of the changes which closely resembles the logistic curve in most analyzed cases, as discussed below. Final paragraph of this section includes the numerical results of the estimations, i.e. values and interpretations of the two key parameters of the logistic substitution model. We assume, as proven by the preliminary analysis, that mutual funds are the 'old' (i.e. losing) products whereas ETFs are 'new' (i.e. winning) products.

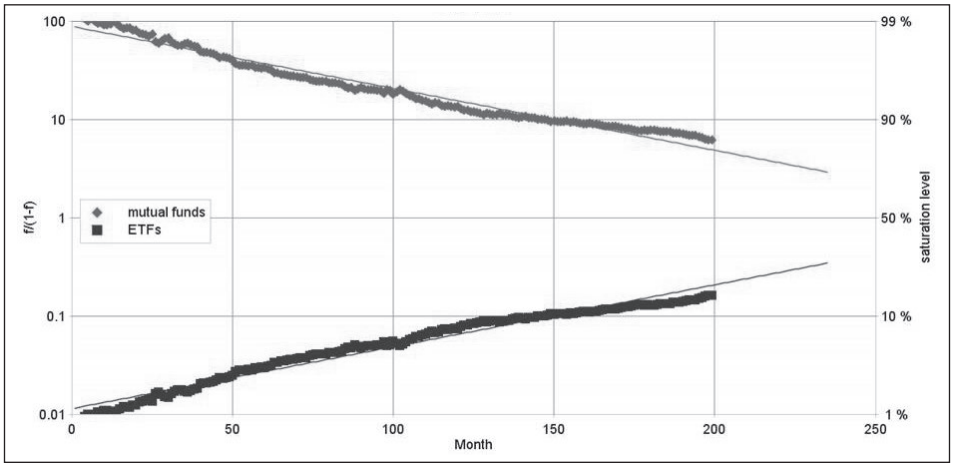
First part of the substitution analysis (see figures 3–5) is conducted for the whole investment company industry, i.e. all types of mutual funds and ETFs. Figure 3 shows that empirical and theoretical curves are close to each other, which proves that changes in market shares follow the logistic curve. It means, therefore, that the logistic growth model may be applied to analyze the substitution in the investment company industry. Mutual funds (as 'old' products) have been substituted by ETFs ('new' products). According to the predictions (see figures 4 and 5) this trend is expected to continue in the near future. Market share of ETFs is expected to grow quickly, reaching the level of ca. 30% by 2020 and 35% by 2022. However, mutual funds will still be the main group of investment companies over the next few years. They are expected to lose the dominant position around 2026. It should be stressed that in the assumed scenarios, changes in the structure of the investment company industry will follow the logistic substitution trajectory.

Figure 3. Market shares of mutual funds and ETFs in the United States between 2000 and 2017 (shares in total net assets). Fisher-Pry transformation



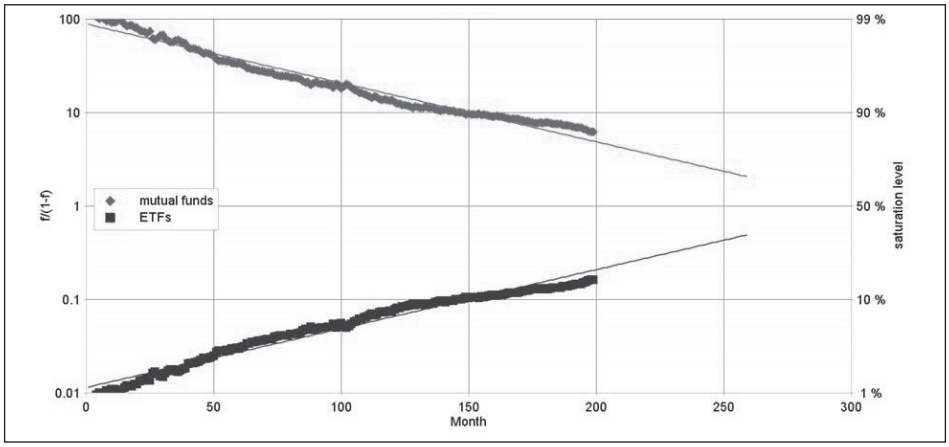
Month 1 = September, 2000. Month 199 = March, 2017.
Source: Author's estimations in IIASA-LSM II software.

Figure 4. Historical and predicted market shares of mutual funds and ETFs in the United States between 2000 and 2020 (shares in total net assets). Fisher-Pry transformation



Month 1 = September, 2000. Month 235 = February, 2020.
Source: Author's estimations in IIASA-LSM II software.

Figure 5. Historical and predicted market shares of mutual funds and ETFs in the United States between 2000 and 2022 (shares in total net assets). Fisher-Pry transformation



Month 1 = September, 2000. Month 259 = February, 2022.
Source: Author's estimations in IIASA-LSM II software.

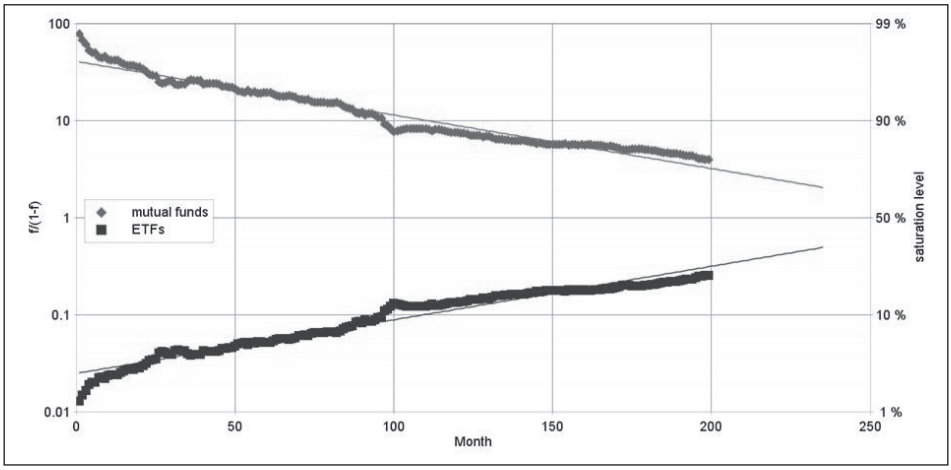
Results of the analysis for the largest category in the investment company industry, domestic equity, show that mutual funds have been to some extent substituted by ETFs (see figure 6); the trajectory of changes and their rate were similar to the ones observed in case of the whole market, discussed in the preceding paragraph. It may be noted, though, that the predictions for the period 2017–2020 indicate that by 2020 the market share of ETFs in this category will be a bit higher than in the whole market – the projected market share of ETFs in the domestic equity category is at ca. 35%. Relatively strong position of ETFs in this category can be explained by the longest history of such ETFs and their popularity – domestic equity ETFs constitute ca. 60% of the entire ETF market. However, reaching higher market shares may be difficult due to the strong competition from a large number of mutual funds. Slowdown in the expected growth of ETFs' market share is also proven by relatively highest value of the estimated specific duration (see table 3).

Table 3. Logistic substitution parameters

Parameter	Total market	Domestic equity	International equity	Bond
T_{mi} – midpoint	V 2026	XI 2024	III 2020	VI 2023
Δt_i – specific duration	302,9	345,8	228,9	201,5

Estimated values of midpoints were transformed from numbers of months into exact points in time (month and year).
Source: Author's estimations in IIASA-LSM II software.

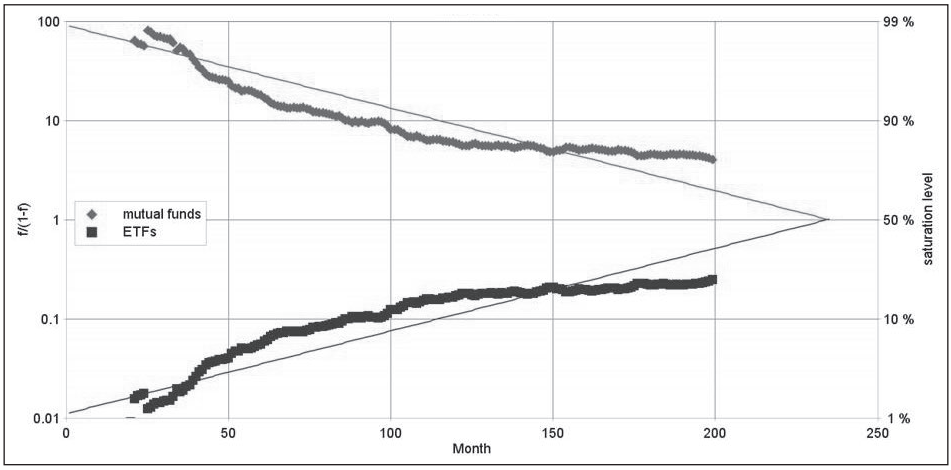
Figure 6. Historical and predicted market shares of mutual funds and ETFs in the United States between 2000 and 2020 (shares in total net assets). Fisher-Pry transformation. Domestic equity category



Month 1 = September, 2000. Month 235 = February, 2020.
Source: Author's estimations in IIASA-LSM II software.

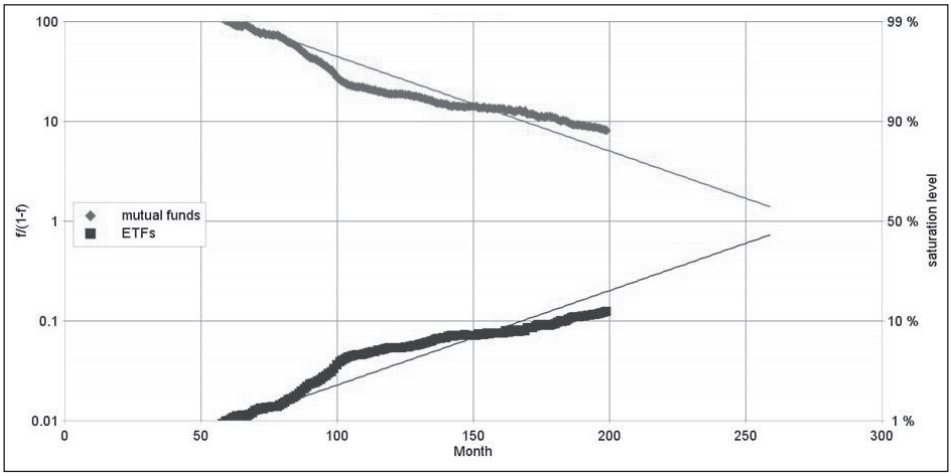
Substitution between mutual funds and ETFs in two next categories of investment companies – international equity and bond funds (see figures 7 and 8) has been more prevalent than for the entire market or domestic equity funds. Particularly significant trend could be observed for international equity. Market share of ETFs in this group in March, 2017 was similar to the market share of domestic equity ETFs in their category but the predictions show that by 2020 international equity are predicted to have the same share as 'old' products, i.e. mutual funds. Rapid development of international equity ETFs was caused by their diversity (ETFs offering access to almost all international equity markets have been launched) and ability to invest in assets located in countries otherwise difficult to access (e.g. for many years mainland China) at lower cost than in mutual funds [Hill et al., 2015, p.109–111]. It should be noted, though, that in case of the international equity category the empirical and theoretical curves to some degree deviate from each other, which proves that changes in market shares do not follow closely the logistic curve. It means that these results should be interpreted with caution – figure 7 shows that in the last few years of historical period the growth of market share of ETFs has slowed down (similar deviations have been observed for the bond category – see figure 8).

Figure 7. Historical and predicted market shares of mutual funds and ETFs in the United States between 2000 and 2020 (shares in total net assets). Fisher-Pry transformation. International equity category



Month 1 = September, 2000. Month 235 = February, 2020.
Source: Author's estimations in IIASA-LSM II software.

Figure 8. Historical and predicted market shares of mutual funds and ETFs in the United States between 2000 and 2022 (shares in total net assets). Fisher-Pry transformation. Bond category



Month 1 = September, 2000. Month 259 = February, 2022.
Source: Author's estimations in IIASA-LSM II software.

Market share of bond ETFs in March, 2017 was much lower than of equity ETFs but it has been growing very quickly since 2007 (first such funds were launched in 2002 but initially both their net assets were very low). When we compare the changes in market share of ETFs in various categories,

we can see that bond ETFs have reached market share of 10% much faster than any other group of ETFs. Moreover, according to the substitution prediction they are expected to reach market share of slightly above 40% by 2020. This high dynamics of bond ETFs in the last years may be explained by referring to the general rules of diffusion of innovations. Innovative products initially gain market share slowly but after certain time, when market participants become more aware, their diffusion accelerates. Bond ETFs have to some extent skipped the first stage of diffusion as equity ETFs were already popular among investors when the actual development of bond ETFs began.

We do not conduct the substitution analysis for the hybrid and money-market funds due to the very low market share of ETFs. Mutual funds are still the dominant investment companies and no substitution has been observed.

More accurate analysis of the substitution, which supports the results presented in the preceding paragraphs, may be performed using the logistic substitution parameters (see table 3; for methodological discussion and limitations such as impossibility to estimate estimation errors see section 2.1). There is no formal procedure to verify their reliability but they may evaluated with regard to the time period and number of observations in the dataset (in here: 199) – values of parameters significantly exceeding the moment in time or length of the sample may be perceived as highly unrealistic.

According to estimated midpoints (points in time when market shares of mutual funds and ETFs are expected to be equal, i.e. of 50%), international equity ETFs are expected to reach 50% market share most quickly, in March, 2020, whereas for domestic equity ETFs the value of this parameter is much higher, meaning much more distant point in time – November, 2024. For bond ETFs estimated mid-point is also rather distant – June, 2023. Nevertheless, in case of bond ETFs the rate of growth in the market share of ETFs is most rapid as indicated by the lowest specific duration (estimated time needed for market share to grow from 10% to 90%), at ca. 202 months. Even though international equity ETFs are much closer to reaching 50% market share than bond ETFs, their growth is a bit slower as suggested by higher specific duration. The highest specific duration for domestic equity ETFs suggests that in this category ETFs are very far from dominating and saturating the market – the value of ca. 346 months is extremely high and shows that their further growth is uncertain.

For the total market estimated midpoint is more distant than in case of any discussed category which is caused by inclusion of hybrid and money market funds (in these categories ETFs are almost absent).

These results show that ETFs are projected to have strong position in some parts of the market but in the upcoming years mutual funds will still dominate the industry at large. This may be changed by the rapid development of ETFs in the hybrid and money market categories but as of mid 2017 there are no signs of such trend.

Conclusions

During the first several years of the 21st century, US investment company industry has changed very deeply. Despite the considerable increase in the net assets of mutual funds, the growth of the competing investment companies – ETFs has been even more significant. In 2000 ETFs were an almost unknown group of investment companies, with marginal market share. Seventeen years later, in 2017, they are contestant for mutual funds, at least in some parts of the market.

When the structure of the investment company industry is considered, results of the analysis show that ETFs linked with the equity markets have gained more popularity and won higher market share than ETFs based on bonds – the explanation is rather simple: equity ETFs have been present much longer than their bond counterparts. International equity ETFs have expanded most quickly and are expected to win substantial market share (the highest among all categories considered) – they have been used by US investors to gain cheap access to foreign markets. However, the observed and predicted growth in the market share of bond ETFs is even more rapid than of equity ETFs due to their later launch and adoption of large group of users.

References

Abner D. (2016), *The ETF Handbook. How to Value and Trade Exchange-Traded Funds. Second Edition*, John Wiley & Sons, Hoboken, New Jersey.

Agapova A. (2011), *Conventional mutual index funds versus exchange-traded funds*, "Journal of Financial Markets", Vol. 14, No. 2.

Amenc N., Goltz F., Le Sourd V. (2017), *The EDHEC European ETF and Smart Beta Survey*, EDHEC-Risk Institute.

Ben-David I., Franzoni F.A., Moussawi R. (2017), *Exchange Traded Funds*, "Annual Review of Financial Economics", Vol. 9.

Cameron R. (2015), *ETFs Exchange Traded Funds: Everything to know about trading exchanges traded funds*, Amazon, UK.

Deville L. (2008), *Exchange Traded Funds: History, Trading and Research*, in: Doumpos M., Pardalos P., Zopounidis C. (eds.), *Handbook of Financial Engineering*, Springer US.

Deutsche Bank (2010), *The Race for Assets in the European Commodity Exchange-Traded Products Space*, London.

Deutsche Bank (2015), *European Monthly ETF Market Review,* London.

Deutsche Bank (2017), *ETF Annual Review & Outlook*, London.
ETFGI (2017), *ETFGI Global ETF and ETP Industry highlights – June 2017*, http://www.etfgi.com/publications/reports, accessed: 20.07.2017.
Fama, E.F., French K.R. (2010), *Luck versus Skill in the Cross-Section of Mutual Fund Returns*, "The Journal of Finance", Vol. 65, No. 5.
Fisher J.C., Pry R.H. (1972), *A simple substitution model of technological change*, "Technological Forecasting and Social Change", Vol. 3.
Gastineau G. (2010), *The Exchange-Traded Funds Manual*, John Wiley & Sons, Hoboken, New Jersey.
Hill J.M., Nadig D., Hougan M., Fuhr D. (2015), *A comprehensive guide to exchange-traded funds (ETFs)*, CFA Institute Research Foundation, Charlottesville.
International Monetary Fund (2011), *Global Financial Stability Report: Durable Financial Stability. Getting There from Here*, Washington, DC.
Investment Company Institute (2009), *Investment Company Fact Book 2009*, Washington, DC.
Investment Company Institute (2017), *Investment Company Fact Book 2017*, Washington, DC.
Kucharavy D., De Guio R. (2011), *Logistic substitution model and technological forecasting*, "Procedia Engineering", Vol. 9.
Lechman E. (2015), *ICT Diffusion in Developing Countries: Towards a New Concept of Technological Takeoff*, Springer.
Lechman E., Marszk A. (2015), *ICT technologies and financial innovations: the case of Exchange Traded Funds in Brazil, Japan, Mexico, South Korea and the United States*, "Technological Forecasting and Social Change", Vol. 99, Issue C.
Madhavan A.N., Sobczyk A. (2016), *Price Dynamics and Liquidity of Exchange-Traded Funds*, "Journal of Investment Management", Vol. 14, No. 2.
Marchetti C., Nakicenovic N. (1979), *The dynamics of energy systems and the logistic substitution model*, International Institute for Applied Systems Analysis, Laxenburg.
Marszk A., Lechman E., Kaur H. (2017), *Financial markets diffusion patterns. The case of Mexican investment funds*, "Equilibrium", Vol. 12, No. 1.
Meyer P.S., Yung J.W., Ausubel J.H. (1999), *A primer on logistic growth and substitution: the mathematics of the Loglet Lab software*, "Technological Forecasting and Social Change", Vol. 61, No. 3.
Samuelson P.A. (1974), *Challenge to judgment*, "The Journal of Portfolio Management", Vol. 1, No. 1.
Sharpe W. (1991), *The Arithmetic of Active Management*, "Financial Analysts Journal", Vol. 47, No. 1.
Sharpe W. (2013), *The Arithmetic of Investment Expenses*, "Financial Analysts Journal", Vol. 69, No. 2.
World Bank (2017), *Global Financial Development Database*, Washington, DC.

Structure of the US investment company industry over the period 2000 to 2017: substitution analysis (Summary)

Over the last years the investment company industry has undergone deep transformation in the majority of the most advanced economies, including the United States. Traditional investment companies, such as mutual funds, have lost their market share to innovative exchange-traded funds (ETFs). The main distinctive features of ETFs are listing and trading in their units on the stock exchanges as well as lower costs for investors. The aim of this paper is to present the in-depth insights into the changes in the structure of the US investment company industry and to determine whether the mutual funds have lost their market share due to the growth of ETFs; supplementary aim is the prediction of future changes. In the major In the major part of the empirical research logistic substitution model is applied, modified to conduct the analysis of changes on the financial markets. Key parameters of the logistic curve are estimated, showing the rate of the substitution. It is the first study to empirically verify the occurrence and intensity of substitution between mutual funds and ETFs in the US investment company industry with the detailed analysis for various groups of funds (classified according to their exposure). Results of the analysis prove that between 2000 and 2017 ETFs have increased their market share and this trend is expected to continue in the future. ETFs linked with the equity markets have won higher market share than ETFs based on bonds. International equity ETFs are projected to win the highest market share among all categories considered.

Keywords

exchange-traded funds, mutual funds, financial innovations, United States, substitution

Barbara Batóg*
Jacek Batóg**

Analiza czynników wpływających na zmiany eksportu krajów Unii Europejskiej

Wstęp

Jednym z istotnych aspektów rozwoju gospodarki jest jej zdolność do intensyfikacji handlu zagranicznego, a zwłaszcza eksportu. Potwierdza to stwierdzenie J. Siedleckiego: „Z obserwacji gospodarki światowej wynika, że żaden kraj nie może osiągnąć sukcesów gospodarczych bez weryfikacji sprawności gospodarczej na międzynarodowych, wielkich, światowych rynkach, tzn. w warunkach ostrej konkurencji technologicznej, organizacyjnej i promocyjnej. Weryfikację taką uzyskuje się za pośrednictwem eksportu" [Siedlecki, 2000, s. 87]. Teoria ekonomii wskazuje wiele czynników, które decydują o wartości i dynamice eksportu danego kraju. Wymienia się wśród nich między innymi stan koniunktury gospodarczej, przynależność do określonego związku państw lub porozumienia handlowego, poziom konkurencyjności eksportowanych dóbr i usług, będący funkcją struktury sektorowej i zaawansowania technologicznego gospodarki, czy lokalizację kraju eksportera. „Tylko ten kraj osiąga dobre rezultaty w eksporcie, który sąsiaduje z dużym, pojemnym rynkiem zbytu" [tamże, s. 89]. Powyższe determinanty eksportu wywodzą się z neoklasycznej teorii ekonomii i klasycznej teorii handlu międzynarodowego [Wernerheim, 2000, s. 112].

Przeprowadzone w artykule badanie ma na celu analizę czynników wpływających na zmiany wartości eksportu towarów w krajach Unii Europejskiej w latach 2002–2015. Zastosowana metoda badawcza w postaci zmodyfikowanej analizy *shift-share* pozwoli na identyfikację, czy czynnikiem sukcesu w handlu zagranicznym krajów członkowskich UE jest ich struktura eksportu czy raczej konkurencyjność eksportowanych dóbr, która utożsamiana jest z wyższą dynamiką eksportu w poszczególnych jego sekcjach[1]. Weryfikacji poddane zostaną dwie hipotezy badawcze. Pierwsza z nich zakłada, że główną przyczyną korzystnego kształtowania

* Dr, Instytut Ekonometrii i Statystyki, Wydział Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński, ul. Mickiewicza 64, 70-376 Szczecin, barbara.batog@usz.edu.pl
** Dr hab., Instytut Ekonometrii i Statystyki, Wydział Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński, ul. Mickiewicza 64, 70-376 Szczecin, jacek.batog@usz.edu.pl
[1] W badaniu nie odnoszono się do eksportu usług.

się eksportu nowych krajów członkowskich UE w latach 2002–2015 była poprawa koniunktury gospodarczej, powodująca wyższą dynamikę eksportu. Według drugiej hipotezy fakt akcesji do Unii Europejskiej pozytywnie wpłynął na wzrost eksportu nowych krajów członkowskich.

1. Przegląd dotychczasowych badań

Analizie prawidłowości charakteryzujących eksport poświęcono w literaturze bardzo wiele uwagi. Świadczą o tym liczne prace z tego zakresu, których wybrane przykłady zostały zamieszczone w tabeli 1 wraz z głównymi wnioskami w nich sformułowanymi.

Tablica 1. Wybrane publikacje odnoszące się do tematyki eksportu

Autor i rok wydania	Wnioski
Wernerheim [2000]	Nie obserwujemy istotnego związku między eksportem i wzrostem gospodarczym, a decydującą rolę dla wzrostu gospodarczego i eksportu danego kraju odgrywa poziom koniunktury gospodarczej w krajach partnerskich (stosowane metody: testy stacjonarności, kointegracji i przyczynowości w sensie Grangera)
Balaguer, Cantavella-Jordá [2004]	Zmiana struktury eksportu na bardziej „nowoczesną" pozytywnie wpływa na wzrost gospodarczy (stosowane metody: testy stacjonarności, kointegracji i przyczynowości w sensie Grangera)
Cieślik i inni [2014a]	Przystąpienie nowych krajów członkowskich do UE, a zwłaszcza strefy euro, zwiększa skłonność przedsiębiorstw do eksportu (stosowana metoda: model probitowy)
Cieślik i inni [2014b]	Skłonność przedsiębiorstw do eksportu w krajach EŚ-W wykazuje silny związek z poziomem: wydajności, kapitału ludzkiego, innowacyjności, nakładów B+R oraz wiekiem i wielkością firmy, pochodzeniem kapitału i stosowanymi technologiami (stosowana metoda: model probitowy)
Dudziński, Narękiewicz [2014]	Struktura eksportu, wraz z długookresowymi relacjami cenowymi w handlu zagranicznym, determinuje jego dynamikę
Lechman [2014]	Brak wpływu wzrostu udziału dóbr zaawansowanych technologicznie, w tym produktów ICT w eksporcie, na ogólną konkurencyjność gospodarek (Global Competitiveness Index) Europy Środkowo-Wschodniej
Michalski [2014a]	Istnieje silne uzależnienie polskiego eksportu od aktywności zagranicznych inwestorów i importowanej technologii

Autor i rok wydania	Wnioski
Michalski [2014b]	Brak wpływu wzrostu polskiego eksportu na poprawę strukturalnej konkurencyjności gospodarki – jej technologicznego zaawansowania
Olczyk, Kordalska [2015]	Występuje silne sektorowe zróżnicowanie czynników determinujących dynamikę eksportu (stosowane metody: testy stacjonarności i kointegracji oraz model SUR)
Dawson, Hubbard [2004]; Krawczyk i inni [2016]	Istnieje istotna dodatnia zależność między dynamiką eksportu i wzrostem gospodarczym (stosowane metody: funkcje produkcji dla danych panelowych, model VAR)
Głodowska [2016]	W krajach UE nie obserwujemy konwergencji stopnia specjalizacji sektorowej eksportu w obszarze towarów wysoko przetworzonych (stosowane metody: klasyczne miary konwergencji typu sigma i beta)
Vu i inni [2016]	Przedsiębiorstwa o wyższym poziomie wydajności wykazują wyższą skłonność do eksportu (stosowane metody: funkcja produkcji typu translog dla danych panelowych)

Źródło: Opracowanie własne.

Cel niniejszej pracy nawiązuje do dotychczasowych badań w dwóch aspektach. Pierwszym z nich jest ocena, czy za główny czynnik determinujący wzrost eksportu można uznać ogólny poziom koniunktury gospodarczej? Natomiast drugi odnosi się do wpływu sektorowej struktury gospodarki na poziom i dynamikę eksportu. Przy czym należy podkreślić, że stosowana metoda badawcza pozwala udzielić odpowiedzi na oba powyższe problemy badawcze, podczas gdy w większości prac są one traktowane osobno.

W publikacjach nawiązujących do zagadnienia eksportu znajdziemy również takie, w których wykorzystana została metoda *shift-share*. Na przykład R. Green i A. Allaway zastosowali tę metodę do identyfikacji rynków i produktów charakteryzujących się wysokim potencjałem wzrostu [Green, Allaway, 1985][2]. F. Herschede wykorzystał wspomnianą metodę w analizie stopnia rywalizacji eksportowej między krajami azjatyckimi, a uzyskany wynik potwierdził spadek eksportu krajów ASEAN do Japonii spowodowany ekspansywną polityką eksportową Chin [Herschede, 1991]. W innej pracy metoda *shift-share* pozwoliła zauważyć wpływ specyficznych

[2] Autorzy ci zwrócili również uwagę na pewne ograniczenia metody *shift-share*.

czynników regionalnych na dynamikę eksportu poszczególnych terytoriów Australii [Neri, Jayanthakumaran, 2005, s. 164–174][3].

2. Stosowana metoda badawcza

Metoda *shift-share* pozwala ustalić, jaki wpływ na zmiany eksportu krajów UE (tzw. efekty krajowe *REGj*) miał efekt strukturalny *SEj* (*sectoral-mix effect*), który mierzy wpływ różnicy między początkową strukturą rodzajową eksportu kraju *j* i całej UE oraz efekt konkurencyjny *CEj* (*competitive effect*), który pozwala określić oddziaływanie różnicy między stopą wzrostu poszczególnych sekcji eksportu dla kraju *j* i stopą wzrostu charakteryzującą te sekcje eksportu w całej UE. W przeprowadzonym badaniu wykorzystana została modyfikacja powyższej metody zaproponowana w pracy [Artige, Neuss, 2013][4]. Procedurę obliczania wszystkich trzech efektów przedstawiają formuły (1)–(4).

$$REG_j = CE_j + SE_j \tag{1}$$

REG_j – efekt krajowy (względna stopa wzrostu eksportu) w *j*-ym kraju:

$$REG_j = g^j_{t+1} - r_{t+1} \tag{2}$$

gdzie:
$j = 1, 2, \ldots, N$,
N – liczba krajów,
g^j_{t+1} – stopa wzrostu eksportu w *j*-ym kraju w stosunku do roku poprzedniego,
r_{t+1} – stopa wzrostu eksportu w Unii Europejskiej w stosunku do roku poprzedniego,
CEj – efekt konkurencyjny w *j*-ym kraju:

$$CE_j = \frac{1}{I}\sum_{i=1}^{I}\left(g^j_{i,t+1} - r_{i,t+1}\right) \tag{3}$$

gdzie:
$g^j_{i,t+1}$ – stopa wzrostu *i*-tej sekcji eksportu w *j*-ym kraju w stosunku do roku poprzedniego,
$r_{i,t+1}$ – stopa wzrostu *i*-tej sekcji eksportu w Unii Europejskiej w stosunku do roku poprzedniego,
SE_j – efekt strukturalny w *j*-ym kraju:

[3] Warto jednocześnie zauważyć, że metoda *shift-share* miała również wiele innych zastosowań, w tym w analizie zanieczyszczenia powietrza [Antczak, 2014], procesów migracyjnych [Antczak, Lewandowska-Gwarda, 2014] czy polityce zdrowia [Jewczak, Żółtaszek, 2011].
[4] Modyfikacja ta pozwala wyeliminować „nakładanie" się efektów strukturalnego i konkurencyjnego występujące w jej wariancie podstawowym.

$$SE_j = \sum_{i=1}^{I}\left(\omega_{i,t}^{j} - \frac{1}{I}\right)g_{i,t+1}^{j} - \sum_{i=1}^{I}\left(\theta_{i,t} - \frac{1}{I}\right)r_{i,t+1} \quad (4)$$

gdzie:
$\omega_{i,t}^{j}$ – udział *i*-tej sekcji eksportu w *j*-ym kraju w eksporcie w *j*-ym kraju w roku bazowym,
$\theta_{i,t}$ – udział *i*-tej sekcji eksportu w Unii Europejskiej w eksporcie Unii Europejskiej w roku bazowym.

Należy zauważyć, że metoda *shift-share* była już wykorzystywana przez autorów prac poświęconych badaniu działalności eksportowej. Można tu między innymi wymienić następujące publikacje: [Green, Allaway, 1985; Herschede, 1991; Coughlin, Pollard, 2001][5].

3. Wyniki empiryczne

Dane statystyczne wykorzystane w obliczeniach zostały pozyskane z bazy danych EUROSTATU [ec.europa.eu/eurostat/data/database] i obejmowały informacje o strukturze eksportu 25 krajów UE (bez Cypru, Luksemburga i Malty ze względu na występujące braki danych) w latach 2002–2015 według 6 sekcji, uzyskanych w oparciu o SITC Rev. 4 (bez sekcji 9 tej klasyfikacji dotyczącej towarów i transakcji niesklasyfikowanych):

SEKCJA 1 = Sekcja 0 Żywność i zwierzęta żywe + **Sekcja 1** Napoje i tytoń

SEKCJA 2 = Sekcja 2 Surowce niejadalne z wyjątkiem paliw + **Sekcja 4** Oleje, tłuszcze, woski zwierzęce i roślinne

SEKCJA 3 = Sekcja 3 Paliwa mineralne, smary i materiały pochodne

SEKCJA 4 = Sekcja 5 Chemikalia i produkty pokrewne

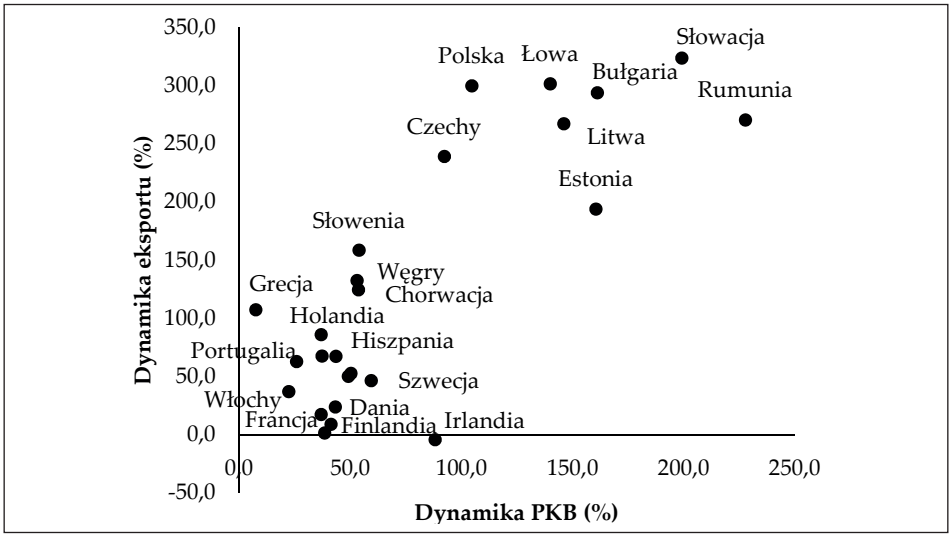
SEKCJA 5 = Sekcja 7 Maszyny, urządzenia i sprzęt transportowy

SEKCJA 6 = Sekcja 6 Towary przemysłowe sklasyfikowane gł. wg surowca + **Sekcja 8** Różne wyroby przemysłowe.

W latach 2002–2015 przeciętny wzrost eksportu w 14 analizowanych starych krajach członkowskich wyniósł 46,0%, natomiast w grupie 11 nowych członków UE kształtował się na poziomie 238,5%. Ta dysproporcja była przede wszystkim efektem dostępu krajów Europy Środkowo-Wschodniej do nowych rynków zbytu oraz transformacji strukturalno-technologicznej ich gospodarek, która pozwoliła zwiększyć konkurencyjność ich produktów. W badanym okresie najwyższą dynamiką PKB charakteryzowała się Rumunia (227,7%), a najniższą Grecja (7,5%). W przypadku dynamiki eksportu były to odpowiednio Słowacja (323,5%) i Irlandia (-4,5%). Wykres korelacyjny tych dwóch zmiennych zaprezentowany został na rysunku 1.

[5] Wśród innych zastosowań tej metody można wymienić analizę zmian zatrudnienia [Batóg, Batóg, 2007] oraz zmian wydajności pracy [Batóg, Batóg, 2008].

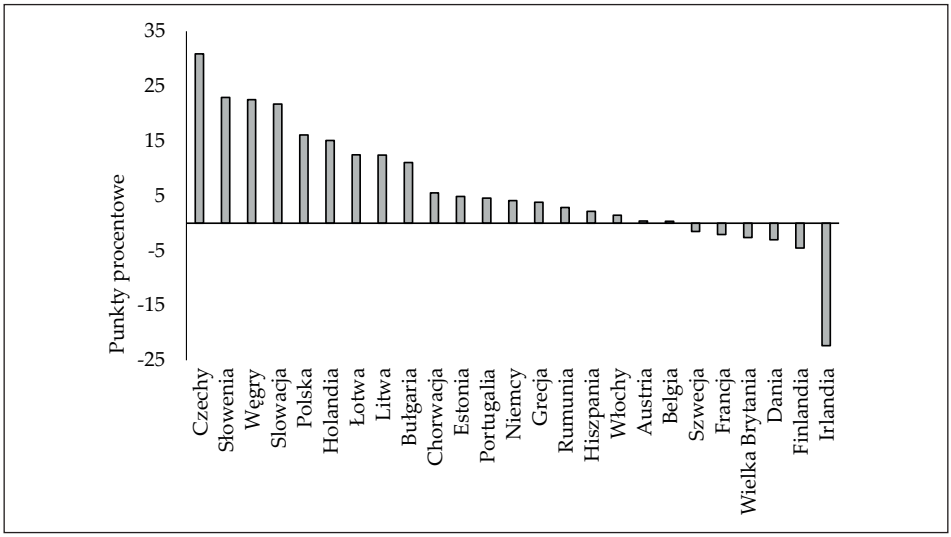
Rysunek 1. Związek między dynamiką PKB i eksportu w krajach UE w latach 2002–2015



Źródło: Opracowanie własne.

Na rysunku 2 przedstawiono zmianę udziału eksportu w PKB dla krajów UE, która miała miejsce w roku 2015, w porównaniu do roku 2002. Warto dodatkowo zauważyć, że udział ten w poszczególnych był stosunkowo zróżnicowany. W roku 2002 wahał się od 4,1% (Grecja) do 62,8% (Belgia), a w roku 2015 od 7,1% (Wielka Brytania) do 73,9% (Słowacja).

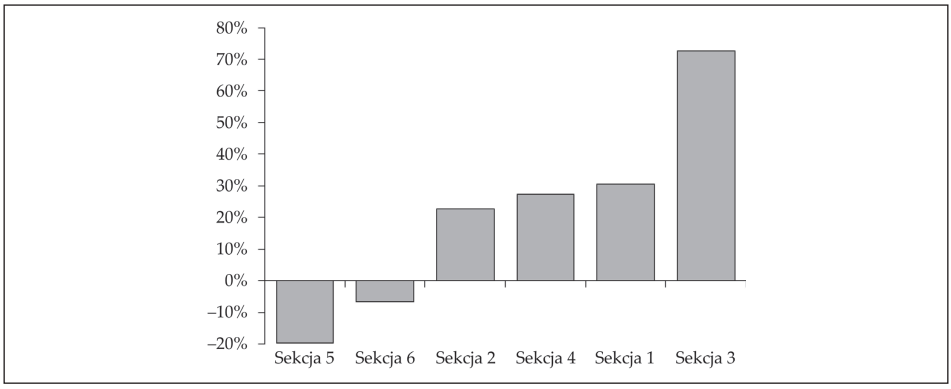
Rysunek 2. Zmiana udziału eksportu w PKB w krajach UE w latach 2002–2015



Źródło: Opracowanie własne.

Najwyższą dynamiką eksportu w okresie 2002–2015 charakteryzowały się produkty Sekcji 3 Paliwa mineralne, smary i materiały pochodne, a najniższą Sekcji 5 Maszyny, urządzenia i sprzęt transportowy, dla której odnotowano spadek eksportu w stosunku do ogólnej dynamiki unijnego eksportu wynoszącej 62,16% o prawie 20% (rys. 3).

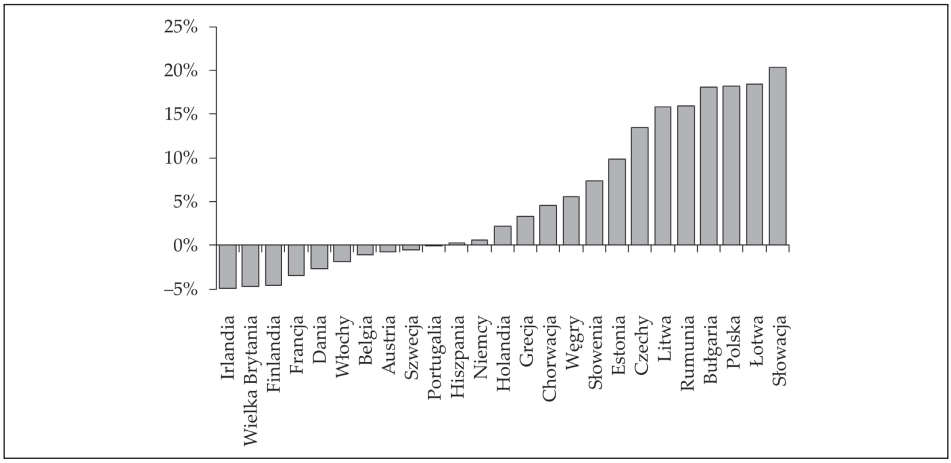
Rysunek 3. Zmiana eksportu w krajach UE według sekcji w latach 2002–2015 (jako odchylenie od stopy wzrostu eksportu ogółem 25 analizowanych krajów UE)



Źródło: Opracowanie własne.

Zmianę ogólną, która zaszła w eksporcie badanych krajów w latach 2002–2015 prezentuje rysunek 4. Najkorzystniejszą sytuację odnotowano w przypadku Słowacji (dodatni efekt krajowy równy 20,3%), a najmniej korzystną dla Irlandii (ujemny efekt krajowy na poziomie 4,9%).

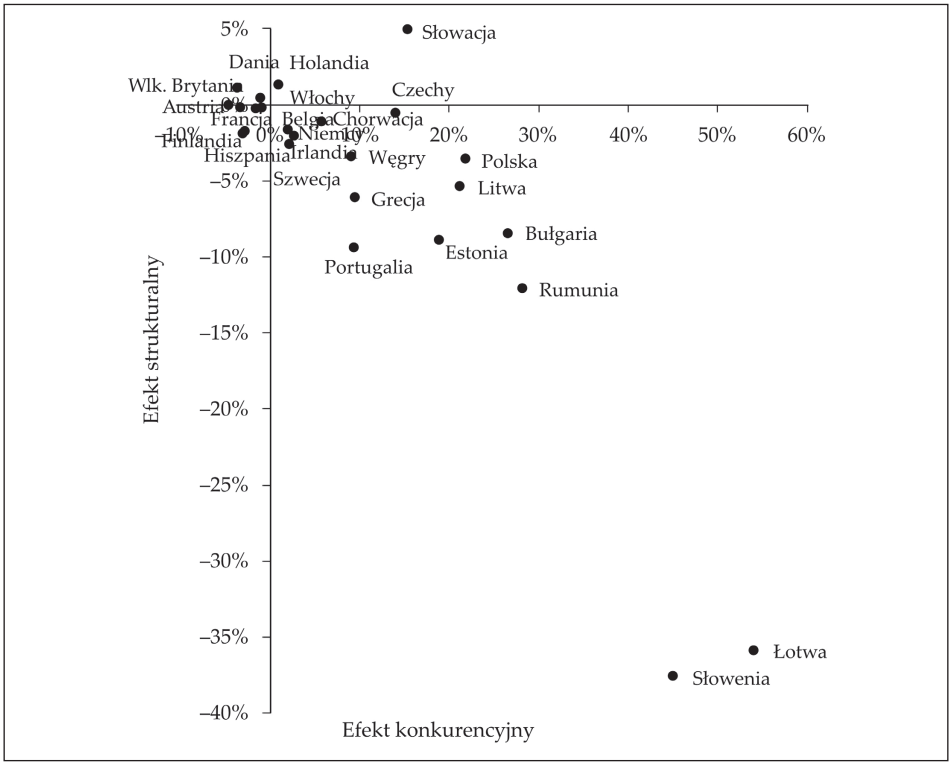
Rysunek 4. Efekt krajowy *REGj* dla krajów UE w latach 2002–2015



Źródło: Opracowanie własne.

Analiza przyczyn takiego kształtowania się zmiany ogólnej eksportu pozwala stwierdzić, że jej główną determinantą był efekt konkurencyjny, natomiast efekt strukturalny odgrywał w tym przypadku znacznie mniejszą rolę. Jest to szczególnie widoczne dla dwóch krajów: Słowenii i Łotwy, które charakteryzowały się bardzo wysokimi wartościami tego pierwszego efektu, przy ujemnym wpływie efektu strukturalnego (rys. 5). Świadczy to o niekorzystnej początkowej strukturze towarowej eksportu w tych krajach – posiadały one niski udział sekcji, dla których występował wysoki wzrost w całej Unii Europejskiej. Tego typu relacja może stanowić w przyszłości znaczące ograniczenie wewnątrzunijnego eksportu tych krajów. Najlepszą „pozycją" eksportową odznaczała się Słowacja, która posiadała dodatnie wartości zarówno efektu strukturalnego (1 pozycja wśród wszystkich badanych krajów), jak i efektu konkurencyjnego.

Rysunek 5. Efekt konkurencyjny i strukturalny dla krajów UE w latach 2002–2015



Źródło: Opracowanie własne.

Zakończenie

Do podstawowych tendencji charakteryzujących kształtowanie się eksportu krajów Unii Europejskiej w latach 2002–2015 zaliczyć można: silny spadek eksportu w okresie kryzysu gospodarczego oraz wyższe stopy wzrostu wartości eksportu nowych krajów członkowskich w porównaniu do starych członków UE – por. na przykład [Cieślik i inni, 2014a], przy czym te pierwsze posiadały również wyższe udziały eksportu w produkcie krajowym brutto. Główną przyczyną korzystniejszej dynamiki eksportu krajów Europy Środkowo-Wschodniej jest jednak szybszy wzrost eksportu w jego poszczególnych sekcjach w porównaniu do przeciętnej dynamiki eksportu dla całej UE (dodatni efekt konkurencyjny), a nie poprawa struktury towarowej eksportu (niski dodatni lub wręcz ujemny efekt strukturalny). Uzyskany wynik jest zgodny z rezultatami wybranych wcześniejszych badań – por. na przykład [Wernerheim, 2000]. Potwierdza on również wyniki otrzymywane przez innych autorów, wskazujące na niski poziom konkurencyjności eksportowanych dóbr przez te kraje [Michalski, 2014b]. W przypadku czynnika strukturalnego jest to spowodowane głównie niskim udziałem wysoko przetworzonych dóbr w eksporcie ogółem, w którym dominują sekcje o charakterze surowcowym. Natomiast w odniesieniu do czynnika konkurencyjnego charakteryzującego kraje Europy Środkowo-Wschodniej, oprócz stopnia aktywności gospodarczej stanowiącego czynnik egzogeniczny, należy zwrócić uwagę na dodatkowe determinanty w postaci: wzrostu otwartości gospodarek i zwiększonego dostępu do rynków zbytu, znacznego importu nowych technologii związanego w dużym stopniu z napływem bezpośrednich inwestycji zagranicznych, jak również występowania konkurencyjności cenowej eksportowanych dóbr, będącej w dużej mierze wynikiem niskich kosztów pracy.

W dalszych badaniach interesującym kierunkiem analiz mogłaby być ocena wpływu: wyboru początkowego okresu analizy oraz zmiany stopnia agregacji struktury towarowej eksportu na uzyskiwane wyniki. Innym obszarem zastosowania metody *shift-share* może być poszukiwanie nowych rynków zbytu oraz identyfikacja produktów o najwyższym potencjale wzrostu eksportu.

Literatura

Antczak E. (2014), *Analiza zanieczyszczenia powietrza w Polsce z wykorzystaniem przestrzennej dynamicznej metody przesunięć udziałów*, „Ekonomia i Środowisko", nr 2(49).

Antczak E., Lewandowska-Gwarda K. (2015), *Analysis of Internal Emigration in Poland with Using the Spatial Dynamic Shift-Share Method*, „Argumenta Oeconomica", nr 2(35).

Artige L., van Neuss L. (2015), *A New Shift-Share Method, Growth and Change*, January.

Balaguer J., Cantavella-Jordá M. (2004), *Structural Change in Exports and Economic Growth: Cointegration and Causality Analysis for Spain (1961–2000)*, „Applied Economics", Vol. 36, No. 5.

Batóg J., Batóg B. (2007), *Employment Growth within BSR Countries: Some Predictive Aspects of Shift-share Analysis,* Baltic Business and Socio-Economic Development, Berliner Wissenschafts-Verlag, Berlin.

Batóg J., Batóg B. (2008), *Analiza regionalnych zmian wydajności pracy w Polsce,* „Wiadomości Statystyczne", nr 6.

Cieślik A., Michałek A., Michałek J. (2014b), *European Integration and Firm's Export Performance in the New EU Member Countries*, „Poznań University of Economics Review", Vol. 14, No. 4.

Cieślik A., Michałek J., Michałek A. (2014a), *The Influence of Firm Characteristics and Export Performance in Central and Eastern Europe: Comparisons of Visegrad, Baltic and Caucasus States*, „Entrepreneurial Business and Economics Review", Vol. 2, No. 1.

Coughlin C.C., Pollard P.S. (2001), *Comparing Manufactiring Export Growth Across States: What Accounts for the Differences?*, Federal Reserve Bank of St. Louis, February.

Dawson P.J., Hubbard L.J. (2004), *Exports and Economic Growth in Central and East European Countries During Transition*, „Applied Economics", Vol. 36, No. 16.

Dudziński J., Narękiewicz J. (2014), *Struktura eksportu jako determinanta jego dynamiki we współczesnym handlu międzynarodowym (na przykładzie Ameryki Łacińskiej)*, „Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania", nr 35, t. 1, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

ec.europa.eu/eurostat/data/database, dostęp: 9.01.2017.

Głodowska A. (2016), *Struktura specjalizacji eksportowej a proces konwergencji państw Unii Europejskiej*, „Finanse, Rynki Finansowe, Ubezpieczenia", nr 3(81).

Green R.T, Allaway A.W. (1985), *Identification of Export Opportunities: A Shift-share Approach*, „Journal of Marketing", Vol. 49, No. 1.

Herschede F. (1991), *Competition among ASEAN, China, and the East Asia NICs. A Shift-Share Analysis*, „ASEAN Economic Bulletin", Vol. 7, No. 3.

Jewczak M., Żółtaszek A. (2011), *Spatial Shift-share Analysis as a Health Policy Tool*, Acta Universitatis Lodziensis, "Folia Oeconomica", nr 252.

Krawczyk A., Nogaj Ł., Wojcieszczyk M. (2016), *Impact of Exports on the GDP Growth in Poland in Years 2009–2014*, „Central European Review of Economics & Finance", Vol. 16, No. 6.

Lechman E. (2014), *Changing Patterns in the Export of Goods Versus International Competitiveness. A Comparative Analysis for Central-East European Countries in the Period 2000–2011*, „Comparative Economic Research" Vol. 17, No. 2.

Michalski B. (2014a), *Competitiveness of Polish Mid-tech and High-tech Exports to the European Union (EU-27) in the First Decade of the 21st Century*, „Poznań University of Economics Review", Vol. 14, No. 4.

Michalski B. (2014b), *Polski eksport w warunkach światowej dekoniunktury (2008–2012). Identyfikacja polskich hitów eksportowych*, „Biznes Międzynarodowy w Gospodarce Globalnej", nr 33.

Neri F., Jayanthakumaran K. (2005), *Trends in Manufactured Exports Across the States and Territories of Australia 1989/1990–2000/2001: A Shift-share Analysis*, „Economic Papers", Vol. 24, No. 2.

Olczyk M., Kordalska A. (2015), *Determinanty wzrostu polskiego eksportu przemysłu przetwórczego – analiza długookresowa*, „Studia Ekonomiczne", nr 4(85).

Siedlecki J. (2000), *Równowaga a wzrost gospodarczy*, Wydawnictwo Naukowe PWN, Warszawa–Wrocław.

Vu V.H., Holmes M., Tuyen T.Q., Lim S. (2016), *Firm Exporting and Productivity: What if Productivity is no Longer a Black Box*, „Baltic Journal of Economics", Vol. 16, No. 2.

Wernerheim C. M. (2000), *Cointegration and Causality in the Exports-GDP Nexus: The Post-war Evidence for Canada*, „Empirical Economics", Vol. 25, No. 1.

Streszczenie

Jednym z istotnych aspektów rozwoju gospodarki jest jej zdolność do intensyfikacji handlu zagranicznego, a zwłaszcza eksportu. Ważna jest nie tylko jego wartość, ale również struktura. W teorii ekonomii definiowanych jest wiele zróżnicowanych czynników determinujących poziom eksportu. Za jeden z kluczowych czynników uważana jest przynależność do określonego związku państw lub porozumienia handlowego. W artykule zweryfikowana została hipoteza o pozytywnym wpływie wejścia nowych krajów członkowskich do Unii Europejskiej na wartość ich eksportu. Zastosowana zmodyfikowana analiza *shift-share* w ujęciu dynamicznym pozwoliła również na identyfikację, czy czynnikiem sukcesu w handlu zagranicznym jest struktura eksportu czy raczej konkurencyjność eksportowanych dóbr. Badanie przeprowadzono na podstawie danych dotyczących eksportu dla 25 krajów Unii Europejskiej w latach 2002–2015.

Słowa kluczowe

eksport, Unia Europejska, analiza *shift-share*

Analysis of Factors Influencing on the Export of European Union Countries (Summary)

One of the important aspects of economic development is the ability of the economy to intensification of the foreign trade and especially export. Both volume and structure of the export have great meaning. In the literature many factors determining the level of export are indicated. The memberships of union of countries or trade agreement are of a big importance. In the paper the hypothesis that EU accession had positive impact on the level of export for the new members of EU was verified. The dynamic *shift-share* analysis was applied. This method allows to determine whether the key success factor in foreign trade is the structure of the export or competiveness of exported goods and products. The research was conducted on the base of data on export for 25 European countries in 2002–2015.

Keywords

export, European Union, *shift-share* analysis

Hanna Godlewska-Majkowska*
Dorota Perło**

Zastosowanie metody wagowo-korelacyjnej i modelowania miękkiego do analizy atrakcyjności inwestycyjnej powiatów w Polsce

Wstęp

W dobie globalizacji i większej mobilności zasobów pracy decyzje o wyborze miejsca zamieszkania czy inwestowania są bardziej płynne, choć zależne od wielu determinant. Ważnym czynnikiem decydującym o wyborze miejsca na świecie jest jakość życia, która wiąże się z potrzebą rozwoju, realizacji własnych planów i wykorzystania doświadczeń. Jest ona skorelowana z jakością pracy. Wiele miast w Polsce określiło swoją misję rozwoju jako miejsce o wysokiej jakości życia i atrakcyjnym rynku pracy. Udziały w podatku z PIT i CIT stanowią kluczowe źródło dochodów własnych gminy, dzięki którym mogą się one rozwijać. Oznacza to, że zarządzanie takimi jednostkami skupia się na przyciąganiu inwestorów i osiąganiu dodatniego salda migracji.

Dlatego tak dużego znaczenia nabiera obecnie dostęp do kompleksowej informacji o miejscach atrakcyjnych inwestycyjnie, w których są relatywnie lepsze warunki do rozwoju własnej firmy lub znalezienia interesującej pracy, a także o tym, jakie czynniki je kształtują. Potrzeba tworzenia takich analiz i zestawień wynika zarówno z oczekiwań mieszkańców, inwestorów, jak i decydentów, odpowiedzialnych za kształtowanie warunków poprawiających walory lokalizacyjne, przyciągające potencjalnych inwestorów. Naukowcy wychodzą im naprzeciw, tworząc różne rankingi regionów, podregionów, powiatów, czy gmin pod względem szeroko rozumianej atrakcyjności inwestycyjnej. Ich odbiorcami są wszystkie zainteresowane strony. Uczeni, konstruując modele, opracowując analizy i interpretując wpływ określonych zmiennych, zależności między nimi, biorą pod uwagę różne wyspecyfikowane obszary badawcze (mikroklimaty), dobierają często odmienne czynniki decydujące o podjęciu konkretnej decyzji

* Prof. dr hab., Zakład Otoczenia Biznesu, Instytut Przedsiębiorstwa, Szkoła Główna Handlowa w Warszawie, al. Niepodległości 162, 02-554 Warszawa, hgodle@sgh.waw.edu.pl
** Dr, Zakład Ekonometrii i Statystyki, Wydział Ekonomii i Zarządzania, Uniwersytet w Białymstoku, ul. Marii Skłodowskiej-Curie 14, 15-097 Białystok, dorota.perlo@uwb.edu.pl

lokalizacyjnej, a także stosują w tym celu różne metody badawcze, najczęściej wywodzące się z wielowymiarowej analizy porównawczej (WAP)[1].

Efektem zastosowanej niejednorodnej metodyki badań są zazwyczaj miary syntetyczne, wykorzystywane do budowy rankingów, a skonstruowane na bazie różnych narzędzi. Przykładowo, w badaniach nad potencjałem atrakcyjności inwestycyjnej (PAI) pod kierunkiem H. Godlewskiej-Majkowskiej stosowana jest metoda wagowo-korelacyjna [Godlewska-Majkowska, 2012]. Natomiast w analizach prowadzonych przez Instytut Badań nad Gospodarką Rynkową – procedura jednocechowa [Tarkowski, 2015, s. 13]. W województwie śląskim w celu zbadania atrakcyjności inwestycyjnej został wykorzystany syntetyczny miernik atrakcyjności inwestycyjnej gmin (SMAIG) [*Atrakcyjność…*, 2009, s. 16]. Z kolei atrakcyjność inwestycyjną gmin pod względem miejsca lokalizacji hoteli w Południowo-Zachodniej Polsce badano za pomocą taksonomicznego miernika wg wzorca i regresji logistycznej [Puciato, Dziedzic, 2016, s. 1–10]. Inny przykład stanowi wskaźnik atrakcyjności inwestycyjnej Ukrainy (EBA), który jest obliczany w prosty sposób, jako średnia nieważona ocen opartych na pięciu wyodrębnionych obszarach wskazujących na klimat inwestycyjny [*EBA…*, 2017]. Jednak konstrukcję miary syntetycznej poprzedzają w tym przypadku badania jakościowe. Praktycznie w każdym z wymienionych przykładowych badań nad atrakcyjnością inwestycyjną wykorzystywane są narzędzia, które bazują na odmiennych podstawach metodologicznych. Analizowane są również odmienne obszary badawcze. W obecnej sytuacji na rynku globalnym czynniki lokalizacyjne są bardzo zróżnicowane, zależą w znacznym stopniu od specyfiki inwestycji (czy ma powstać główna siedziba, czy jej oddział, jakiego rodzaju branży dotyczy, jakie znaczenie posiadają koszty i dostępność zasobów pracy, jakie jest oddziaływanie inwestycji na środowisko, jaki jest klimat inwestycyjny itp.), co uzasadnia potrzebę prowadzonych badań naukowych.

Celem artykułu jest analiza porównawcza wyników badania atrakcyjności inwestycyjnej powiatów w Polsce wykonanego za pomocą dwóch metod wyznaczania wskaźnika syntetycznego: wagowo-korelacyjnej i modelowania miękkiego. Postawiono hipotezę, że zastosowanie różnych sposobów wyznaczania miernika syntetycznego daje w przybliżeniu takie same wyniki klasyfikacji powiatów w Polsce pod względem atrakcyjności inwestycyjnej. Skonfrontowanie wyników badań uzyskanych za pomocą dwóch analizowanych sposobów wyznaczania miernika syntetycznego ma na celu zdiagnozowanie ich podobieństw i różnic oraz pozwala na

[1] Należą do nich m.in. publikacje z kanonu badań statystyczno-ekonometrycznych, jak: [Borys, Walesiak, 1983; Grabiński i inni, 1983; Kolenda, 2006; Ostasiewicz, 1989; Zaucha i inni, 2015].

zdywersyfikowane, w określonym stopniu, podejście do analizowanego zagadnienia. Umożliwia również dokonanie weryfikacji uzyskanych różnic/podobieństw pod kątem zakłóceń wynikających z konieczności wykorzystania „uproszczeń" w stosowanych modelach. Analizy porównawcze zastosowania różnych metod statystyczno-ekonometrycznych stanowią przedmiot zainteresowania wielu naukowców[2].

Analizę porównawczą przeprowadzono na podstawie danych statystycznych z 2015 r., zagregowanych na poziomie 379 powiatów ziemskich i grodzkich.

1. Sposoby wyznaczania miernika syntetycznego

W badaniu wykorzystano dwa sposoby wyznaczania miernika syntetycznego: metodę wagowo-korelacyjną i modelowanie miękkie.

1.1. Metoda wagowo-korelacyjna w pomiarze atrakcyjności inwestycyjnej powiatów w Polsce

O atrakcyjności inwestycyjnej regionów różnych szczebli taksonomicznych świadczą cechy poszczególnych miejsc wpływające na koszty prowadzenia działalności w danym miejscu, możliwe do uzyskania przychody, a także wysokość podatków. Dlatego na liście najważniejszych dla inwestorów czynników determinujących decyzje inwestycyjne umieszczonej w *World Investment Report 2010 UNCTAD* znajdują się takie elementy, jak: dynamiczny wzrost rynku wewnętrznego, wielkość rynku wewnętrznego, dostęp do wykwalifikowanych pracowników, dostęp do międzynarodowych rynków, obecność poddostawców i partnerów, atrakcyjny system zachęt inwestycyjnych. Wagę tych czynników potwierdzają także wyniki badań decyzji lokalizacyjnych inwestorów zagranicznych w Polsce[3].

Atrakcyjność inwestycyjną rozpatrywaną łącznie dla całej gospodarki regionalnej powinno opisywać możliwie dużo zmiennych diagnostycznych,

[2] Przykładowo w artykule [Guzik, 2008] porównywane są wyniki badania atrakcyjności inwestycyjnej uzyskane przez Instytut Badań nad Gospodarką Rynkową [Kalinowski, 2006] z badaniami własnymi wykonanymi za pomocą modeli ekonometrycznych (regresyjnych). W. Guzik w badaniu naukowym zakłada, podobnie jak autorzy raportów IBnGR, że w empirii ocena atrakcyjności inwestycyjnej jest zazwyczaj dokonywana przez pryzmat poniesionych nakładów inwestycyjnych, a analizy na gruncie wielowymiarowej analizy porównawczej stanowią jedynie jej modelową ocenę, dlatego pojawiają się znaczne rozbieżności w wynikach badań realizowanych w ww. kategoriach. Inny przykład stanowi artykuł [Migdał-Najman, Najman, 2013], w którym autorzy wykonują analizę porównawczą klasycznych i hybrydowych metod grupowania, wykonując eksperymenty bazujące na jednakowych zbiorach, wykazując wyższość metod hybrydowej sieci neuronowej typu SOM+GNG nad klasycznymi.

[3] Najważniejszymi czynnikami lokalizacji kapitału zagranicznego są korzystne położenie i połączenia komunikacyjne miejscowości, duży rynek zbytu w regionie, znaczna podaż i niska cena siły roboczej, a także możliwość przejęcia wolnych obiektów, możliwość znalezienia nowych rynków zbytu, związek polskiego partnera z regionem oraz związki partnera zagranicznego z regionem [Durka, 2000].

odnoszących się do szerokiego spektrum rozwoju społeczno-gospodarczego. Czynniki te znajdują odzwierciedlenie w doborze 47 zmiennych, składających się na wskaźnik potencjalnej atrakcyjności inwestycyjnej regionów dla inwestycji w gospodarkę narodową (PAI2_GN). Składa się on z mikroklimatów: „zasoby pracy", „infrastruktura techniczna", „infrastruktura społeczna", „rynkowy", „administracja" – patrz tablica A w aneksie.

Przy doborze zmiennych i konstrukcji wskaźnika potencjalnej atrakcyjności inwestycyjnej kierowano się następującymi zasadami:

- w ocenie atrakcyjności inwestycyjnej dopuszczone są zmienne jednocechowe i pseudojednocechowe, stanowiące w sposób jednoznaczny stymulanty lub destymulanty potencjalnej atrakcyjności inwestycyjnej;
- wszystkie zmienne jednocechowe i pseudojednocechowe nie są ze sobą silnie skorelowane ($r < 0{,}70$); jeśli jednak informacja o skorelowanych zmiennych jest potrzebna do dalszych analiz, np. analizy strategicznej jednostek samorządu terytorialnego (JST), wówczas do modelu można wprowadzić skorelowane zmienne z przelicznikiem ⅓, gdy trzy zmienne są skorelowane, lub ½, gdy korelacja dotyczy pary zmiennych;
- wskaźniki pseudojednocechowe składają się na mikroklimaty, a mikroklimaty na wskaźnik syntetyczny potencjalnej atrakcyjności inwestycyjnej;
- wskaźniki pseudojednocechowe są konstruowane przy dążeniu do eliminacji subiektywności wag.

Za podstawę prac nad parametryzacją atrakcyjności inwestycyjnej została przyjęta metoda wagowo-korelacyjna. Jej istota tkwi w zastosowaniu współczynników korelacji liniowej Pearsona pomiędzy zmiennymi pseudojednocechowymi a wskaźnikiem końcowym, jako wag umożliwiających budowanie wskaźnika syntetycznego, służącego ocenie jednostek terytorialnych.

Procedura jest następująca:

1. Standaryzacja zmiennych składowych, tworzących wskaźnik pseudojednocechowy atrakcyjności, w oparciu o następujące wzory:
   - w przypadku cech o charakterze negatywnym,

$$\frac{X_{\max} - X_{\mathrm{i}}}{X_{\max} - X_{\min}} \cdot 100\% \tag{1}$$

   - w przypadku cech o charakterze pozytywnym,

$$\frac{X_{\mathrm{i}} - X_{\min}}{X_{\max} - X_{\min}} \cdot 100\% \tag{2}$$

gdzie:
X_{max} – wartość maksymalna cechy w badanym zbiorze jednostek,
X_{min} – wartość minimalna cechy w badanym zbiorze jednostek,
X_i – *i*-ta wartość cechy w badanym zbiorze jednostek.

2. Obliczenie wartości poszczególnych wskaźników pseudojednocechowych (mikroklimatów) jako średniej ważonej wystandaryzowanych zmiennych wyjściowych. Wagi pojawiające się na tym etapie są stosowane, w przypadku gdy w zestawie zmiennych wyjściowych następuje nadreprezentatywność określonych zmiennych wyjściowych lub gdy są one nadmiernie skorelowane.
3. Obliczanie końcowego wskaźnika syntetycznego na podstawie poniższych kroków:
   1) wyznaczenie nieważonej średniej arytmetycznej zmiennych pseudojednocechowych (mikroklimatów) – początkowej postaci wskaźnika syntetycznego;
   2) obliczenie wskaźników korelacji pomiędzy zmiennymi pseudocechowymi a wskaźnikiem syntetycznym, czyli pomiędzy mikroklimatami a ich średnią arytmetyczną w postaci wskaźnika syntetycznego;
   3) następnie, wykorzystując obliczone współczynniki, na nowo przelicza się miernik pseudojednocechowy (wskaźnik syntetyczny) potencjalnej atrakcyjności inwestycyjnej jako średnią ważoną mierników cząstkowych, gdzie poziom wag wyznaczony został przez kolejne iteracje współczynnika korelacji (pomiędzy danym mikroklimatem a średnią ważoną wszystkich mikroklimatów), prowadzące do jego stabilizacji.

W celu ułatwienia porównania wyników osiągniętych przez poszczególne jednostki terytorialne wielkości poszczególnych wskaźników cząstkowych i miernika pseudojednocechowego zostały przyporządkowane do 6 klas, których zakres został wyznaczony przez lewostronnie domknięte przedziały o następujących dolnych granicach:

- klasa A: Av + S(x),
- klasa B: Av + 0,5 S(x),
- klasa C: Av,
- klasa D: Av – 0,5 S(x),
- klasa E: Av – S(x),
- klasa F: 0,

gdzie:
Av – średnia arytmetyczna,
S(x) – odchylenie standardowe[4].

[4] Opis metody zawiera m.in. publikacja [Poniatowska-Jaksch, 2006].

Zgodnie z przedstawioną metodą stworzono wskaźniki potencjalnej atrakcyjności inwestycyjnej w wersji zredukowanej – PAI1, które są porównywalne na wszystkich szczeblach taksonomicznych. Jest to możliwe dzięki budowie na wszystkich szczeblach terytorialnych kompatybilnych baz danych, zawierających identyczne zestawy zmiennych wyjściowych.

Wskaźnik ten jest przede wszystkim użyteczny w badaniu atrakcyjności inwestycyjnej dowolnych konglomeratów jednostek statystycznych. Pozwala zatem na badanie jednostek administracyjnych w ujęciu zarówno *a priori*, jak i *ex post* – czyli regionów planowanych lub już nieistniejących. Ma to bardzo duże znaczenie dla prowadzenia analiz przestrzennych, wyprzedzających korekty podziału administracyjnego.

W oparciu o badania nad atrakcyjnością inwestycyjną prowadzone w Instytucie Przedsiębiorstwa SGH systematycznie od 2007 r. można stwierdzić, że wskaźniki te wykazują niezmiennie wysoki poziom ocen mikroklimatu rynkowego i infrastruktury technicznej, co potwierdza duże znaczenie wyposażenia infrastrukturalnego danego terenu, jak również jego zamożność lub zdolność do generowania efektu dochodowego w sferze przedsiębiorstw. Również ta zasada znajduje potwierdzenie w ocenie wagi poszczególnych mikroklimatów w wyznaczonych wskaźnikach PAI1 dla gospodarki narodowej dla wszystkich szczebli podziału na regiony statystyczne (tab. 1).

Tablica 1. Wagi mikroklimatów tworzących wskaźnik potencjalnej atrakcyjności inwestycyjnej dla gospodarki narodowej (PAI1_GN) dla roku 2015

JST	Mikroklimat zasoby pracy	Mikroklimat infra-struktura techniczna	Mikroklimat infra-struktura społeczna	Mikroklimat rynkowy	Mikroklimat administra-cja
Woje-wództwo	0,245961	0,895526	0,289616	0,819305	0,686149
Podregion	0,370055	0,915134	0,450115	0,949205	0,639077
Powiat	0,350096	0,903024	0,402357	0,929154	0,658598
Gmina	0,245957	0,895534	0,289616	0,819305	0,686145

Źródło: Opracowanie własne.

Walory lokalizacyjne poszczególnych miejsc składają się na atrakcyjność inwestycyjną. Z uwagi na to, że satysfakcję inwestora można rozpatrywać z punktu widzenia walorów lokalizacyjnych danego miejsca, jak i oczekiwanych efektów dokonanej inwestycji, proponuje się rozróżnienie atrakcyjności inwestycyjnej na potencjalną atrakcyjność inwestycyjną oraz rzeczywistą atrakcyjność inwestycyjną.

Potencjalna atrakcyjność inwestycyjna to ujęcie nawiązujące do oceny lokalizacji przed przystąpieniem do inwestycji. W tym ujęciu są analizowane walory lokalizacyjne. Konstrukcja wskaźników opisujących potencjalną atrakcyjność inwestycyjną nawiązuje do wiodących czynników lokalizacji, pogrupowanych na komponenty opisujące czynnik demograficzny, rynkowy, infrastrukturę techniczną, infrastrukturę społeczną oraz proinwestycyjne zachowania samorządów[5].

Zaletą tego wskaźnika jest konstrukcja umożliwiająca pełną porównywalność ocen atrakcyjności inwestycyjnej na wszystkich szczeblach podziału statystycznego kraju. Wadą jest natomiast ograniczenie zestawu zmiennych diagnostycznych, z uwagi na konieczność ich dostępności dla wszystkich szczebli podziału statystycznego kraju. Mimo tych trudności w wyniku ewaluacji wskaźników i włączenia do miary syntetycznej wskaźnika administracyjnego w wersji podstawowej wskaźnik ten opiera się na łącznie 47 zmiennych diagnostycznych, co wydaje się wystarczająco szerokim fundamentem porównań (por. tablica A).

Wszystkie wskaźniki są wyznaczane na podstawie metody wagowo-korelacyjnej, umożliwiającej wyznaczenie wag zmiennych pseudojednocechowych w oparciu o cechy rozkładu statystycznego, dzięki czemu następuje minimalizacja subiektywizmu ocen końcowych.

Przyjęta metoda pozwala na zminimalizowanie wpływu subiektywnej oceny autora na wyniki końcowe, przy jednoczesnym uwzględnieniu niejednakowego wpływu poszczególnych zmiennych na wielkość wskaźnika końcowego. Odpowiada to w większym stopniu rzeczywistości niż w przypadku pominięcia rang poszczególnych zmiennych cząstkowych.

Obliczone wartości syntetycznego wskaźnika pseudojednocechowego są podstawą podziału zbioru jednostek przestrzennych na klasy od A do F, których zakres został wyznaczony przez lewostronnie domknięte przedziały o następujących dolnych granicach: Klasa A: Av + S(x), Klasa B: Av + 0,5S(x), Klasa C: Av, Klasa D: Av – 0,5S(x), Klasa E: Av – S(x), Klasa F: 0, gdzie: Av – średnia arytmetyczna, S(x) – odchylenie standardowe.

1.2. Modelowanie miękkie w pomiarze atrakcyjności inwestycyjnej powiatów w Polsce

W celu dokonania analizy porównawczej wyników pomiaru atrakcyjności inwestycyjnej powiatów w Polsce badanych na podstawie różnych metod jako drugą z nich zastosowano modelowanie miękkie. Specyfikacji modelu miękkiego dokonano na bazie analogicznego jak w przypadku metody wagowo-korelacyjnej zbioru 47 zmiennych diagnostycznych. Do obliczeń wykorzystano program MM (modelowanie miękkie)

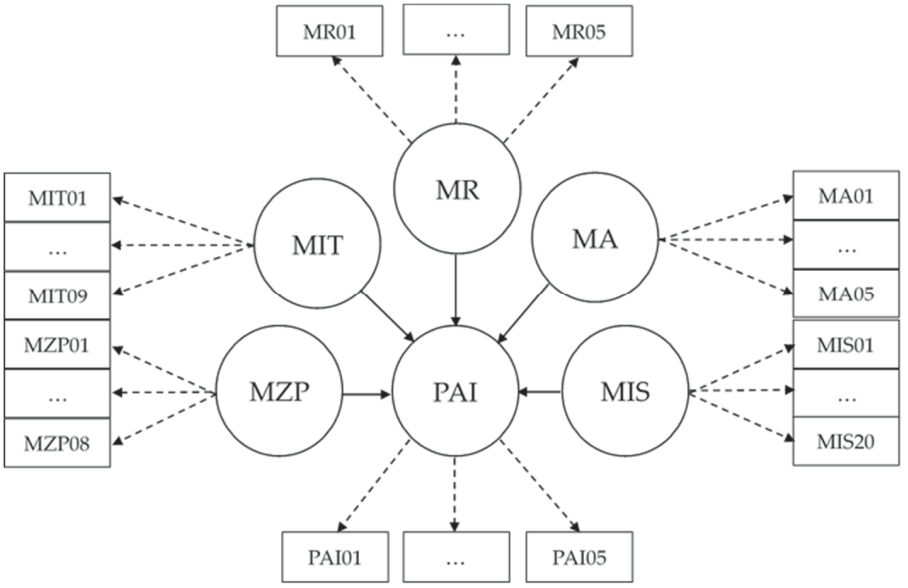
[5] Więcej na ten temat metodyki pomiaru w: [Godlewska-Majkowska, 2012; Godlewska-Majkowska, 2011; Godlewska-Majkowska, 2010].

autorstwa M. Karaś (SGH) opracowany na bazie programu PLS autorstwa J. Rogowskiego.

Każdy model miękki składa się z dwóch części: modelu wewnętrznego (teoretycznego) oraz zewnętrznego (miary) [Joreskog, Wold, 1982; Rogowski, 1990; Rogowski, 2002, s. 23–34; Perło, 2014, s. 86–100; Kock, Mayfield, 2015, s. 113–130; Lee i inni, 2013, s. 1705–1722; Servera-Francés i inni, 2012; Tenenhaus i inni, 2005, s. 159–205; Wold, 1982]. Obydwa modele są ze sobą powiązane, tzn. obydwa jednocześnie są wykorzystywane w procesie estymacji parametrów. Model wewnętrzny opisuje relacje zachodzące między zmiennymi nieobserwowalnymi, czyli opisuje relacje teoretyczne wynikające z przyjętej teorii. W modelu wewnętrznym potencjału atrakcyjności inwestycyjnej powiatów w Polsce (PAI) zdefiniowano identyczne mikroklimaty jak w przypadku metody wagowo-korelacyjnej, tj.: mikroklimat zasobów pracy (MZP), mikroklimat infrastruktury technicznej (MIT), mikroklimat infrastruktury społecznej (MIS), mikroklimat rynkowy (MR) i mikroklimat administracja (MA). Model teoretyczny bada wpływ wszystkich ww. mikroklimatów na potencjalną atrakcyjność inwestycyjną. Zastosowanie analizowanej metody wymaga założenia, że relacje w modelach miękkich mają charakter liniowy. Drugi model – model miary – pokazuje, w jaki sposób indykatory odzwierciedlają swoje zmienne ukryte, czyli zawiera on definicje zmiennych teoretycznych. W przypadku modelu miękkiego przyjmuje się, że zmienne ukryte są kombinacjami liniowymi swoich indykatorów (są to tzw. relacje wagowe).

Zmienne nieobserwowalne można definiować dwoma sposobami: dedukcyjnie i indukcyjnie. W zależności od podejścia otrzymuje się różnice w sposobie i w wynikach estymacji. W artykule przyjęto podejście dedukcyjne, które zakłada, że zmienna ukryta, jako pojęcie teoretyczne, jest punktem wyjścia do poszukiwania danych empirycznych (przejście od pojęć teoretycznych do indykatorów), co oznacza, że jest to zmienna pierwotna w stosunku do danego indykatora. Wskaźniki tego typu zmiennych nieobserwowalnych nazywa się odbijającymi (odzwierciedlającymi). Indykatory odbijające, z założenia, powinny charakteryzować się wysoką korelacją między sobą, dlatego przy doborze zmiennych należy kierować się względami merytorycznymi, a nie opierać się na klasycznych metodach doboru. Metody klasyczne zakładają bowiem niską zależność między zmiennymi objaśniającymi [Kuszewski, 2000, s. 14–16]. W modelu pomiaru atrakcyjności inwestycyjnej wszystkie zmienne ukryte zostały zdefiniowane na podstawie zmiennych diagnostycznych wyspecyfikowanych metodą wagowo-korelacyjną (por. tablica A). Schemat modelu wewnętrznego i zewnętrznego prezentuje rysunek 1.

Rysunek 1. Schemat modelu miękkiego atrakcyjności inwestycyjnej



W którym:
——► relacja wewnętrzna,
- - - - - ► relacja zewnętrzna,
○ zmienna ukryta,
▭ indykator.
Źródło: Opracowanie własne.

Model wewnętrzny prezentuje wpływ badanych mikroklimatów na atrakcyjność inwestycyjną. Relację przedstawia poniższe równanie liniowe:

$$\mathbf{PAI}_t = \alpha_1\mathbf{MZP}_t + \alpha_2\mathbf{MIT}_t + \alpha_3\mathbf{MIS}_t + \alpha_4\mathbf{MR}_t + \alpha_5\mathbf{MA}_t + \alpha_6 + \varepsilon_t \qquad (3)$$

gdzie:
$\mathbf{PAI}_t$ – zmienna ukryta endogeniczna w momencie t,
$\mathbf{MZP}_t$, $\mathbf{MIT}_t$, $\mathbf{MR}_t$, $\mathbf{MA}_t$, $\mathbf{MIS}_t$ – zmienne ukryte egzogeniczne w momencie t,
α_i – parametry strukturalne modelu, $i = 1, 2, \ldots, 6$;
ε_t – składnik losowy.

Model miękki jest szacowany częściową metodą najmniejszych kwadratów PLS (*partial least squares*). W metodzie PLS szacuje się jednocześnie parametry modelu miary i modelu teoretycznego. W wyniku estymacji, oprócz tych parametrów, otrzymuje się również oszacowania wartości zmiennej ukrytej, które można traktować jako miarę syntetyczną. Wielkości te zależą nie tylko od relacji zewnętrznych, ale również od założonych w modelu wewnętrznym związków między zjawiskami złożonymi. Oznacza to, że proces poznania jest uzależniony od opisu teoretycznego. Oszacowania zmiennych nieobserwowalnych nie mają interpretacji merytorycznej, ale

można interpretować zmiany ich wartości. Zmienne analizowane w modelowaniu miękkim określane są jako stymulanty lub destymulanty. Jeżeli estymatory wag i ładunków czynnikowych dla indykatorów będących stymulantami danej zmiennej obserwowalnej są dodatnie, a dla będących destymulantami ujemne, to większa wartość tej zmiennej wskazuje na wyższy poziom badanego zjawiska na danym obiekcie. Interpretując kolejność tych liczb, dokonuje się analizy porównawczej.

Przed przystąpieniem do analizy wyników porządkowania otrzymanych na podstawie modelu miękkiego należy dokonać jego weryfikacji. Na początku, podobnie jak w klasycznym modelu ekonometrycznym, weryfikacji merytorycznej, a dopiero później statystycznej. Oceny parametrów relacji zewnętrznych i wewnętrznych muszą być zgodne z przyjętym opisem teoretycznym. W celu sprawdzenia jakości modelu miękkiego można posłużyć się miernikami należącymi do „klasyki" ekonometrii, traktując oszacowane wartości zmiennych ukrytych jako wartości zaobserwowane w rzeczywistości. W ten sposób wyznacza się kwadraty współczynników korelacji wielorakiej (R^2), czy też na przykład odchylenia standardowe estymatorów parametrów.

Własności statystyczne modelu miękkiego sprawdza się głównie w oparciu o test Stone'a-Geissera (test S-G), który jest miarą dokładności predykcji dokonanej na podstawie modelu w stosunku do predykcji „trywialnej" oraz metodą tzw. cięcia Tuckey'a [Wold, 1980; Rogowski, 1990]. Są to metody specyficzne dla modelowania miękkiego. Test S-G, nazwany w ten sposób przez H. Wolda, nie jest typowym testem statystycznym, nie zawiera bowiem ani hipotez, ani statystyki, ani wartości krytycznej. Procedura badania istotności zmiennych jest zupełnie inna, stanowiąca pewnego rodzaju sprawdzian, a nie test statystyczny. Wartości testu S-G nie są ograniczone z dołu, zaś górne ograniczenie wynosi 1. Jeśli wartość testu jest ujemna, to model miękki ma gorszą własność predykcyjną w stosunku do prognozy trywialnej. W przypadku wartości testu równej 1 prognozy są bezbłędne, natomiast dla wartości równej 0 jakości prognozy z modelu i prognozy trywialnej są identyczne.

Do wyznaczenia wartości testu S-G należy L-krotnie oszacować parametry modelu miękkiego. Dla każdego parametru relacji zewnętrznej i wewnętrznej można wyznaczyć odchylenie standardowe uzyskanych L jego estymatorów: $b_1, b_2,...,b_L$ według wzoru:

$$s_\beta = \sqrt{\frac{\sum_{l=1}^{L}(b_l - \bar{b})^2}{L}} \tag{4}$$

gdzie $\bar{b} = \frac{\sum_{l=1}^{L} b_l}{L}$.

Wówczas ocenę istotności parametrów modelu przeprowadza się stosując regułę „2s". Metoda ta nazywana jest „metodą cięć Tukey'a".

2. Interpretacja wyników pomiaru atrakcyjności inwestycyjnej

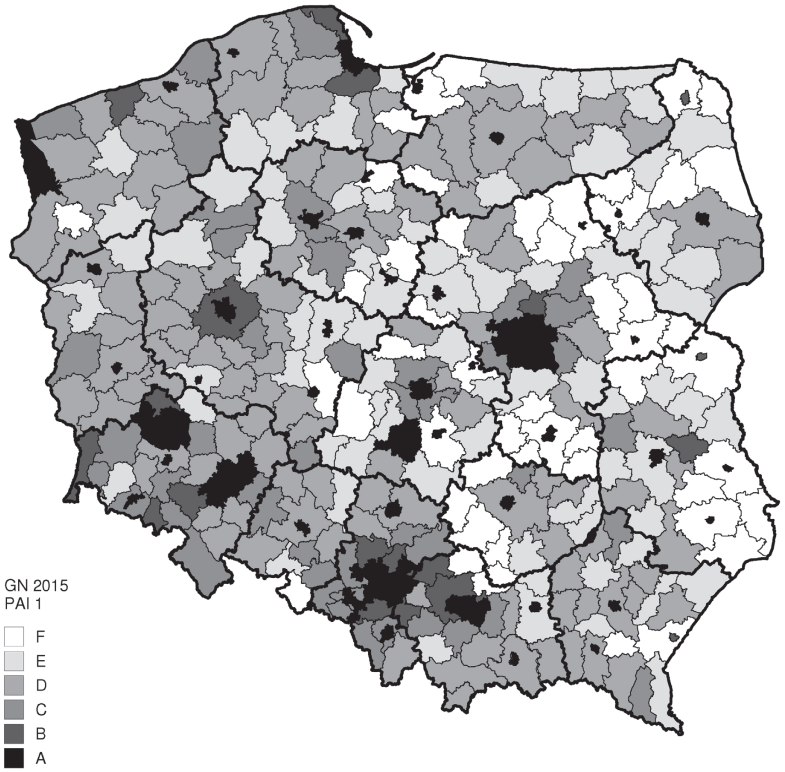
Zastosowanie metody wagowo-korelacyjnej pozwala na osiągnięcie licznych możliwości analitycznych w procesie zarządzania lokalizacją lub ich zbiorem. Dotyczy to monitorowania lokalizacji już wybranych, w których istnieje dane przedsiębiorstwo, w tym wielozakładowe o rozbudowanej strukturze przestrzennej. Umożliwia także rozpatrywanie przyszłych lokalizacji na tle istniejącego rozmieszczenia własnych firm, jak i konkurencji. Wynika to z faktu, że wskaźniki atrakcyjności są liczone wg tej samej metody dla wszystkich jednostek samorządu terytorialnego w Polsce, w podziale wg gmin, powiatów, podregionów, województw i makroregionów.

Utworzone na ich podstawie kartogramy pozwalają na ocenę atrakcyjności inwestycyjnej nie tylko konkretnych miejsc, gdzie jest zlokalizowana firma, ale także zbioru potencjalnych lokalizacji. Bardzo ważne jest określenie na podstawie analizy mapy atrakcyjności inwestycyjnej jednostek pierwszego rzędu sąsiedztwa, drugiego rzędu sąsiedztwa, czy w razie potrzeby dalszych jednostek. Tak samo jest możliwe dokonanie analizy porównawczej atrakcyjności inwestycyjnej wg innego klucza, np. położenia względem głównych szlaków komunikacyjnych czy kompleksów leśnych (ważne dla turystyki)[6].

Przykładowym opracowaniem kartograficznym jest prezentacja wskaźnika potencjalnej atrakcyjności inwestycyjnej dla polskich powiatów, w oparciu o dane za rok 2015 – por. rys. 2.

[6] Bardzo pomocne do tego celu są wskaźniki potencjalnej atrakcyjności inwestycyjnej, które są tworzone dla sekcji: hotele i gastronomia.

Rysunek 2. Atrakcyjność inwestycyjna powiatów w Polsce w roku 2015 na podstawie metody wagowo-korelacyjnej



Źródło: Opracowanie własne na podstawie obliczeń H. Godlewskiej-Majkowskiej.

Z zaprezentowanego rysunku wynika, że najbardziej atrakcyjne lokalizacje tworzą duże aglomeracje miejskie i miejsko-przemysłowe, przy czym im większy jest dany zespół miejski, tym większy obszar zajmuje koncentracja atrakcyjnych lokalizacji. Największą koncentrację przestrzenną atrakcyjnych gmin tworzy konurbacja górnośląska wraz z aglomeracją krakowską. Wyjątkiem jest aglomeracja łódzka, której oddziaływanie jako bieguna wzrostu jest ograniczone przestrzennie. Może mieć to związek ze specjalizacją przemysłu w dziedzinie tradycyjnych rodzajów przemysłu. Również aglomeracja warszawska obejmuje stosunkowo mały obszar wpływów. Może to wynikać z silnej specjalizacji usługowej w Warszawie i tym samym braku wystarczająco silnych mechanizmów pobudzających rozwój sektorów pokrewnych w strefie podmiejskiej. Warto także zauważyć, iż w województwie mazowieckim występują bardzo silne dysproporcje wewnątrzwojewódzkie pod względem poziomu rozwoju gospodarczego i tym samym atrakcyjności inwestycyjnej.

Na podstawie wskaźników można także dokonać ich uszeregowania od najbardziej atrakcyjnych do najmniej atrakcyjnych oraz wykonać profil atrakcyjności w oparciu o podstawowe grupy czynników lokalizacji, nazwane w tej metodzie mikroklimatami. W ten sposób można też wykorzystać wskaźniki i ich składowe do pogłębionej analizy mocnych i słabych stron, odnosząc cechy diagnostyczne do profilu obszaru referencyjnego lub określonych wielkości granicznych (np. za dopuszczalne lokalizacje można uznać te, które mają do 90% wielkości lidera w danym mikroklimacie).

Stosowanie tej metody w odniesieniu do przynajmniej kilkuletniego okresu pozwala na identyfikację jednostek, które podwyższają swoją atrakcyjność w porównaniu do innych, utrzymują wysoką pozycję, względnie ją tracą.

Wskaźniki potencjalnej atrakcyjności inwestycyjnej mają zastosowanie na etapie wstępnej selekcji zbioru potencjalnych lokalizacji inwestycji. Ocena syntetyczna wyrażona za pomocą wskaźnika PAI 1 pozwala nie tylko na ocenę ogólną miejsca, w którym może zostać przeprowadzona dana inwestycja. Oceny otoczenia danego miejsca ukazują także dostęp do zasobów oraz rynków położonych w sąsiedztwie. Konstrukcja wskaźników pozwala też na ich dekompozycję do poziomu subagregatów, dzięki czemu każdą jednostkę terytorialną można ocenić pod względem atrakcyjności walorów uwzględnionych w poszczególnych mikroklimatach (zasobów pracy, infrastruktury technicznej, infrastruktury społecznej, rynku zbytu oraz administracji). Dowolna jednostka może mieć zatem stworzony profil, ukazujący jej silne i słabe strony. Porównywalność metodyczna ocen dla wszystkich poziomów podziału statystycznego umożliwia także dokonywanie ocen dowolnych agregatów jednostek. W razie potrzeby można sięgnąć głębiej do profilu danej jednostki oraz określić mocne i słabe strony na poziomie zmiennych źródłowych. Daje to bardzo duże możliwości analizy dowolnego zbioru dopuszczalnych lokalizacji, rozpatrywanych pod względem różnego rodzaju filtrów, np. w postaci zakresu objęcia danej jednostki miejscowym planem zagospodarowania przestrzennego.

Tego typu analizy są przydatne dla obu stron regionalnego rynku inwestycyjnego, zarówno przedsiębiorstw poszukujących siedziby, jak i władz samorządowych przygotowujących ofertę terenów inwestycyjnych.

Przedsiębiorca otrzymuje informację zbiorczą, w oparciu o gotowe wskaźniki PAI ogólnie dla gospodarki narodowej lub dla poszczególnych rodzajów usług lub dla działalności produkcyjnej. Ponadto może otrzymać informację sprofilowaną, stosownie do znaczenia poszczególnych walorów lokalizacyjnych w danym konkretnym przypadku. Z kolei władze jednostki samorządu terytorialnego otrzymują informacje o mocnych i słabych stronach ofert terenów inwestycyjnych rozpatrywanych na tle konkurencji, na

dowolnym poziomie analizy lokalizacji ogólnej (gmina, powiat, podregion, województwa, makroregion).

Wskaźniki potencjalnej atrakcyjności inwestycyjnej można analizować w czasie pod warunkiem ujednolicenia ich metodyki. W ten sposób można określić, czy zmienia się, i w jakim kierunku, atrakcyjność inwestycyjna danego miejsca, jak również można określić bardzo konkretnie rodzaj zmian i ich wielkość w oparciu o porównanie ze zmianami w dowolnie określonym obszarze referencyjnym (powiat, województwo, kraj, Unia Europejska). Dlatego analiza wskaźników potencjalnej atrakcyjności inwestycyjnej ma bardzo duże walory praktyczne.

Podobne efekty można uzyskać, stosując metodę modelowania miękkiego. W wyniku budowy takiego modelu otrzymuje się oszacowania relacji wewnętrznej i zewnętrznej. Oszacowania modelu wewnętrznego mają dwie zalety. Pokazują wpływ poszczególnych mikroklimatów na atrakcyjność inwestycyjną, a także umożliwiają tworzenie rankingów i podział na klasy wszystkich badanych obiektów.

Oszacowania modelu wewnętrznego atrakcyjności inwestycyjnej powiatów w Polsce kształtują się w następujący sposób:

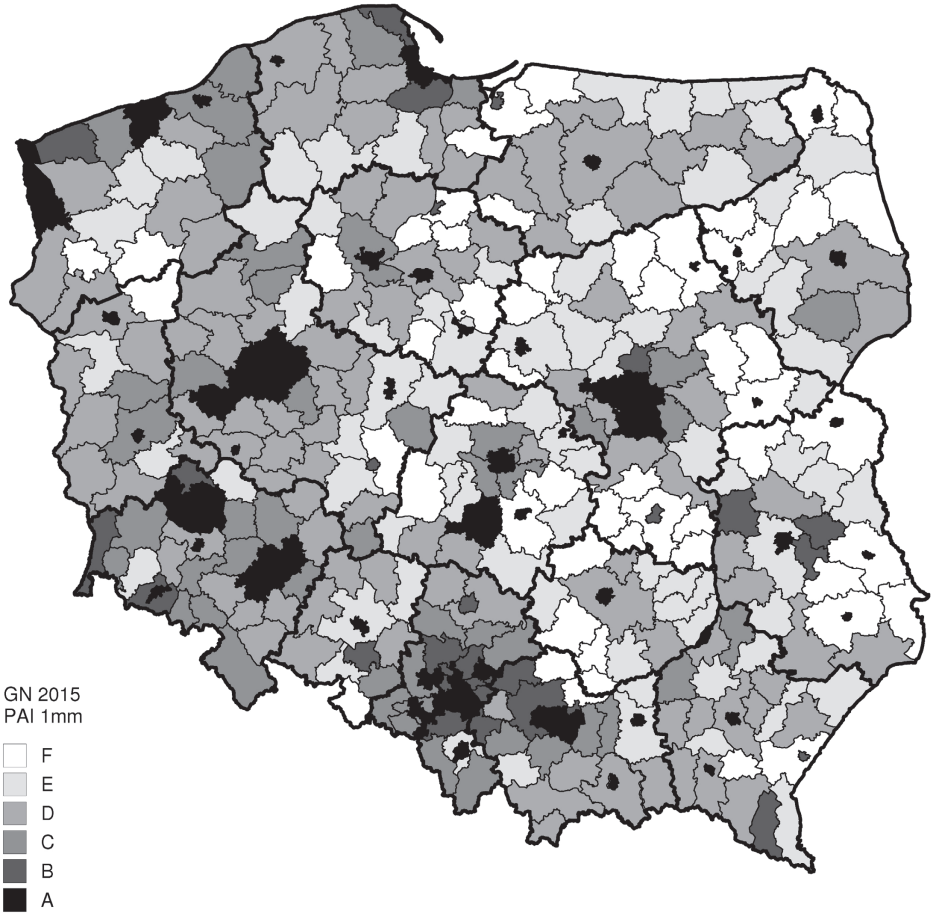
$$\begin{array}{lllllll} \mathbf{P\hat{A}I} = & 0{,}1136\,\mathbf{MZP} + & 0{,}2911\,\mathbf{MIT} + & 0{,}1640\,\mathbf{MIS} + & 0{,}2460\,\mathbf{MRY} + & 0{,}3107\,\mathbf{MAD} + & 1{,}4186 \quad \mathbf{R^2} = 0{,}9491 \\ & (0{,}0402) & (0{,}0207) & (0{,}0026) & (0{,}0211) & (0{,}0261) & (0{,}1350) \end{array} \tag{5}$$

Oznacza to, że wpływ analizowanych mikroklimatów na atrakcyjność inwestycyjną powiatów w Polsce jest bardzo podobny – dodatni i umiarkowany. Wszystkie oszacowane parametry (ładunki czynnikowe) są weryfikowane pozytywnie na podstawie testu cięć Tuckey'a. Największy wpływ na atrakcyjność inwestycyjną ma mikroklimat administracyjny (0,3107), mikroklimat infrastruktury technicznej (0,2911) i mikroklimat rynkowy (0,2460). Potwierdza to tezę o dużym wpływie sposobu zarządzania jednostką samorządu terytorialnego na jej potencjalną atrakcyjność inwestycyjną. Decyzje decydentów o realizacji na terenie JST konkretnych inwestycji (rozwój infrastruktury komunikacyjnej, uzbrojenie terenów itp.) w zakresie poprawy warunków lokalizacji inwestycji wpływa bezpośrednio na wzrost jej atrakcyjności. Jakość modelu jest bardzo wysoka. Kwadrat współczynnika korelacji wielorakiej oznacza, że w 94,91% zmienność atrakcyjności inwestycyjnej jest determinowana zmiennością badanych mikroklimatów. Model jest również weryfikowany pozytywnie na podstawie testu Stone'a-Geissera (ogólny test SG wynosi 0,5404), co wskazuje na pozytywną wartość prognostyczną otrzymanych oszacowań.

Oszacowania wewnętrzne umożliwiły dokonanie rankingu powiatów w Polsce pod względem atrakcyjności inwestycyjnej i ich podział na klasy

A-F, zgodnie z metodologią zastosowaną w analizie wyników uzyskanych metodą wagowo-korelacyjną. Najwyższe pozycje w rankingu, identycznie jak w przypadku metody wagowo-korelacyjnej, zajmują powiaty grodzkie, a także te, które skupione są wokół metropolii (rys. 3).

Rysunek 3. Atrakcyjność inwestycyjna powiatów w Polsce w roku 2015 na podstawie metody modelowania miękkiego

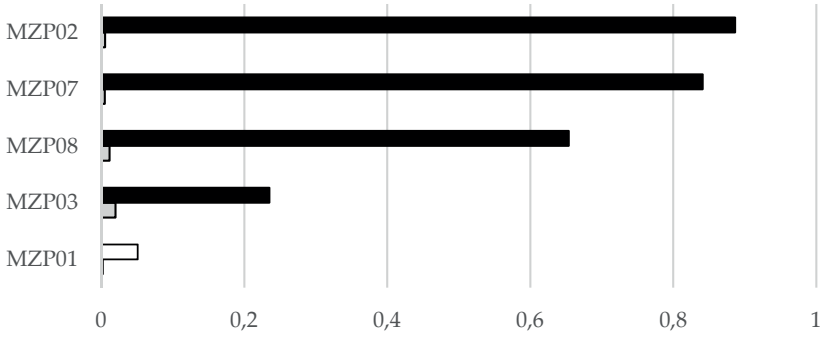


Źródło: Opracowanie własne na podstawie wyników modelowania miękkiego.

Oszacowania modelu zewnętrznego są również zgodne z oczekiwaniami. Na etapie estymacji i weryfikacji modelu zostały usunięte zmienne diagnostyczne, które były ocenione negatywnie. W ostatecznej wersji modelu pozostało 35 zmiennych diagnostycznych. Wartości bezwzględne ładunków czynnikowych i błędy ich szacunku zostały zaprezentowane na rysunkach 4–8.

Najwyższe ładunki czynnikowe występują w mikroklimacie infrastruktura techniczna. Widoczny jest bardzo wysoki (ok. 0,9 i więcej) dodatni wpływ na badany mikroklimat zmiennych diagnostycznych opisujących uzbrojenie techniczne w postaci sieci kanalizacyjnych, gazociągowych i wodociągowych. Z kolei mikroklimat rynkowy najsilniej odbijają dochody budżetów gmin z podatku PIT na mieszkańca (0,9336), co potwierdza ważny wpływ jakości życia (i jakości pracy) na wzrost atrakcyjności inwestycyjnej. Współcześnie w dobie powszechnego dostępu do Internetu i globalizacji wiele branż, np. związanych z szeroko pojętym rynkiem nowych technologii (IT), może rozwijać się praktycznie w każdym miejscu na świecie. Wyznacznikiem lokalizacji miejsca inwestycji jest wówczas np. jakość życia. Potwierdza to np. dynamiczny rozwój tego rynku w Białymstoku, który zajmuje bardzo wysokie pozycje w rankingach jakości życia, a relatywnie niskie pod względem rozwoju infrastrukturalnego. Ponadto wyższe dochody mieszkańców wpływają na wyższe możliwości dynamicznego wzrostu rynku endogenicznego, co również przekłada się na decyzje lokalizacyjne. Udział w podatku dochodowym od osób fizycznych stanowi często główne źródło dochodów własnych gmin. Zmienna ta (MA03) ma najwyższy wpływ na mikroklimat administracyjny (0,7935). Przekłada się to bezpośrednio na większe możliwości rozwojowe gmin w zakresie promocji gospodarczej, a także działalności inwestycyjnej gmin mającej na celu poprawę warunków lokalizacyjnych, poprzez rozwój badanych mikroklimatów.

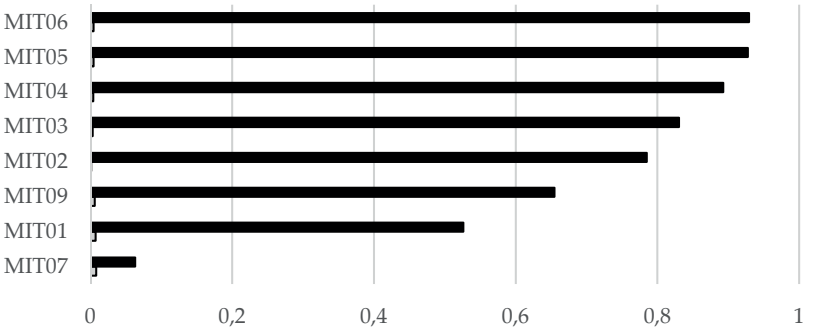
Rysunek 4. Wartości bezwzględne ładunków czynnikowych mikroklimatu zasoby pracy (MZP)*



	MZP01	MZP03	MZP08	MZP07	MZP02
MZP	0,051	0,2352	0,6539	0,8409	0,8862
Błąd	0,0021	0,0198	0,0118	0,0049	0,0054

* Kolor biały – oznacza destymulantę, a czarny – stymulantę.
Źródło: Opracowanie własne na podstawie wyników modelowania miękkiego.

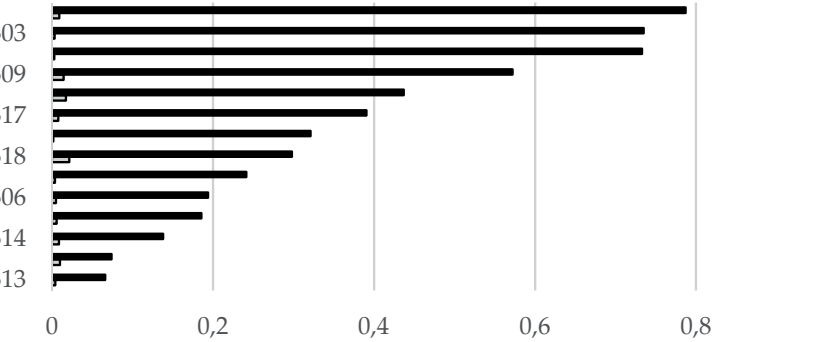
Rysunek 5. Wartości bezwzględne ładunków czynnikowych mikroklimatu infrastruktura techniczna (MIT)*



	MIT07	MIT01	MIT09	MIT02	MIT03	MIT04	MIT05	MIT06
MIT	0,0628	0,5259	0,6548	0,7849	0,8303	0,8927	0,9275	0,9292
Błąd	0,008	0,007	0,0057	0,001	0,0029	0,0037	0,0043	0,0044

* Kolor czarny oznacza stymulantę.
Źródło: Opracowanie własne na podstawie wyników modelowania miękkiego.

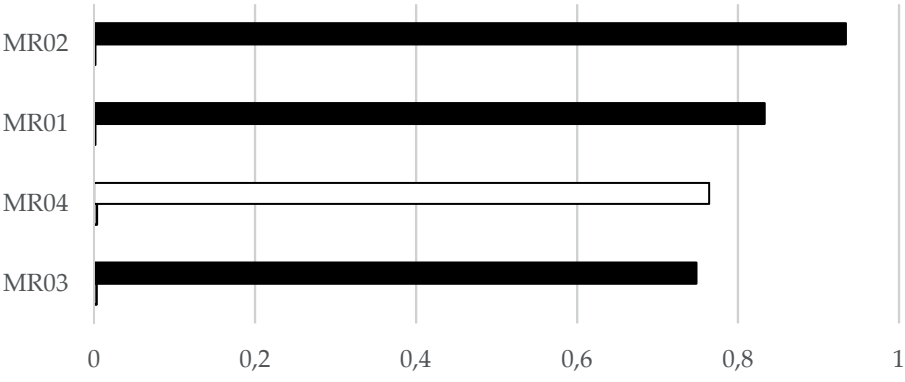
Rysunek 6. Wartości bezwzględne ładunków czynnikowych mikroklimatu infrastruktura społeczna (MIS)*



	MIS 13	MIS 04	MIS 14	MIS 19	MIS 06	MIS 15	MIS 18	MIS 05	MIS 17	MIS 16	MIS 09	MIS 02	MIS 03	MIS 11
MIS	0,066	0,074	0,138	0,185	0,194	0,241	0,298	0,321	0,39	0,437	0,572	0,733	0,735	0,787
Błąd	0,004	0,01	0,008	0,006	0,005	0,003	0,021	0,001	0,007	0,017	0,014	0,003	0,003	0,009

* Kolor czarny oznacza stymulantę.
Źródło: Opracowanie własne na podstawie wyników modelowania miękkiego.

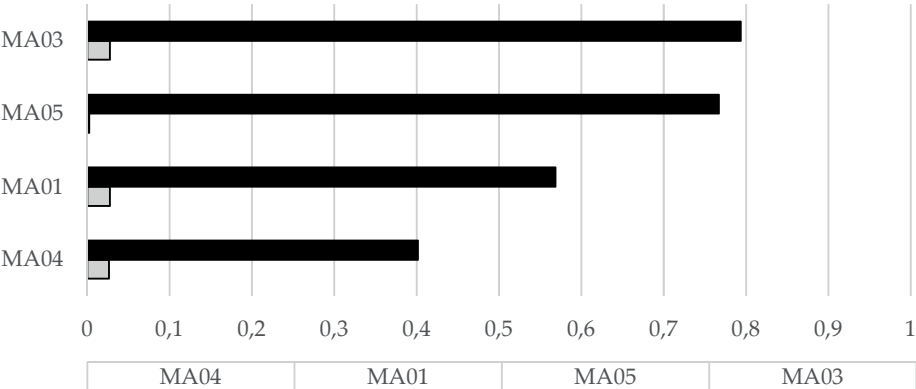
Rysunek 7. Wartości bezwzględne ładunków czynnikowych mikroklimatu rynkowego (MR)*



	MR03	MR04	MR01	MR02
MR	0,7483	0,7640	0,8331	0,9336
Błąd	0,0032	0,0035	0,0012	0,0015

* Kolor biały – oznacza destymulantę, a czarny – stymulantę.
Źródło: Opracowanie własne na podstawie wyników modelowania miękkiego.

Rysunek 8. Wartości bezwzględne ładunków czynnikowych mikroklimatu administracyjnego (MA)*



	MA04	MA01	MA05	MA03
MA	0,4016	0,5685	0,767	0,7935
Błąd	0,0262	0,0274	0,0023	0,0276

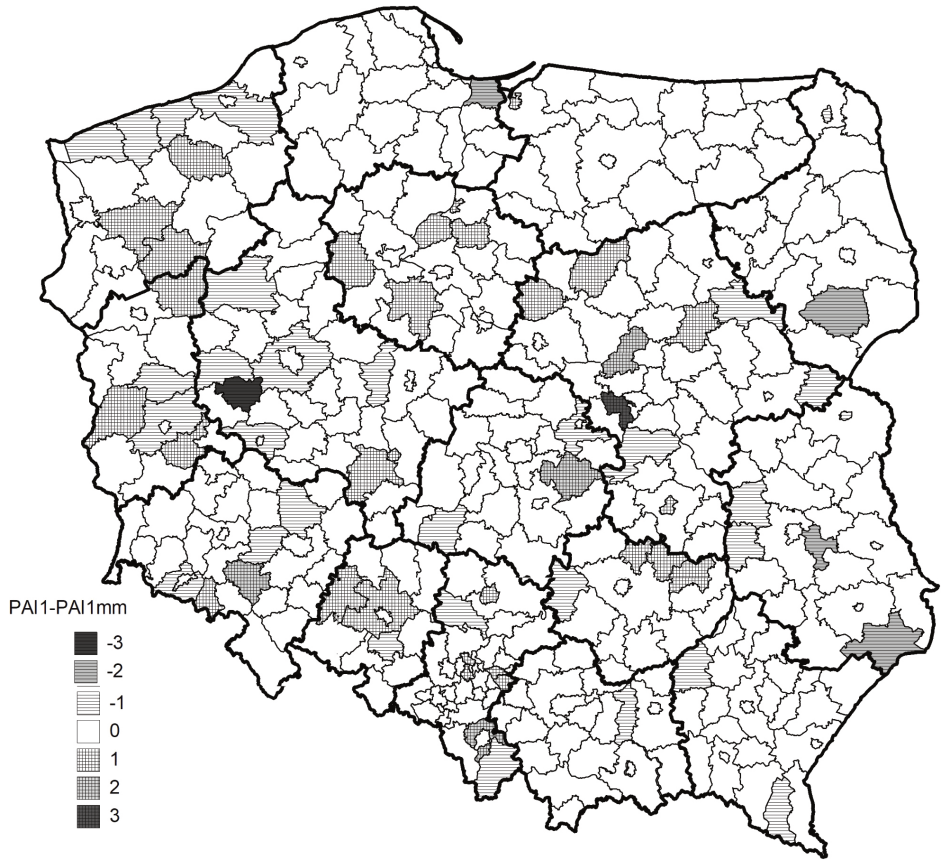
* Kolor czarny oznacza stymulantę.
Źródło: Opracowanie własne na podstawie wyników modelowania miękkiego.

Modelowanie miękkie, w porównaniu do metody wagowo-korelacyjnej, daje dodatkowe możliwości analityczne. Oszacowania modelu zewnętrznego pozwalają na określenie, które ze zmiennych diagnostycznych najsilniej, a które najsłabiej wpływają na kształtowanie się konkretnego mikroklimatu. Model miękki umożliwia również pełną weryfikację, zarówno merytoryczną, jak i statystyczną. W związku z tym modelowanie miękkie może być traktowane jako źródło walidacji zmiennych diagnostycznych w analizach wykonywanych metodą wagowo-korelacyjną. Z kolei model wewnętrzny, poza możliwością uporządkowania obiektów i podziałem ich na klasy, pokazuje wpływ badanych mikroklimatów na atrakcyjność inwestycyjną. Informacje te mogą być źródłem pomocnym przy określaniu wag w pierwszej z analizowanych metod. Wadą modelowania miękkiego, w porównaniu do metody wagowo-korelacyjnej, jest wyeliminowanie z analiz, często bardzo ważnych z merytorycznego punktu widzenia, zmiennych diagnostycznych (np. saldo migracji zewnętrznych), z powodu ich negatywnej weryfikacji statystycznej. Ważniejsze są w tym przypadku testy statystyczne niż kryteria merytoryczne.

Wyniki uporządkowania powiatów w Polsce pod względem atrakcyjności inwestycyjnej uzyskane metodą wagowo-korelacyjną i modelowania miękkiego są zbieżne. Współczynnik korelacji rang Spearmana (r_s) kształtuje się na bardzo wysokim poziomie 0,9923. Nieco niższe zależności widoczne są w wynikach dotyczących poszczególnych mikroklimatów. Najbardziej podobne uporządkowania, na podstawie współczynnika r_s, dotyczą mikroklimatów: rynkowego (0,9932), infrastruktury technicznej (0,9601) i administracyjnej (0,9071). Najmniej podobne są pod względem dwóch mikroklimatów: infrastruktury społecznej (0,6523) i zasobów pracy (0,4845). Największe różnice wynikają z eliminacji dwunastu zmiennych diagnostycznych z modelu miękkiego, z uwagi na ich negatywną weryfikację. Dotyczyło to trzech zmiennych z mikroklimatu zasoby pracy, sześciu z mikroklimatu infrastruktura społeczna, a także po jednej zmiennej z mikroklimatów: MIT, MR i MA.

Analiza porównawcza kartogramów ilustrujących wyniki pomiaru atrakcyjności inwestycyjnej wskazuje na dużą zbieżność wyników w skali większości województw (rys. 9).

Rysunek 9. Atrakcyjność inwestycyjna powiatów w Polsce w roku 2015 – w ujęciu porównawczym



Źródło: Opracowanie własne na podstawie wyników obliczeń autorów.

Na 379 wszystkich ocenianych powiatów 311, czyli 82%, uzyskało tę samą klasę atrakcyjności inwestycyjnej. Po 29 powiatów, czyli po 8%, różniło się o +/– jedną klasę, a po 4, czyli po 1%, powiatów uzyskało oceny o dwie wyższe lub niższe. Tylko jeden powiat – grodziski w województwie wielkopolskim – uzyskał ocenę rozbieżną o 3 klasy. Niedoszacowana w oparciu o metodę wagowo-korelacyjną wydaje się być atrakcyjność inwestycyjna terenów przemysłowych wokół Poznania oraz leżących na szlakach komunikacyjnych łączących Poznań z granicą zachodnią Polski, jak również obszary turystyczne Pomorza zachodniego. Natomiast metoda ta ukazuje w większym stopniu niż metoda modelowania miękkiego atrakcyjność ośrodków przemysłowych Dolnego Śląska, np. Opola, Kamiennej Góry, Świdnicy i Brzegu, jak również Mazowsza, zwłaszcza Grodziska Mazowieckiego, Mławy, a także Białego Zagłębia. Jednak są to różnice dotyczące co najwyżej kilku powiatów w każdym województwie.

Zakończenie

Celem artykułu była analiza porównawcza wyników badania atrakcyjności inwestycyjnej powiatów w Polsce wykonanego za pomocą dwóch metod badawczych: wagowo-korelacyjnej i modelowania miękkiego. Postawiono hipotezę, że zastosowanie różnych sposobów wyznaczania miernika syntetycznego nie ma istotnego wpływu na wyniki klasyfikacji powiatów w Polsce pod względem atrakcyjności inwestycyjnej, a stanowi komplementarne źródło wnioskowania umożliwiające osiągnięcie efektu synergii. Hipotezę weryfikowano poprzez zastosowanie metody wagowo-korelacyjnej i modelowania miękkiego do analizy tego samego zestawu 47 zmiennych diagnostycznych, dobranego na podstawie pierwszej z wykorzystanych metod. Dodatkowo specyfikacji modelu miękkiego w zakresie relacji wewnętrznych dokonano na bazie mikroklimatów określonych metodą wagowo-korelacyjną. Uzupełnieniem wykonanych badań była analiza kartograficzna. Warto podkreślić, że w przypadku modelowania miękkiego nie cały zbiór 47 zmiennych diagnostycznych został zweryfikowany pozytywnie. Oznacza to, że metoda ta może stanowić uzupełniające narzędzie oceny przydatności poszczególnych zmiennych diagnostycznych do analizy atrakcyjności inwestycyjnej powiatów. Wskazuje bowiem, które mikroklimaty najsilniej odbijają zmienną endogeniczną w postaci wskaźnika syntetycznego „potencjalna atrakcyjność inwestycyjna", a także, które zmienne diagnostyczne tworzące określone „mikroklimaty" odznaczają się najwyższym ładunkiem czynnikowym.

W artykule wskazano zalety i wady zastosowanych sposobów wyznaczania miernika syntetycznego. Zaletą wskaźnika skonstruowanego metodą wagowo-korelacyjną jest pełna porównywalność ocen atrakcyjności inwestycyjnej na wszystkich szczeblach podziału statystycznego kraju, a także wykorzystanie specjalistycznej wiedzy eksperckiej przy konstruowaniu poziomu wag. Wadą natomiast – podobnie, jak w przypadku innych narzędzi badawczych stosowanych do konstrukcji wskaźników na bazie danych dostępnych na podstawowym poziomie JST – zawężenie zbioru zmiennych diagnostycznych do tych, które są dostępne na najniższym szczeblu – gminy. Zaletą modelowania miękkiego jest możliwość wskazania, który z mikroklimatów ma najwyższy wpływ na potencjalną atrakcyjność inwestycyjną oraz możliwość merytorycznej i statystycznej weryfikacji wyników. Jednak z drugiej strony, weryfikacja statystyczna wymaga często wyeliminowania zmiennych bardzo istotnych z punktu widzenia merytorycznego – wynikającego m.in. z zaangażowania wiedzy eksperckiej. Taka eliminacja zmiennych diagnostycznych z przyczyn statystycznych nie przekłada się na rzeczywistą sytuację gospodarczą, w której większość badanych zjawisk odznacza się zależnością również

od czynników niemierzalnych. Można to określić jako wadę tego modelu. Ograniczeniem jednej i drugiej metody są ponadto relacje liniowe.

Zastosowanie obydwu sposobów wyznaczania miernika syntetycznego wskazało na identyczne, najbardziej atrakcyjne lokalizacje, występujące w dużych aglomeracjach miejskich i miejsko-przemysłowych. Uzyskane wyniki są bardzo podobne. Współczynnik korelacji rang Spearmana między uporządkowaniem powiatów wg atrakcyjności inwestycyjnej uzyskanej metodą wagowo-korelacyjną i modelowania miękkiego wskazuje na silną zależność miedzy nimi na poziomie 0,9923, co było przewidywalne z uwagi na zbliżony zestaw zmiennych diagnostycznych . Różnice w uzyskanych rankingach przeanalizowano na podstawie metody kartograficznej. Analizując różnice w ostatecznym składzie zmiennych diagnostycznych, poziomy wag w metodzie wagowo-korelacyjnej, poziomy ładunków czynnikowych w metodzie modelowania miękkiego, można wyciągać bardzo szczegółowe wnioski, dotyczące konkretnej jednostki samorządu terytorialnego, co przekłada się na jej duże znaczenie aplikacyjne i jest zaletą przeprowadzonej analizy porównawczej. Dotyczy to w szczególności czynników lokalizacji inwestycji, które są bardzo zróżnicowane i trudne do całkowitego zidentyfikowania. Włodarze jednostek samorządu terytorialnego powinni odznaczać się proinwestycyjnym i innowacyjnym myśleniem, warunkującym dynamiczny rozwój ich lokalnych gospodarek w burzliwym i konkurencyjnym otoczeniu. Kluczowe znaczenie w zakresie przyciągania inwestorów mają nadal warunki związane z dostępnością transportową, infrastrukturą techniczną, zasobami i kosztami pracy, cenami nieruchomości, położeniem geograficznym. Jednak coraz częściej istotną rolę odgrywa wizerunek JST, klimat do rozwoju przedsiębiorczości, aktywność ośrodków naukowo-badawczych, poziom obsługi inwestorów przez instytucje publiczne, jakość życia, jakość środowiska, profesjonalizm firm doradczych i konsultingowych. Już sam proces pobudzania atrakcyjności inwestycyjnej jest zjawiskiem korzystnym, ponieważ wymusza szereg pozytywnych zmian, przyczyniając się do rozwoju społeczno-gospodarczego jednostek samorządu terytorialnego.

Literatura

Atrakcyjność inwestycyjna i potencjał gospodarczy gmin województwa śląskiego. Raport końcowy (październik 2009), Urząd Marszałkowski Województwa Śląskiego, Katowice.

Borys T., Walesiak M. (red.) (1983), *Statystyczna analiza porównawcza*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.

Durka B. (2000), *Udział podmiotów z kapitałem zagranicznym w polskim handlu za-granicznym*, w: B. Durka (red.), *Inwestycje zagraniczne w Polsce*, Instytut Koniunktur i Cen Handlu Zagranicznego, Warszawa.

EBA Investment Attractiveness Index, http://www.eba.com.ua/en/press-and-media/press-room/indices/investment-attractiveness-index, dostęp: 25.09.2017.

Godlewska-Majkowska H. (red.) (2012), *Atrakcyjność inwestycyjna regionów Polski jako źródło przedsiębiorczych przewag konkurencyjnych*, Oficyna Wydawnicza SGH, Warszawa.

Godlewska-Majkowska H. (red.) (2010), *Innowacyjność jako czynnik wzrostu atrakcyjności inwestycyjnej polskich regionów w latach 2002–2007*, Studia i Analizy Instytutu Przedsiębiorstwa, Oficyna Wydawnicza SGH, Warszawa.

Godlewska-Majkowska H. (red.) (2011), *Atrakcyjność inwestycyjna a przedsiębiorczość regionalna w Polsce*, Studia i Analizy Instytutu Przedsiębiorstwa, Oficyna Wydawnicza SGH, Warszawa.

Godlewska-Majkowska H. (red.) (2012), *Atrakcyjność inwestycyjna regionów Polski jako źródło przedsiębiorczych przewag konkurencyjnych*, Studia i Analizy Instytutu Przedsiębiorstwa, Oficyna Wydawnicza SGH, Warszawa.

Grabiński T., Wydymus S., Zeliaś A. (1983), *Metody prognozowania rozwoju społeczno-gospodarczego*, PWE, Warszawa.

Guzik W. (2008), *Atrakcyjność inwestycyjna województw*, „Gospodarka Narodowa", nr 3.

Joreskog K.G., Wold H. (1982), *Systems under indirect observation. Causality – Structure – Prediction*, Part II, North-Holland.

Kalinowski T. (red.), *Atrakcyjność inwestycyjna województw i podregionów Polski 2006*, IBnGR, Gdańsk.

Kock N., Mayfield M. (2015), *PLS-based SEM Algorithms: The Good Neighbour Assumption*, Collinearity, and Nonlinearity, Texas A&M International University, USA, „Information Management and Business Review", Vol. 7, No. 2.

Kolenda M., (2006), *Taksonomia numeryczna: klasyfikacja, porządkowanie i analiza obiektów wielocechowych*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.

Kuszewski T. (2000), *Wprowadzenie do modelowania ekonometrycznego*, w: M. Gruszczyński, M. Podgórska (red.), *Ekonometria*, Wydawnictwo SGH, Warszawa.

Lee Ch.-S., Chen Y.-Ch., Tsui P.-L., Yu T.-H. (2013), *Examining the relations between open innovation climate and job satisfaction with a PLS path model*, „Quality & Quantity", Vol. 48, No. 3.

Migdał-Najman K., Najman K. (2013), *Analiza porównawcza wybranych metod analizy skupień w grupowaniu jednostek o złożonej strukturze grupowej*, „Zarządzanie i Finanse", nr 3, cz. 2.

Ostasiewicz W. (red.) (1989), *Statystyczne metody analizy danych*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.

Perło D. (2014), *Modelowanie zrównoważonego rozwoju regionów*, Trans Humana, Białystok.

Poniatowska-Jaksch M. (2006), *Przemysłowe bezpośrednie inwestycje zagraniczne źródłem konkurencyjności regionu*, „Monografie i Opracowania", nr 544, Oficyna Wydawnicza SGH, Warszawa.

Puciato D., Dziedzic E. (2016), *Attractiveness of South-West Poland municipalities for independent hotel investors*, „Tourism Economics", Vol. 23, No. 3.

Rogowski J. (1990), *Modele miękkie. Teoria i zastosowanie w badaniach ekonomicznych*, Wydawnictwo Filii UW w Białymstoku, Białystok.

Rogowski J. (2002), *Modele miękkie w budowie strategii finansowej regionu*, „Optimum. Studia Ekonomiczne", nr 1.

Servera-Francés D., Arteaga-Moreno F., Gil-Saura I., Gallarza M.G. (2012), *A multiblock PLS-based algorithm applied to a causal model in marketing*, „Applied Stochastic Models in Business and Industry", Vol. 29, No. 3.

Tarkowski M. (red.) (2015), *Atrakcyjność inwestycyjna województw i podregionów Polski 2015*, Instytut Badań nad Gospodarką Rynkową przy współpracy z Fundacją Konrada Adenauera, Gdańsk.

Tenenhaus M., Esposito Vinzi V., Chatelin Y.-M., Lauro C. (2005), *PLS path modeling*, „Computational Statistics & Data Analysis", Vol. 48, No. 1.

Wold H. (1980) *Soft Modelling: Intermediate between Traditional Model Building and Data Analysis*, Banach Centre Publication 6, Mathematical Statistics.

Wold H. (1982), *Soft modeling: The basic design and some extensions*, w: K.G. Jöreskog, S. Wold (eds.), *Systems under indirect observation: causality, structure, prediction*, Amsterdam.

Zaucha J., Brodzicki T., Ciołek D., Komornicki T., Mogiła Z., Szlachta J., Zaleski J. (2015), *Terytorialny wymiar wzrostu i rozwoju*, Difin, Warszawa.

Aneks

Tablica A. Składowe wskaźnika potencjalnej atrakcyjności inwestycyjnej (PAI1_ GN) dla gospodarki narodowej

Symbol	Wyszczególnienie	Charakter zmiennej (stymulanta: S, destymulanta: D)	Przelicznik
MZP	MIKROKLIMAT ZASOBY PRACY		
MZP01	Odsetek ludności w wieku nieprodukcyjnym na 100 os. w wieku produkcyjnym	D	1
MZP02	Wskaźnik aktywności zawodowej – liczba osób pracujących w przeliczeniu na 100 os. w wieku produkcyjnym	S	1
MZP03	Saldo migracji stałej wewnętrznej na 1000 mieszkańców	S	1
MZP04	Saldo migracji zagranicznej na 1000 mieszkańców	S	1
MZP05	Ludność w wieku poprodukcyjnym na 100 osób w wieku przedprodukcyjnym	D	1
MZP06	Udział % ludności w wieku produkcyjnym	S	1

Symbol	Wyszczególnienie	Charakter zmiennej (stymulanta: S, destymulanta: D)	Przelicznik
MZP07	Wydatki na oświatę i wychowanie w zł/mieszk.	S	1
MZP08	Wydatki na kulturę i ochronę dziedzictwa narodowego w zł/ mieszk.	S	1
MIT	MIKROKLIMAT INFRASTRUKTURA TECHNICZNA		
MIT01	Udział % ludności obsługiwanej przez wodociąg	S	1
MIT02	Udział % mieszkań z podłączeniem do gazociągu	S	1
MIT03	Udział % ludności obsługiwanej przez kanalizację	S	1
MIT04	Gęstość sieci wodociągowej w km na 100 km^2	S	0,333
MIT05	Gęstość sieci gazociągowej w km na 100 km^2	S	0,333
MIT06	Gęstość sieci kanalizacyjnej w km na 100 km^2	S	0,333
MIT07	Udział % odpadów wytworzonych w ciągu roku i unieszkodliwionych do odpadów wytworzonych w ciągu roku	S	1
MIT08	Udział % ścieków oczyszczonych w ściekach wymagających oczyszczenia	S	1
MIT09	Wydatki na transport i łączność w zł/mieszk.	S	1
MIS	MIKROKLIMAT INFRASTRUKTURA SPOŁECZNA		
MIS01	Praktyki lekarskie na wsi i w mieście na 100 tys. mieszkańców	S	1
MIS02	Liczba zakładów opieki zdrowotnej ogółem na 100 tys. mieszkańców	S	1
MIS03	Liczba aptek na 100 tys. mieszkańców	S	1
MIS04	Powierzchnia użytkowa mieszkań w m^2 *per capita*	S	1
MIS05	Liczba komputerów podłączonych do internetu do ogółu komputerów w szkołach podstawowych (%)	S	1

Symbol	Wyszczególnienie	Charakter zmiennej (stymulanta: S, destymulanta: D)	Przelicznik
MIS06	Liczba komputerów podłączonych do internetu do ogółu komputerów w gimnazjach (%)	S	1
MIS07	Liczba uczniów na komputer z dostępem do internetu w szkołach podstawowych	D	1
MIS08	Liczba uczniów na komputer z dostępem do internetu w gimnazjach	D	1
MIS09	Wypożyczenie księgozbioru na zewnątrz na 1000 mieszkańców	S	1
MIS10	Liczba mieszkańców przypadających na 1 kino stałe	D	1
MIS11	Liczba widzów w kinach stałych na 100 mieszkańców	S	1
MIS12	Kubatura nowych budynków mieszkalnych w m^3 na 100 mieszkańców	S	1
MIS13	Liczba mieszkańców przypadających na 1 muzeum z oddziałami	d	1
MIS14	Liczba zwiedzających muzea z oddziałami na 1000 mieszkańców	S	1
MIS15	Hale sportowe o wymiarach od 36x19 do 44x22 m oraz 44x22 m i większe na 1000 mieszkańców	S	1
MIS16	Korty tenisowe otwarte i kryte na 1000 mieszkańców	S	1
MIS17	Pływalnie otwarte i kryte na 1000 mieszkańców	S	1
MIS18	Aquaparki na 1000 mieszkańców	S	1
MIS19	Skateparki na 1000 mieszkańców	S	1
MIS20	Długość ścieżek rowerowych na 1000 mieszkańców	S	1
MR	MIKROKLIMAT RYNKOWY		
MR01	Gęstość zaludnienia (km^2 na os.)	S	1
MR02	Dochody budżetów gmin z podatku PIT na mieszkańca (zł)	S	1
MR03	Dochody budżetów gmin z podatku CIT na tysiąc pracujących (zł)	S	1

Symbol	Wyszczególnienie	Charakter zmiennej (stymulanta: S, destymulanta: D)	Przelicznik
MR04	Udział % wpływów z podatku rolnego w dochodach podatkowych	D	1
MR05	Udział % wydatków na pomoc społeczną i pozostałe zadania w zakresie polityki społecznej w wydatkach gmin	D	1
MA	MIKROKLIMAT ADMINISTRACJA		
MA01	Powierzchnia objęta planem zagospodarowania przestrzennego odniesiona do powierzchni gminy (%)	S	1
MA02	Środki na dofinansowanie własnych zadań pozyskane z innych źródeł w zł na 1 mieszkańca	S	1
MA03	Udział % dochodów własnych w dochodach ogółem	S	1
MA04	Wydatki majątkowe ogółem na wydatki bieżące ogółem (%)	S	1
MA05	Udział % wydatków na gospodarkę komunalną i ochronę środowiska, kulturę i ochronę dziedzictwa narodowego, bezpieczeństwo publiczne i ochronę przeciwpożarową	S	1

Źródło: Opracowanie własne.

Streszczenie

Celem artykułu jest analiza porównawcza wyników badania atrakcyjności inwestycyjnej powiatów w Polsce wykonanego za pomocą dwóch metod badawczych: wagowo-korelacyjnej i modelowania miękkiego. Postawiono hipotezę, że zastosowanie różnych sposobów wyznaczania miernika syntetycznego daje w przybliżeniu takie same wyniki klasyfikacji powiatów w Polsce pod względem atrakcyjności inwestycyjnej. Analizę porównawczą przeprowadzono na podstawie danych statystycznych z 2015 r., zagregowanych na poziomie 379 powiatów ziemskich i grodzkich. Artykuł składa się z dwóch kluczowych części: opisu metod badawczych ze wskazaniem ich zalet i wad, a także interpretacji otrzymanych wyników. Kluczowe znaczenie poznawcze posiada metoda wagowo-korelacyjna, na podstawie której zostały wyspecyfikowane zmienne diagnostyczne. Atrakcyjność inwestycyjną rozpatrywaną łącznie dla całej gospodarki regionalnej powinno opisywać możliwie dużo zmiennych diagnostycznych, odnoszących się do szerokiego spektrum rozwoju społeczno-gospodarczego. Czynniki te znajdują

odzwierciedlenie w doborze 47 zmiennych, składających się na wskaźnik potencjalnej atrakcyjności inwestycyjnej regionów dla inwestycji w gospodarkę narodową (PAI2_GN). Składa się on z mikroklimatów: „zasoby pracy", „infrastruktura techniczna", „infrastruktura społeczna", „rynkowy", „administracja". Badania i analizy przeprowadzone za pomocą metody wagowo-korelacyjnej stanowiły podstawę do konstrukcji modelu miękkiego. Zastosowanie badanych metod pozwala na osiągnięcie licznych możliwości analitycznych w procesie zarządzania lokalizacją lub ich zbiorem. Dotyczy to monitorowania lokalizacji już wybranych, w których istnieje dane przedsiębiorstwo, w tym wielozakładowe o rozbudowanej strukturze przestrzennej. Umożliwia także rozpatrywanie przyszłych lokalizacji na tle istniejącego rozmieszczenia własnych firm, jak i konkurencji.

Słowa kluczowe
atrakcyjność inwestycyjna, metoda wagowo-korelacyjna, modelowanie miękkie, powiaty

Implementation of weight-correlative method and soft modeling in investment attractiveness analysis of districts in Poland

The aim of the article is a comparative analysis of the research results of Polish districts' investment attractiveness conducted by means of two research methods: weight-correlative and soft modeling. It has been assumed that the implementation of various methods for determining the synthetic measure gives approximately the same results of the Polish districts' classification in terms of investment attractiveness. Comparative analysis was carried out on the basis of statistical data of 2015, aggregated at the level of 379 urban and rural districts. The article is composed of two key parts: description of research methods together with their pros and cons as well as interpretation of the obtained results. The key cognitive significance was given to the weight-correlative method, on the basis of which the diagnostic changes were specified. Investment attractiveness, considered jointly for the whole regional economy, should be described by a possibly vast number of diagnostic variables, pertaining to a wide spectrum of social and economic development. These factors are reflected in the selection of 47 variables, comprising the index of potential investment attractiveness of regions in terms of investment in the national economy (PAI2_GN). It is composed of microclimates: 'labour resources', 'technical infrastructure', 'social infrastructure', 'market', 'administration'. The research and analysis conducted with the use of the weight-correlative method constituted the basis for the construction of soft model. The implementation of the studied methods allows to achieve numerous analytical possibilities in the process of location management and their collection. It applies to monitoring the already selected locations, where the given enterprises exist, including the multi-plant ones with developed spatial structure. It also enables the consideration on the future locations against the background of the existing arrangement of own companies as well as the competition.

Keywords
investment attractiveness, weight-correlative method, soft modelling, districts

Aleksandra Łuczak*

Modelowanie rozwoju lokalnego z wykorzystaniem opinii wybranych grup ekspertów

Wstęp

Modelowanie rozwoju jest jednym z głównych etapów planowania rozwoju jednostek administracyjnych. Można je prowadzić zarówno na poziomie regionalnym (województwa), jak i lokalnym (powiaty, gminy). Na modelowanie rozwoju w skali regionalnej istotny wpływ ma modelowanie na szczeblu lokalnym. W tym procesie budowany jest model, który umożliwia właściwe planowanie rozwoju i związane z tym wydatkowanie środków finansowych.

Celem badań jest zbudowanie modelu rozwojowego dla gminy z wykorzystaniem opinii ekspertów i rozmytego analitycznego procesu hierarchicznego (FAHP). W badaniach postawiono hipotezę, ściśle związaną z celem badania, która sprowadza się do stwierdzenia, że FAHP może być pomocnym narzędziem w modelowaniu rozwoju lokalnego. Konstruowany model obejmuje najistotniejsze czynniki strategiczne: cele, zadania i scenariusze rozwoju. Czynniki strategiczne zostają porównane parami pod względem ich ważności w procesie rozwoju lokalnego. Porównań dokonują wybrane grupy ekspertów (np. radni gminy, radni powiatu). Dla każdej z grup ekspertów agreguje się ich indywidualne oceny. Stanowią one podstawę do obliczenia cząstkowych ocen ważności czynników strategicznych (wartości priorytetów) za pomocą FAHP. W efekcie zostają utworzone cząstkowe modele, które są podstawą do opracowania zagregowanego modelu rozwoju gminy. Badania przeprowadzono na danych pochodzących z badań ankietowych przeprowadzonych wśród wybranych grup ekspertów dla gminy wiejskiej Chocz w perspektywie do 2020 r. Grupy ekspertów stanowili radni gminy i radni powiatu. W badaniach wykorzystano też opinie wójta gminy i starosty powiatu.

1. Metodyka badań

Planowanie rozwoju lokalnego związane jest z rozwiązywaniem wielu zagadnień decyzyjnych. Mają one charakter złożony i wielokryterialny. Wśród tych zagadnień znajduje się opracowanie strategii. Opracowanie dokumentu strategii jest pomocne w procesie pozyskiwania dodatkowych

* Dr hab., Katedra Finansów i Rachunkowości, Wydział Ekonomiczno-Społeczny, Uniwersytet Przyrodniczy w Poznaniu, ul. Wojska Polskiego 28, 60-637 Poznań, aleksandra.luczak@up.poznan.pl

zewnętrznych źródeł finansowania, a także stanowi wymóg dla pozyskania środków pomocowych z funduszy Unii Europejskiej [Łuczak, 2016]. W ramach budowania strategii rozwoju ocenia się ważność czynników strategicznych (celów, zadań i scenariuszy rozwoju) oraz dokonuje się wyboru najlepszego scenariusza rozwoju. Ważność czynników strategicznych rozumiana jest jako ich pilność w realizacji z punktu widzenia rozwoju społeczno-gospodarczego.

Ważność czynników strategicznych można ocenić różnymi metodami, tj.: metody addytywne, metoda analitycznej hierarchizacji i metody pokrewne, metody werbalne, metody PROMETHEE, metody wykorzystujące punkty referencyjne i metody interaktywne [Trzaskalik, 2014a; 2014b]. W pracy zaproponowano do ich oceny rozmyty analityczny proces hierarchiczny (FAHP) [Chang, 1996]. Proponowana procedura obejmuje następujące etapy postępowania:

Etap 1. Budowa hierarchicznego schematu czynników strategicznych, które wywierają wpływ na rozwój jednostki samorządu terytorialnego.

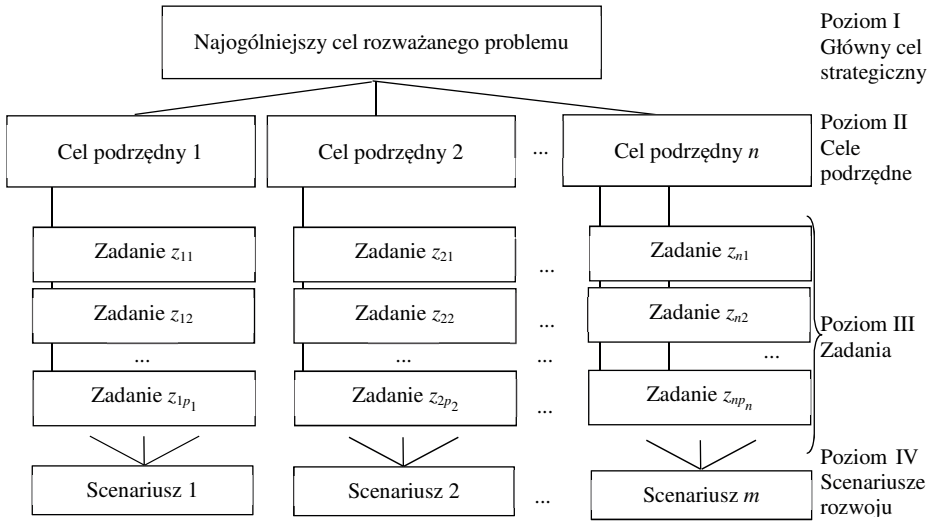
Etap 2. Porównania parami ważności czynników strategicznych.

Etap 3. Sprawdzenie zgodności porównań dokonanych przez ekspertów.

Etap 4. Obliczenie priorytetów lokalnych i globalnych czynników strategicznych i wybór najlepszego scenariusza rozwoju.

W pierwszym etapie buduje się hierarchiczny schemat decyzyjny, który obejmuje najważniejsze czynniki strategiczne wywierające wpływ na rozwój jednostki samorządu terytorialnego (rys. 1). Na szczycie hierarchii umieszczony jest główny cel strategiczny (poziom I), który zostaje uszczegółowiony przez cele podrzędne (poziom II). W ramach każdego celu podrzędnego zostają wyodrębnione najważniejsze zadania strategiczne (poziom III), które są niezbędne do osiągnięcia celów podrzędnych. Zadania również mogą zostać rozłożone na bardziej szczegółowe działania. Ostatni poziom hierarchii tworzą alternatywne decyzje, czyli scenariusze rozwoju. Z jednej strony należy unikać nadmiernego uszczegółowiania i skomplikowania hierarchii ze względu na możliwość pojawienia się problemów na dalszych etapach analiz związanych z interpretacją otrzymanych wyników. Z drugiej strony nadmierne uproszczenie modelu może prowadzić do trywializacji opisywanego zjawiska [Waszczyk, Szczerbicki, 2003].

Rysunek 1. Hierarchia czynników strategicznych dla jednostki samorządu terytorialnego



Źródło: Opracowanie własne.

W drugim etapie proponowanej procedury czynniki strategiczne zostają ocenione przez ekspertów. Jako że rozpatrywane problemy są złożone, a do ich rozwiązania potrzebne są doświadczenie i znajomość problematyki z różnych dziedzin życia społecznego, gospodarczego i finansów publicznych, niezbędne staje się wykorzystanie metod grupowego podejmowania decyzji. W grupowym podejmowaniu decyzji wyróżnia się dwa główne podejścia do agregacji opinii ekspertów [Prusak, Stefanów, 2014]:

- behawioralne (jakościowe) – opinie ekspertów są łączone lub syntetyzowane w procesie wymiany poglądów (dyskusji) w grupie, np. debata, burza mózgów, grupy fokusowe;
- matematyczne (ilościowe) – polegają na matematycznym syntetyzowaniu indywidualnych ocen za pomocą np.: średniej ważonej, średniej geometrycznej, mediany.

W proponowanej procedurze czynniki strategiczne zostają ocenione przez ekspertów poprzez dokonanie porównań parami ich ważności na każdym poziomie hierarchii. Na poziomie II porównuje się cele podrzędne w odniesieniu do ich wpływu (ważności) na osiąganie celu głównego. Na poziomie III dokonuje się porównań ważność zadań strategicznych, których realizacja przyczyni się do osiągnięcia celów podrzędnych. Natomiast na ostatnim poziomie porównuje się parami scenariusze rozwoju, badając ich siłę oddziaływania w realizacji poszczególnych zadań strategicznych.

W porównaniach wykorzystuje się dziewięciostopniową skalę preferencji między dwoma porównywanymi czynnikami strategicznymi (tab. 1).

Tablica 1. Dziewięciostopniowa skala preferencji między dwoma porównywanymi czynnikami strategicznymi

Przewaga ważności	Preferencje opisane słownie – ocena werbalna	Siła przewagi ważności – liczba rozmyta
Równoważność	Oba czynniki strategiczne są tak samo ważne w osiąganiu celu	$\tilde{1} = (1, 1, 1)$
Słaba	Pierwszy czynnik strategiczny jest niewiele ważniejszy od drugiego	$\tilde{3} = (1, 3, 5)$
Mocna	Pierwszy czynnik strategiczny jest istotnie ważniejszy od drugiego	$\tilde{5} = (3, 5, 7)$
Bardzo mocna	Pierwszy czynnik strategiczny jest zdecydowanie ważniejszy od drugiego	$\tilde{7} = (5, 7, 9)$
Absolutna	Pierwszy czynnik strategiczny jest absolutnie ważniejszy od drugiego	$\tilde{9} = (7, 9, 9)$
Dla porównań kompromisowych pomiędzy powyższymi wartościami	W przypadku gdy decydent nie jest zdecydowany pomiędzy dwoma sąsiadującymi ocenami, wtedy stosuje się pośrednie wartości	$\tilde{2} = (1, 2,4)$; $\tilde{4} = (2, 4, 6)$; $\tilde{6} = (4, 6, 8)$; $\tilde{8} = (6, 8, 9)$
Przechodniość ocen	Jeżeli *i*-ty czynnik strategiczny ma przypisany jeden z powyższych stopni podczas porównania do *j*-tego czynnika strategicznego, wtedy *j*-ty czynnik strategiczny ma odwrotną wartość, gdy porównuje się do *i*-tego	Odwrotności powyższych wartości

Źródło: [Saaty, 1980; Wang i inni, 2009].

Preferencje opisane słownie zostają przekształcone w trójkątne liczby rozmyte. Wyniki zapisuje się w postaci macierzy porównań parami (1):

$$\tilde{\mathbf{A}} = \begin{bmatrix} \tilde{1} & \tilde{x}_{12} & \dots & \tilde{x}_{1(\bullet)} \\ \dfrac{\tilde{1}}{\tilde{x}_{12}} & \tilde{1} & \dots & \tilde{x}_{2(\bullet)} \\ \vdots & \vdots & \dots & \vdots \\ \dfrac{\tilde{1}}{\tilde{x}_{1(\bullet)}} & \dfrac{\tilde{1}}{\tilde{x}_{2(\bullet)}} & \dots & \tilde{1} \end{bmatrix} \tag{1}$$

gdzie (•) – oznacza n dla celów podrzędnych, p_k dla zadań ($k = 1, 2, \ldots n$), m dla scenariuszy; $\tilde{x}_{ij} = (l_{ij}, m_{ij}, u_{ij})$, $i, j = 1,\ldots.(\bullet)$, $\tilde{x}_{ij}$ jest to siła przewagi ważności i-tego czynnika strategicznego nad j-tym, są to wartości uśrednione za pomocą średniej geometrycznej z ocen porównań parami dokonanymi przez ekspertów. Ponadto $\tilde{x}_{ij} = \tilde{1} / \tilde{x}_{ij}$ $(1/u_{ij}, 1/m_{ij}, 1/l_{ij})$ oraz $\tilde{x}_{ij} = \tilde{1}(1, 1, 1)$.

W etapie trzecim następuje sprawdzenie zgodności porównań parami [zob. też Alonso, Lamata, 2006]. Sprawdzenia zgodności porównań można dokonać za pomocą wskaźnika niezgodności *CR* [Saaty, 1980]:

$$CR = CI / RI \cdot 100\% \tag{2}$$

gdzie $CI = (\lambda_{\max} - n)/(n - 1)$ jest indeksem niezgodności, przy czym $\lambda_{\max}$ jest maksymalną lub główną wartością własną macierzy porównań **A**[1], a n jest liczbą wierszy (kolumn) w macierzy **A**, natomiast *RI* jest średnim losowym indeksem niezgodności obliczonym z losowo generowanej macierzy o wymiarach $n \times n$ (tab. 2).

Tablica 2. Średni losowy indeks niezgodności

Rząd macierzy (n)	1	2	3	4	5	6	7	8	9
Indeks niezgodności *(RI)*	0,0000	0,0000	0,5245	0,8815	1,1086	1,2479	1,3417	1,4056	1,4499

Źródło: [Alonso, Lamata, 2006].

Wskaźnik niezgodności *CR* nie powinien przekroczyć 10%. Jeżeli w macierzy **A** utworzonej z wartości m_{ij} porównania parami są zgodne, wtedy w rozmytej macierzy $\tilde{\mathbf{A}}$ porównania parami również są zgodne [Csutora, Buckley, 2001].

Jeśli porównania parami zostały przeprowadzone poprawnie[2], można obliczyć wartości priorytetów lokalnych i globalnych (etap 4). Oblicza się je w następujący sposób[3]:

Krok 1. Obliczenie rozmytej sumy dla każdego wiersza z rozmytej macierzy porównań $\tilde{\mathbf{A}}$ i ich normalizacja za pomocą operacji na liczbach rozmytych:

$$\tilde{Q}_i = \left(l_i, m_i, u_i\right) = \frac{\sum_{j=1}^{p_k} \left(l_{ij}, m_{ij}, u_{ij}\right)}{\sum_{i=1}^{p_k} \sum_{j=1}^{p_k} \left(l_{ij}, m_{ij}, u_{ij}\right)}, \; i = 1, 2, \ldots, p_k; k = 1, 2, \ldots, n \tag{3}$$

[1] Macierz **A** jest macierzą zawierającą tylko wartości m_{ij}.

[2] W przypadku gdy wartość wskaźnika *CR* jest znacznie większa od 10%, wskazuje to na zbyt dużą niespójność porównań parami czynników strategicznych. W takiej sytuacji zaleca się powtórzyć porównania parami [Saaty, 1980].

[3] Sposób obliczania priorytetów lokalnych zostanie przedstawiony na przykładzie zadań.

Krok 2. Obliczenie stopni możliwości, że $\tilde{Q}_i \geq \tilde{Q}_g$ ($i, g = 1, 2,\ldots, p_k, i \neq g$), według następującej formuły:

$$V\left(\tilde{Q}_i \geq \tilde{Q}_g\right) = \mathrm{hgt}\left(\tilde{Q}_i \cap \tilde{Q}_g\right) = \begin{cases} 1, & dla \quad m_i \geq m_g \\ 0, & dla \quad l_g \geq u_i \\ \dfrac{l_g - u_i}{(m_i - u_i) - (m_g - l_g)} & \text{w innych przypadkach,} \end{cases} \tag{4}$$

gdzie hgt jest to wysokość liczb rozmytych na przecięciu $\tilde{Q}_i$ i $\tilde{Q}_g$.

Następnie dokonuje się wyboru minimum z powyższych wartości $w_i^S = \min V(\tilde{Q}_i \geq \tilde{Q}_g)$. Wartości w_i^S po znormalizowaniu są priorytetami lokalnymi dla p_k ($k = 1, 2,\ldots, n$) zadań w ramach k-tego celu podrzędnego: $w_i = w_i^S / \sum_{i=1}^{p_k} w_i^S$.

Analogiczne oblicza się priorytety lokalne w_i dla celów podrzędnych[4] i scenariuszy rozwoju[5] według kroków 1 i 2. Priorytety lokalne poziomów II i III wyrażają udział danego czynnika strategicznego (celu podrzędnego lub zadania) w osiąganiu celu na poziomie bezpośrednio wyższym. Priorytety lokalne są podstawą do obliczenia priorytetów globalnych w_i^g, które reprezentują udział każdego czynnika strategicznego (z poszczególnych poziomów) w osiąganiu celu głównego. Priorytet globalny uzyskuje się przez przemnożenie wartości priorytetu lokalnego czynnika strategicznego z danego poziomu hierarchii przez wartość priorytetu globalnego czynnika strategicznego z nim powiązanego z poziomu bezpośrednio wyższego.

W przypadku ostatniego poziomu hierarchii dotyczącego scenariuszy priorytety lokalne mnoży się przez odpowiadające im priorytety globalne dla zadań. Otrzymane wielkości nazywa się cząstkowymi priorytetami globalnymi. Pokazują one udział danego scenariusza w osiąganiu celu głównego poprzez realizację rozpatrywanego zadania. Suma wszystkich cząstkowych priorytetów globalnych danego scenariusza jest jego priorytetem globalnym. Za najwłaściwszy scenariusz rozwoju uznaje się ten, dla którego priorytet globalny jest najwyższy.

Dodać należy, że wartości priorytetów globalnych uzyskane za pomocą FAHP mogą przyjmować wartości równe zero zarówno dla celów podrzędnych, jak i zadań. Z jednej strony jest to zaletą metody, gdyż pomaga ona dokonać selekcji czynników strategicznych i daje możliwość wyeliminowania zadań, które są mało istotne. Z drugiej strony stanowi to pewną wadę tej metody, gdyż może prowadzić do usunięcia niektórych celów podrzędnych.

[4] Lokalne i globalne priorytety dla danego celu podrzędnego są identyczne.

[5] Priorytety lokalne scenariuszy rozwoju oblicza się w odniesieniu do każdego zadania.

2. Wyniki badań empirycznych

W pracy podjęto próbę oceny czynników strategicznych (celów, zadań i scenariuszy rozwoju) dla gminy Chocz w powiecie pleszewskim w perspektywie do 2020 r. za pomocą rozmytego analitycznego procesu hierarchicznego i danych zebranych wśród lokalnych ekspertów. Ważnym problemem jest znalezienie odpowiedniej grupy ekspertów, wśród których będą zbierane dane. Istotna jest wielkość grupy eksperckiej i właściwy dobór jej członków. Jeśli chodzi o liczbę ekspertów to przyjmuje się grupy składające się od kilku do kilkunastu osób. W przypadku małej liczby ekspertów mogą pojawić się problemy związane z konsekwencjami porównań czynników strategicznych. Natomiast przy znacznej liczbie ekspertów (powyżej 20) może wystąpić zjawisko równoważenia ocen. Członkowie grupy eksperckiej powinny być przedstawicielami różnych dziedzin i środowisk oraz reprezentować wspólne interesy gminy. Te kryteria spełniają grupy eksperckie mające w składzie radnych gminnych lub powiatowych [zob. też Łuczak, 2016]. Badania przeprowadzono na podstawie danych zebranych wśród wybranych ekspertów: 14 radnych gminy wiejskiej Chocz oraz jej wójta, 18 radnych powiatu pleszewskiego i jego starosty. Porównania parami radnych gminy i radnych powiatu zostały uśrednione za pomocą średniej geometrycznej.

Pierwszym etapem była budowa hierarchii czynników strategicznych wywierających wpływ na rozwój gminy (rys. 2).

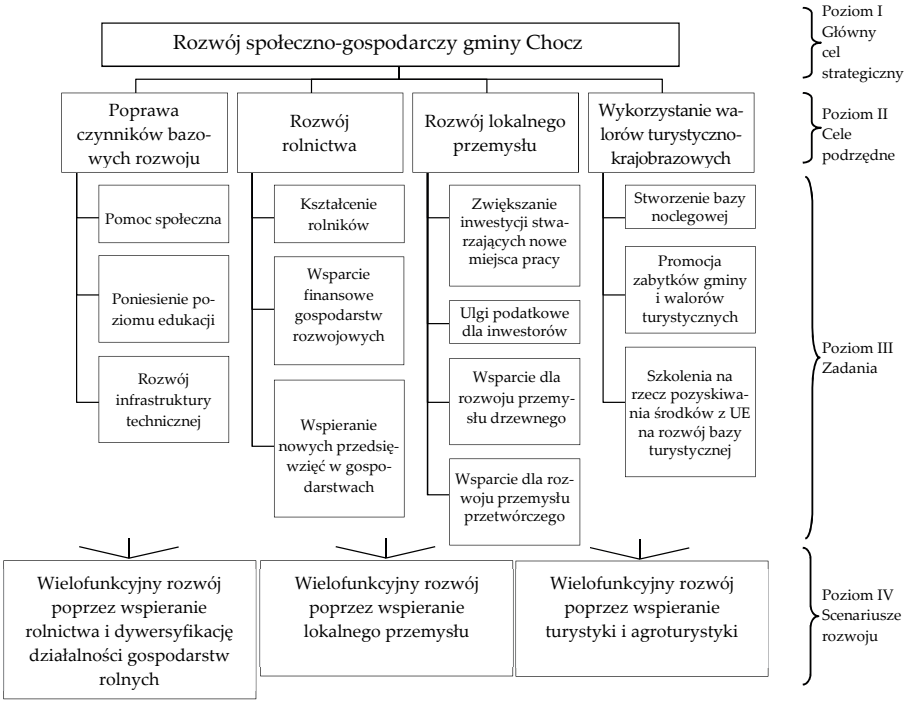
Przyjęto, że głównym celem strategicznym będzie rozwój społeczno-gospodarczy gminy prowadzący do poprawy jakości życia mieszkańców gminy, a na jego osiągnięcie będą wpływały cztery cele podrzędne dotyczące: czynników bazowych rozwoju, rozwoju rolnictwa, rozwoju lokalnego przemysłu i wykorzystania walorów turystyczno-krajobrazowych. W ramach każdego celu podrzędnego wyróżniono niezbędne do ich osiągnięcia zadania strategiczne (rys. 2).

Cele i zadania wpływają na scenariusze rozwoju, które dotyczą wielofunkcyjnego rozwoju poprzez wspieranie:

- rolnictwa i dywersyfikację działalności gospodarstw rolnych (scenariusz 1),
- lokalnego przemysłu (scenariusz 2),
- turystyki i agroturystyki (scenariusz 3).

Na każdym poziomie hierarchii czynniki strategiczne zostały porównane parami przez ekspertów lokalnych: radnych gminy, wójta, radnych powiatu i starostę.

Rysunek 2. Schemat hierarchiczny czynników strategicznych wpływających na rozwój gminy Chocz



Źródło: Opracowanie własne.

Na podstawie ocen ekspertów obliczono z wykorzystaniem metod rozmytego analitycznego procesu hierarchicznego (FAHP) priorytety lokalne i globalne dla celów podrzędnych, zadań i scenariuszy rozwoju. W omówieniu ważności czynników strategicznych według opinii ekspertów lokalnych zostaną wykorzystane wartości priorytetów globalnych (tab. 3 i 7). Za najważniejszy cel podrzędny należy uznać poprawę czynników bazowych rozwoju. W opinii wszystkich ekspertów lokalnych priorytet globalny osiągnął dość dużą wartość. Były to wartości od 0,242 według opinii radnych powiatu do 0,473 według opinii starosty powiatu pleszewskiego (tab. 3). Oznacza to, że osiągnięcie tego celu będzie miało wpływ na cel główny w 24,3% według opinii radnych powiatu, a według opinii starosty powiatu wpływ ten będzie prawie dwukrotnie większy – 47,3%. Osiągnięcie tego celu jest istotne ze względu na słabe strony infrastruktury technicznej i społecznej.

Drugim ważnym celem podrzędnym jest rozwój lokalnego przemysłu. Radni gminy (priorytet globalny – 0,264), radni powiatu (0,266) i wójt gminy (0,348) ocenili ten cel jako ważny. Według danych GUS w 2012 r. były

597 podmioty wpisane do rejestru REGON na 10 tys. ludności. Nie mniej ważny według opinii radnych gminy (0,306) i radnych powiatu (0,255) okazał się cel podrzędny związany z rozwojem rolnictwa. Ważności tego celu podrzędnego nie podkreślili ani wójt gminy, ani starosta powiatu. Dostrzegli oni jednak znaczne walory krajobrazowe gminy. W 2012 r. według danych GUS ponad 40% powierzchni gminy zajmowały lasy. Stąd też wójt i starosta przykładali wagę do celu dotyczącego wykorzystania walorów turystyczno-krajobrazowych. Należy zauważyć, że oceny wszystkich celów podrzędnych według radnych powiatu są raczej zachowawcze. Wszystkie cele zostały ocenione na zbliżonym poziomie (około 0,25).

Wartości priorytetów globalnych dla celów podrzędnych mają przełożenie na wartości priorytetów globalnych dla zadań. W obrębie celu podrzędnego związanego z poprawą czynników bazowych rozwoju wszyscy eksperci lokalni dostrzegają też potrzebę poniesienia poziomu edukacji. W opinii radnych gminy i starosty powiatu jest to zadanie bardzo ważne. Stąd też wartości priorytetów globalnych odpowiednio 0,171 i 0,244. Dla pozostałych ekspertów zadanie to było ważne. Stąd nieco niższe wartości priorytetów globalnych (0,130 – wójt, 0,143 – radni powiatu). Należy zauważyć, że w opinii prawie wszystkich ekspertów zadania w ramach celu dotyczącego poprawy czynników bazowych rozwoju są ważne co najmniej na poziomie średnim, gdyż wartości priorytetów w większości przypadków przekraczają 0,05.

Tablica 3. Ocena ważności celów podrzędnych i zadań (priorytety globalne) wpływających na osiągnięcie celu głównego dla gmin Chocz

Cele podrzędne	Eksperci lokalni				Zadania	Eksperci lokalni			
	radni gminy	wójt	radni powiatu	starosta		radni gminy	wójt	radni powiatu	starosta
Poprawa czynników bazowych rozwoju	**0,285**[a)]	**0,339**	**0,242**	**0,473**	Pomoc społeczna	0,072	0,139	0,023	0,102
					Podniesienie poziomu edukacji	0,171	0,130	0,143	0,244
					Rozwój infrastruktury technicznej	0,042	0,070	0,076	0,127
Rozwój rolnictwa	**0,306**	0,024	**0,255**	0,081	Kształcenie rolników	0,006	0,000	0,092	0,000
					Wsparcie finansowe gospodarstw rozwojowych	0,300	0,023	0,050	0,029
					Wspieranie nowych przedsięwzięć w gospodarstwach	0,000	0,002	0,114	0,052

Cele podrzędne	Eksperci lokalni				Zadania	Eksperci lokalni			
	radni gminy	wójt	radni powiatu	starosta		radni gminy	wójt	radni powiatu	starosta
Rozwój lokalnego przemysłu	**0,264**	**0,348**	**0,266**	0,000	Zwiększanie inwestycji stwarzających nowe miejsca pracy	0,233	0,219	0,156	0,000
					Ulgi podatkowe dla inwestorów	0,031	0,000	0,053	0,000
					Wsparcie dla rozwoju przemysłu drzewnego	0,000	0,053	0,000	0,000
					Wsparcie dla rozwoju przemysłu przetwórczego	0,000	0,075	0,056	0,000
Wykorzystanie walorów turystyczno-krajobrazowych	0,145	**0,289**	**0,237**	**0,446**	Stworzenie bazy noclegowej	0,041	0,087	0,000	0,179
					Promocja zabytków gminy i walorów turystycznych	0,069	0,000	0,096	0,267
					Szkolenia na rzecz pozyskiwania środków z UE na rozwój bazy turystycznej	0,035	0,202	0,141	0,000

[a] Pogrubioną czcionką zaznaczono wartości priorytetów dla najważniejszych celów podrzędnych.

Kolorem ciemnym szarym zaznaczono najwyższe wartości priorytetów globalnych dla zadań $w_i^g \in\langle 0{,}15;\ 0{,}30\rangle$, średnim szarym – ważne $w_i^g \in\langle 0{,}10;\ 0{,}15)$, jasnym szarym – średnio ważne $w_i^g \in\langle 0{,}05;\ 0{,}10)$, białym – zadania, które nie muszą być realizowane $w_i^g \in\langle 0;\ 0{,}05)$.

Źródło: Obliczenia własne na podstawie danych zebranych za pomocą kwestionariusza statystycznego wśród wybranych ekspertów dla gminy Chocz [Kaczmarek, 2012].

W ramach celu Rozwój rolnictwa znacząco wyróżnia się zadanie Wsparcie finansowe gospodarstw rozwojowych, które zostało ocenione przez radnych gminy jako bardzo ważne (0,300). Z perspektywy pozostałych ekspertów lokalnych zadanie nie jest aż tak istotne.

Do najważniejszych zadań dla gminy Chocz należy też zaliczyć zwiększanie inwestycji stwarzających nowe miejsca pracy, które wpływają na osiągnięcie celu związanego z rozwojem lokalnego przemysłu. Radni gminy (0,233), wójt (0,219) i radni powiatu (0,156) uznali to zadanie za bardzo ważne.

W ramach celu Wykorzystanie walorów turystyczno-krajobrazowych według starosty powiatu pleszewskiego bardzo ważne są zadania dotyczące stworzenie bazy noclegowej (0,179) oraz promocja zabytków gminy i walorów turystycznych (0,267). Natomiast wójt gminy Chocz dostrzegł bardzo dużą wagę szkoleń na rzecz pozyskiwania środków z Unii Europejskiej na rozwój bazy turystycznej (0,202).

Natomiast do najmniej istotnych zadań należy zaliczyć wsparcie dla rozwoju przemysłu drzewnego. Według radnych gminy, radnych powiatu i starosty powiatu zadanie to nie musi być realizowane, gdyż priorytet globalny tego zadania wynosi zero (tab. 3).

Tablica 4. Wartości wskaźnika podobieństwa struktur Bray'a-Curtisa dla wartości priorytetów globalnych celów podrzędnych, zadań i scenariuszy rozwoju według opinii lokalnych ekspertów

Eksperci	radni gminy	wójt	radni powiatu	starosta	
radni gminy	1	0,718	**0,906**	0,511	Cele podrzędne
wójt		1	0,769	0,652	
radni powiatu			1	0,560	
starosta				1	
Eksperci	radni gminy	wójt	radni powiatu	starosta	
radni gminy	1	0,562	0,555	0,424	Zadania
wójt		1	0,601	0,414	
radni powiatu			1	0,419	
starosta				1	
Eksperci	radni gminy	wójt	radni powiatu	starosta	
radni gminy	1	0,830	**0,910**	0,722	Scenariusze rozwoju
wójt		1	0,750	0,717	
radni powiatu			1	0,733	
starosta				1	

Źródło: Obliczenia własne na podstawie tab. 3 i 7.

Analizując podobieństwo wartości priorytetów globalnych dla czynników strategicznych według lokalnych ekspertów, należy zauważyć, że opinie radnych gminy i radnych powiatu dotyczące celów podrzędnych i scenariuszy rozwoju wykazują dużą zgodność. Świadczą o tym wartości wskaźnika podobieństwa struktur Bray'a-Curtisa, które przekraczają 0,9[6] (tab. 4). Natomiast nie można wskazać znaczącego podobieństwa w wartościach priorytetów globalnych zadań dla żadnych ekspertów lokalnych.

Dla gminy Chocz wyróżniono zadania najważniejsze (wartość priorytetu globalnego $w_i^g \in\langle 0{,}15; 0{,}30\rangle$), ważne ($w_i^g \in\langle 0{,}10; 0{,}15\rangle$), średnio ważne ($w_i^g \in\langle 0{,}05; 0{,}10\rangle$) i te, które nie muszą być realizowane ($w_i^g \in\langle 0; 0{,}05\rangle$) (tab. 5). Hierarchia ta posłużyła do zbudowania modelu rozwoju lokalnego (tab. 6).

[6] Im wartość wskaźnika podobieństwa struktur Bray'a-Curtisa jest bliższa 1, tym bardziej do siebie podobne są wartości priorytetów globalnych. Im bliżej 0, tym wartości priorytetów są mniej podobne.

Model obejmuje: siłę pilności zadań wyrażoną w punktach według opinii wszystkich ekspertów lokalnych, pilność zadań przedstawioną symboliczne, wskazanie zadań, które nie muszą być pilnie realizowane, i określenie priorytetu realizacji zadania.

Tablica 5. Hierarchia ważności zadań dla gmin Chocz według grup ekspertów lokalnych

Zadania	Eksperci lokalni			
	radni gminy	wójt	radni powiatu	staro-sta
Pomoc społeczna	+	++	0	++
Poniesienie poziomu edukacji	+++	++	++	+++
Rozwój infrastruktury technicznej	0	+	+	++
Kształcenie rolników	0	0	+	0
Wsparcie finansowe gospodarstw rozwojowych	+++	0	+	0
Wspieranie nowych przedsięwzięć w gospodarstwach	0	0	++	+
Zwiększanie inwestycji stwarzających nowe miejsca pracy	+++	+++	+++	0
Ulgi podatkowe dla inwestorów	0	0	+	0
Wsparcie dla rozwoju przemysłu drzewnego	0	+	0	0
Wsparcie dla rozwoju przemysłu przetwórczego	0	+	+	0
Stworzenie bazy noclegowej	+	+	0	+++
Promocja zabytków gminy i walorów turystycznych	+	0	+	+++
Szkolenia na rzecz pozyskiwania środków z UE na rozwój bazy turystycznej	0	+++	++	0

Oznaczenia: +++ zadanie najważniejsze – priorytetowe, ++ zadanie ważne, + zadanie średnio ważne, 0 zadanie nie musi być pilnie realizowane.
Źródło: Opracowanie własne na podstawie tab. 3.

Uzupełnieniem hierarchii ważności zadań dla gmin Chocz są scenariusze rozwoju wraz z ich ocenami – priorytetami globalnymi (tab. 7). Ocena najważniejszego ze scenariuszy w opinii radnych gminy i radnych powiatu była zgodna. Radni wskazali najlepszy scenariusz związany z wielofunkcyjnym rozwojem rolnictwa i dywersyfikacją działalności gospodarstw rolnych. Związane jest to z typowo rolniczym charakterem gminy, gdzie użytki rolne zajmują ponad 50% obszaru i dominują małe gospodarstwa rolne od 2 do 10 ha z przewagą słabych gleb. Priorytety globalne tego scenariusza miały wartość ponad 0,5.

Tablica 6. Hierarchia ważności zadań dla gmin Chocz (model rozwoju lokalnego – część I)

Zadania	Siła pilność zadań	Pilność zadań	Zadania nie muszą być pilnie realizowane	Priorytet realizacji zadania
Pomoc społeczna	5	++(2)	0(1)	wysoki
Poniesienie poziomu edukacji	10	+++(2)	0(0)	b. wysoki
Rozwój infrastruktury technicznej	4	+(2)	0(1)	średni
Kształcenie rolników	1	+(1)	0(3)	niski
Wsparcie finansowe gospodarstw rozwojowych	4	++(2)*	0(2)	średni
Wspieranie nowych przedsięwzięć w gospodarstwach	3	+(2)*	0(2)	niski
Zwiększanie inwestycji stwarzających nowe miejsca pracy	9	+++(3)	0(1)	b. wysoki
Ulgi podatkowe dla inwestorów	1	+(1)	0(3)	niski
Wsparcie dla rozwoju przemysłu drzewnego	1	+(1)	0(3)	niski
Wsparcie dla rozwoju przemysłu przetwórczego	2	+(2)	0(2)	niski
Stworzenie bazy noclegowej	5	+(2)	0(1)	średni
Promocja zabytków gminy i walorów turystycznych	5	+(2)	0(1)	średni
Szkolenia na rzecz pozyskiwania środków z UE na rozwój bazy turystycznej	5	++(2)*	0(2)	średni

Oznaczenia: +++ zadanie najważniejsze – priorytetowe, ++ zadanie ważne, + zadanie średnio ważne, 0 zadanie nie musi być pilnie realizowane. (·) w nawiasie podano liczbę powtarzających się ocen grupy ekspertów. * oznacza średnią wartości pilności.
Źródło: Opracowanie własne na podstawie tab. 5.

Tablica 7. Wartości priorytetów globalnych dla scenariuszy rozwoju gminy Chocz (model rozwoju lokalnego – część II)

Scenariusze – Wielofunkcyjny rozwój	Eksperci lokalni			
	radni gminy	wójt	radni powiatu	starosta
rolnictwo i dywersyfikacja działalności gospodarstw rolnych	**0,503**	0,333	**0,583**	0,409
lokalny przemysł	0,362	**0,461**	0,271	0,178
turystyka i agroturystyka	0,135	0,206	0,145	**0,413**

Źródło: Obliczenia własne na podstawie danych zebranych za pomocą kwestionariusza statystycznego wśród wybranych ekspertów dla gminy wiejskiej Chocz [Kaczmarek, 2012].

Opinia wójta gminy ujawniła, że najważniejszy jest wielofunkcyjny rozwój poprzez wspieranie lokalnego przemysłu (0,461). Związane jest to z niewielką liczbą głównie małych podmiotów gospodarczych (zakłady rzemieślnicze i handlowe) funkcjonujących na terenie gminy. Dodać należy, że na terenie gminy zasób terenów dla działalności biznesowej jest niewystarczający i obserwowany jest deficyt rodzimego kapitału. Natomiast ocena starosty powiatowego pokazała, że najważniejszy w jego opinii jest wielofunkcyjny rozwój poprzez turystykę i agroturystykę (0,413). Gmina posiada znaczne walory krajobrazowe, jednak brak jest bazy turystycznej.

Zakończenie

Przeprowadzone badania empiryczne pozwoliły na zweryfikowanie sformułowanej w pracy hipotezy, potwierdzając przydatność FAHP do modelowania rozwoju lokalnego. Proponowane podejście dla gminy Chocz ukazało nowe zastosowania metodologiczne w modelowaniu rozwoju lokalnego. Metoda rozmytego analitycznego procesu hierarchicznego pozwoliła na ocenę ważności czynników strategicznych z wykorzystaniem jakościowych preferencji lokalnych ekspertów, które zostały skwantyfikowane. Podejście to łączy metody ilościowe z wykorzystaniem jakościowych ocen dokonanych przez ekspertów.

Dzięki zastosowaniu FAHP oceniono ważność celów i zadań strategicznych oraz scenariuszy rozwoju gminy na podstawie opinii różnych ekspertów: radnych gminy, radnych powiatu, wójta gminy i starosty powiatu. Opinie grupowe ekspertów – radnych (gminy i powiatu) były w znacznym stopniu zgodne w odniesieniu do najważniejszych celów podrzędnych i scenariuszy rozwoju. W zakresie oceny ważności zadań odnotowano zróżnicowane preferencje.

Należy zauważyć, że wartości priorytetów globalnych czynników strategicznych są dość „wyostrzone". Wśród zadań ujawniły się najważniejsze zadania priorytetowe, tj. podniesienie poziomu edukacji oraz zwiększanie inwestycji stwarzających nowe miejsca pracy. Na podstawie ocen ważności zadań dokonanych przez ekspertów lokalnych ustalono hierarchię ich ważności, wskazując zadania najważniejsze – priorytetowe, ważne, średnio ważne i te, które nie muszą być pilnie realizowane.

Ocena najważniejszego ze scenariuszy w opinii radnych (gminy i powiatu) była zgodna. Radni wskazali najlepszy scenariusz związany z wielofunkcyjnym rozwojem rolnictwa i dywersyfikacją działalności gospodarstw rolnych.

Oceniając użyteczność FAHP, należy podkreślić niewątpliwą jego zaletę, że zastosowanie tej metody pozwala przyporządkować zróżnicowane współczynniki wagowe czynnikom strategicznym. Istnieje też możliwość

wystąpienia wartości priorytetu równego zero, który jest przydzielany, gdy wstępny priorytet w postaci rozmytej trójkątnej liczby nie przecina się co najmniej z jedną z pozostałych trójkątnych liczb rozmytych. Świadczy to o znacznym oddaleniu wartości tego priorytetu od pozostałych. Z merytorycznego punktu widzenia wydaje się, że nie można jednoznacznie stwierdzić, że jego wartość będzie dokładnie równa zero. Stąd też należy uznać to za wadę tej metody.

W końcowym etapie badań zbudowano model rozwoju gminy ujawniający priorytety czynników rozwojowych. Proponowane podejście do modelowania rozwoju lokalnego dotyczącego oceny ważności czynników strategicznych dla gminy może być wykorzystane w procesie planowania rozwoju społeczno-gospodarczego jednostek administracyjnych, jako przesłanki do budowy strategii rozwoju. Oczywiście modelowanie rozwoju lokalnego jest procesem ciągłym, w którym uzyskiwane efekty i zachodzące zmiany uwarunkowań wymuszają odpowiednie korekty. Na szczególną uwagę zasługuje otwarty charakter schematu hierarchicznego czynników strategicznych wpływających na rozwój gminy, który powinien ulegać korektom z upływem czasu. Zrealizowane zadania i osiągnięte cele należy usuwać z modelu, a nowe elementy powinny być do niego wprowadzane.

Literatura

Alonso J.A., Lamata M.T. (2006), *Consistency in the analytic hierarchy process: a new approach*, „International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems", Vol. 14, No. 4.

GUS (2012), *Bank Danych Lokalnych*, https://bdl.stat.gov.pl/BDL/start, dostęp: 5.02.2018.

Chang D.-Y. (1996), *Application of the extent analysis method on fuzzy AHP*, „European Journal of Operational Research", Vol. 95, No. 3.

Csutora R., Buckley J.J. (2001), *Fuzzy hierarchical analysis: the Lambda-Max method*, „Fuzzy Sets and Systems", Vol. 120, No. 2.

Escobar M.T., Moreno-Jiménez J.M. (2007), *Aggregation of individual preference structures in AHP-group decision making*, „Group Decision and Negotiation", Vol. 16, No. 4.

Forman, E., Peniwati K. (1998), *Aggregating individual judgements and priorities with the Analytic Hierarchy Process*, „European Journal of Operational Research", Vol. 108, No. 1.

Kaczmarek K. (2012), *Wybór scenariusza rozwoju społeczno-gospodarczego dla gminy Chocz z wykorzystaniem metody AHP*, Materiał źródłowy. Uniwersytet Przyrodniczy w Poznaniu, Poznań.

Łuczak A. (2016), *Wielokryterialne metody ilościowe w diagnozowaniu i modelowaniu rozwoju lokalnego*, Wyd. Uniwersytetu Przyrodniczego w Poznaniu, Poznań.

Prusak A., Stefanów P. (2014), *AHP – analityczny proces hierarchiczny. Budowa i analiza modeli decyzyjnych krok po kroku*, C.H. Beck, Warszawa.

Saaty T.L., (1980), *The analytic hierarchy process. Planning, priority setting, resource allocation*, MacGraw-Hill, New York International Book Company.
Saaty T.L., Peniwati K., (2007), *Group decision making: drawing out and reconciling differences*, wyd. 1, RWS Publications, Pittsburgh.
Trzaskalik T. (2014a), *Wielokryterialne wspomaganie decyzji. Przegląd metod i zastosowań*, „Zeszyty Naukowe Politechniki Śląskiej, Seria: Organizacja i Zarządzanie", z. 74 (1921).
Wang J.-W., Cheng C.-H., Kun-Cheng H., (2009), *Fuzzy hierarchical TOPSIS for supplier selection*, „Applied Soft Computing", Vol. 9, No. 1.
Trzaskalik T. (red.) (2014b), *Wielokryterialne wspomaganie decyzji* , PWE, Warszawa.
Waszczyk M., Szczerbicki E. (2003), *Metodologiczne aspekty opisowego modelowania w naukach ekonomicznych*, „Zeszyty Naukowe Politechniki Gdańskiej", nr 589, „Filozofia", nr 7.

Streszczenie

Celem badań było zbudowanie modelu rozwojowego dla gminy z wykorzystaniem opinii ekspertów i rozmytego analitycznego procesu hierarchicznego (FAHP). Konstruowany model obejmuje najistotniejsze czynniki strategiczne: cele i zadania oraz scenariusze rozwoju. Czynniki strategiczne zostały porównane parami i oceniono ich ważność. Porównań dokonali eksperci lokalni: radni gminy, radni powiatu oraz wójt i starosta powiatu. Oceny porównań stanowią podstawę do obliczenia ocen ważności czynników strategicznych za pomocą FAHP. W efekcie zostają utworzone cząstkowe modele, które są podstawą do opracowania zagregowanego modelu rozwoju gminy. Badania przeprowadzono na danych pochodzących z badań ankietowych przeprowadzonych wśród wybranych lokalnych ekspertów dla gminy wiejskiej Chocz.

Słowa kluczowe

rozwój lokalny, czynniki strategiczne, lokalna jednostka administracyjna, FAHP

Modeling local development using opinions of selected experts groups (Summary)

The main aim of this study was to build a model of development for the municipality using opinions of experts and fuzzy analytic hierarchy process (FAHP). The constructed model includes the most important strategic factors: goals, tasks and development scenarios. Strategic factors were pairwise compared and their importance was assessed. Comparisons were made by local experts: councilors of the municipality, councilors of the county, commune head and district governor. Comparisons are the basis for calculating the evaluations of importance of strategic factors using FAHP. As a result, partial models are created. They are the basis for the elaboration of an aggregate development model of municipality. The research was carried out on data from the survey conducted among local experts of municipality of Chocz.

Keywords

local development, strategic factors, local administrative unit (LAU), FAHP

Maciej Nowak*
Tadeusz Trzaskalik**

Zarządzanie zdolnością produkcyjną z wykorzystaniem interaktywnego programowania dynamicznego

Wstęp

Zdolność produkcyjną definiuje się zwykle jako maksymalną ilość wyrobów lub usług, którą przedsiębiorstwo jest w stanie zaoferować swoim klientom. Decyzje dotyczące jej zwiększenia powinny być częścią zsynchronizowanego planu, prowadzącego do uzyskania przez firmę przewagi konkurencyjnej. Realizowane projekty inwestycyjne mają prowadzić nie tylko do zdobycia nowych klientów, ale również do zwiększenia elastyczności realizowanych procesów, skrócenia terminów dostaw i poprawy jakości.

Planując działania nakierowane na zwiększenie zdolności produkcyjnej, należy odpowiedzieć sobie na dwa podstawowe pytania: w którym momencie powinny one być zainicjowane, oraz jak duży powinien być ich zakres. Jest to o tyle istotne, że w długim, kilkuletnim horyzoncie czasu zachodzi konieczność wielokrotnego podejmowania projektów inwestycyjnych prowadzących do zwiększenia zasobów produkcyjnych przedsiębiorstwa.

Problem planowania zdolności produkcyjnej należy, obok zagadnienia odnowienia zapasów, programowania produkcji oraz pewnych typów odnowy parku maszynowego, włączając w nie wymianę wyposażenia, do ważnych zagadnień występujących w przemyśle, w stosunku do których wykorzystujemy dynamiczne modele optymalizacyjne. Zagadnienia te, jako problemy jednokryterialne, rozwiązywane były już we wczesnym okresie rozwoju badań operacyjnych [Wagner, 1980]. W okresie późniejszym, wraz z rozwojem metod wielokryterialnego wspomagania decyzji, pojawiły się również wielokryterialne ujęcia tych zagadnień.

Rozpatrywany w niniejszej pracy problem planowania zdolności produkcyjnej był również rozważany jako zagadnienie wielokryterialne. W pracy [Geng, Jiang, 2009] przedstawiono przegląd metod dedykowanych

* Prof. UE dr hab., Katedra Badań Operacyjnych, Wydział Informatyki i Komunikacji, Uniwersytet Ekonomiczny w Katowicach, ul. 1 Maja 50, 40-287 Katowice, maciej.nowak@ue.katowice.pl
** Prof. dr hab., Katedra Badań Operacyjnych, Wydział Informatyki i Komunikacji, Uniwersytet Ekonomiczny w Katowicach, ul. 1 Maja 50, 40-287 Katowice, tadeusz.trzaskalik@ue.katowice.pl

dla producentów półprzewodników. Większość z nich może być zaadaptowanych do warunków panujących w innych sektorach. W pracy [Cheng i inni, 2004] planowanie zdolności produkcyjnej rozważane jest łącznie z zagadnieniem zarządzania zapasami. Autorzy formułują wielokryterialny model podejmowania decyzji w warunkach niepewności oraz proponują metodę opartą na łańcuchach Markowa. Z kolei w pracy [Fangqi i inni, 2011] problem zarządzania zdolnością produkcyjną jest rozważany wspólnie z zagadnieniem planowania produkcji i dystrybucji. Zaprezentowano model wielokryterialnego programowania całkowitoliczbowego oraz zaproponowano metodę jego rozwiązania opartą na relaksacji Lagrange'a.

W niniejszym artykule zagadnienie planowania zdolności produkcyjnej przedstawiamy jako dynamiczny problem wielokryterialny. Dokonując hierarchizacji rozpatrywanych kryteriów, pokażemy, w jaki sposób można go rozwiązać we współpracy analityka z decydentem. Cele pracy są następujące:

1) zaprezentowanie, w jaki sposób problem planowania zdolności produkcyjnej może być opisany za pomocą modelu dyskretnego programowania dynamicznego,
2) zaproponowanie nowej interaktywnej procedury o charakterze iteracyjnym, za pomocą której można rozwiązać tak sformułowany problem,
3) przedstawienie przykładu liczbowego ilustrującego sposób wykorzystania zaproponowanego podejścia.

Opracowanie składa się z czterech części. W części 1 omówiono wybrane zagadnienia, związane z dyskretnym programowaniem dynamicznym, uwzględniając zarówno aspekty jedno-, jak i wielokryterialne. W części 2 przedstawione zostało dynamiczne ujęcie problemu zarządzania zdolnością produkcyjną, wykorzystywane w dalszej części pracy. W kolejnej części przedstawiona została autorska propozycja procedury interaktywnej, pozwalającej na rozwiązanie sformułowanego w części 2 problemu decyzyjnego, natomiast w części 4 zamieszczono ilustrację liczbową zaproponowanej procedury. Całość kończy podsumowanie.

Artykuł powstał w ramach projektu badawczego NCN DEC-2013/11/B/HS4/01471.

1. Dyskretne programowanie dynamiczne

Procesy decyzyjne mają często charakter wieloetapowy. Zajmiemy się dalej procesami, w których liczba etapów jest z góry ustalona. Sterowanie takimi procesami polega na podejmowaniu w poszczególnych etapach kolejnych decyzji, przy czym decyzje podjęte w etapach wcześniejszych mogą mieć istotny wpływ na możliwości podejmowania kolejnych decyzji

w etapach późniejszych, gdyż zawężają lub – przeciwnie – rozszerzają pole decyzyjne.

Wśród zmiennych pojawiających się w procesach wieloetapowych wyróżniamy zmienne stanu, charakteryzujące proces na początku rozpatrywanego etapu, oraz zmienne decyzyjne, na które mamy bezpośredni wpływ. W rozpatrywanym w niniejszej pracy przypadku deterministycznym stan procesu na początku kolejnego etapu określony jest jednoznacznie przez stan procesu na początku rozpatrywanego etapu oraz podjętą w tym etapie decyzję.

Realizacją procesu jest ciąg następujących po sobie stanów i decyzji. Realizacje możemy porównywać ze sobą, wykorzystując wieloetapowe funkcje kryterialne. W najczęściej spotykanym przypadku funkcje te to sumy (lub sumy zdyskontowane) wartości kryteriów etapowych. W przypadku jednokryterialnym, gdy mamy do czynienia z jednym kryterium wieloetapowym, realizację optymalną procesu, czyli taką, dla której rozpatrywane kryterium przyjmuje wartość maksymalną lub minimalną, można znaleźć, wykorzystując zasadę optymalności Bellmana oraz równania optymalności [Bellman, 1957]. Odpowiednio przekształcone równania optymalności pozwalają rekurencyjnie wyznaczyć kolejne realizacje procesu ze względu na rozpatrywane kryterium (drugą, trzecią itd.) [Elmaghraby, 1970; Trzaskalik, 2016].

Wieloetapowe procesy decyzyjne niejednokrotnie rozpatruje się jako procesy wielokryterialne, biorąc pod uwagę więcej niż jedno kryterium oceny. W pracy [Ganguly i inni, 2013] autorzy rozpatrują wielokryterialne podejście w planowaniu systemu dystrybucji energii elektrycznej. Natomiast w publikacji [Kannan i inni, 2013] przedstawione zostało dynamiczne, zintegrowane, rozmyte podejście wielokryterialne w problemie wyboru dostawców. W opracowaniu [Nguyen i inni, 2013] podjęty został wielokryterialny dynamiczny problem harmonogramowania z wykorzystaniem metaheurystyk, zaś w pracy [Soroudi i inni, 2011] przedstawiono probabilistyczny, dynamiczny model w planowaniu wykorzystania zasobów odnawialnych i nieodnawialnych. W pracy [Jafarian-Moghaddam, Ghoseiri, 2011] autorzy przedstawili dynamiczny wielokryterialny model DEA. Z kolei [Huang i inni, 2011] wykorzystali dynamiczne ujęcie wielokryterialne do sterowania zmiennymi w czasie, niestabilnymi modelami, a [Mafakheri i inni, 2011] zaproponowali dwuetapowe wielokryterialne podejście w zarządzaniu łańcuchem dostaw z wykorzystaniem programowania dynamicznego. W opracowaniu [Li i inni, 2010] autorzy przedstawili podejście programowania dynamicznego w wielokryterialnym, grupowym podejmowaniu decyzji.

Przytoczone powyżej prace wskazują na to, że podejście wielokryterialnego, dynamicznego podejmowania decyzji jest bardzo uniwersalne, a zaproponowane sposoby rozwiązywania dynamicznych problemów decyzyjnych są bardzo zróżnicowane. Poniżej przedstawimy ujęcie wykorzystywane dalej w niniejszym artykule.

Przyjmijmy, że rozpatrujemy zadanie wektorowej maksymalizacji (inne zadania, w których pewne kryteria minimalizujemy, łatwo sprowadzić do zadania wektorowej maksymalizacji, opisywanego dalej). Porównując dwie realizacje procesu, powiemy, że pierwsza z nich jest lepsza od drugiej, jeżeli wartości wszystkich rozpatrywanych kryteriów wieloetapowych dla pierwszej z rozpatrywanych realizacji są większe lub równe od wartości odpowiednich kryteriów wieloetapowych dla drugiej z nich oraz przynajmniej w jednym przypadku ma miejsce silna nierówność. Jako rozwiązania optymalne wektorowo przyjmujemy realizacje Pareto-optymalne (w przestrzeni kryterialnej) lub sprawne (w przestrzeni decyzyjnej), czyli takie, dla których nie istnieje inna realizacja procesu lepsza od rozpatrywanej.

Rozwiązanie dynamicznego zadania wektorowej maksymalizacji polega na znalezieniu wszystkich sprawnych realizacji procesu. W tym celu możemy wykorzystać zasadę optymalności Bellmana w wersji wektorowej [Trzaskalik, 1990; 1998]. W przypadku dużych rozmiarów rozpatrywanego zadania zagadnienie to może być trudne rachunkowo. Zbiór realizacji sprawnych może być bardzo obszerny, co powoduje trudności we wspomaganiu decyzji. Dlatego też, rozwiązując zagadnienia praktyczne, wygodnie jest dokonać skalaryzacji problemu lub też wykorzystać podejście interaktywne. Obydwie te możliwości można też wykorzystać równocześnie.

Podejście hierarchiczne polega na określeniu hierarchii kryteriów występujących w rozpatrywanym problemie decyzyjnym, a następnie sekwencyjnym rozwiązywaniu kolejnych problemów jednokryterialnych. Zbiorem rozwiązań dopuszczalnych kolejnego problemu jest zbiór rozwiązań optymalnych zadania poprzedniego, biorącego pod uwagę ważniejsze kryterium.

Bardzo często wykorzystywane jest również podejście quasi-hierarchiczne. Oprócz rozwiązań optymalnych, w kolejnym zadaniu uwzględniamy również te rozwiązania, które są prawie optymalne, czyli takie, dla których wartość rozpatrywanego kryterium mieści się w zadanym przez decydenta przedziale tolerancji. Przy wykorzystaniu tego podejścia w dyskretnych problemach dynamicznych istotna staje się wspominana powyżej możliwość generowania kolejnych rozwiązań prawie optymalnych.

Podejście interaktywne we wspomaganiu rozwiązywania problemów wielokryterialnych stosowane jest począwszy od lat 70. XX wieku [Benayoun i inni, 1971; Steuer, 1977; Miettinen i Makela, 2000; Nowak, 2008; Özpeynirci i inni, 2017]. Może ono być również wykorzystywane w przypadku poszukiwania rozwiązania końcowego w wielokryterialnych, wieloetapowych procesach decyzyjnych [Trzaskalik, 1990]. Polega ono na wzajemnej interakcji pomiędzy decydentem i wspomagającym go analitykiem, wyposażonym w odpowiednie informatyczne narzędzia optymalizacyjne. W kolejnym kroku procedury interaktywnej analityk przedstawia decydentowi aktualnie proponowaną sprawną realizację procesu. Ten ocenia ją ze względu na rozpatrywane kryteria. W przypadku akceptacji przez decydenta proponowanej realizacji procedura zostaje zakończona i rozwiązaniem końcowym staje się ostatnie rozwiązanie, proponowane przez analityka. W przypadku braku takiej akceptacji decydent określa kierunek dalszych poszukiwań, wskazując, które kryteria wieloetapowe mają być poprawione kosztem innych kryteriów. Zazwyczaj istnieje również możliwość cofnięcia się do poprzedniego kroku, gdy okaże się, że wskazany uprzednio kierunek poprawy nie przyniósł akceptowanych przez decydenta wyników. Jednocześnie w każdym kroku procedury interaktywnej istnieje możliwość rozpoczęcia obliczeń od początku lub też odstąpienia od uzyskania rozwiązania końcowego w taki właśnie sposób.

2. Dynamiczny model zarządzania zdolnością produkcyjną

Problem planowania zdolności produkcyjnej można opisać za pomocą modelu programowania dynamicznego. W tym celu analizowany horyzont planowania podzielimy na T etapów. Przez y_t oznaczamy zmienną stanu określającą, jaką zdolnością produkcyjną dysponuje organizacja na początku etapu t ($t = 1, \ldots, T$). Z kolei zmienna decyzyjna x_t wyraża wielkość, o jaką zdolność ta jest zwiększana w etapie t. Przez $y_T + 1$ oznaczamy stan procesu na końcu ostatniego etapu, czyli w analizowanym wypadku – zdolność produkcyjną na koniec okresu objętego planem. Funkcja przejścia $\Omega_t(y_t, x_t)$ opisuje relację zachodzącą pomiędzy stanem, w jakim znajduje się analizowany proces decyzyjny w etapie t, oraz decyzją, która została w tym etapie podjęta, a stanem, w jakim znajdzie się na początku etapu $t + 1$. W analizowanym problemie ma ona postać następującą:

$$y_{t+1} = \Omega_t(y_t, x_t) = y_t + x_t \text{ dla t} \in \overline{1,T} \tag{1}$$

Zakładamy, że na podstawie analizy ograniczeń technicznych, kapitałowych oraz organizacyjnych określono, w jakim zakresie możliwe jest zwiększenie zdolności produkcyjnej w kolejnych etapach rozpatrywanego procesu, a ponadto, że dostępne są na tyle precyzyjne

prognozy zapotrzebowania, że uzasadnione jest zastosowanie modelu deterministycznego.

Przez **D** oznaczymy zbiór wszystkich możliwych realizacji procesu, z których każda opisywana jest przez stany, w jakich proces znajduje się na początku każdego etapu, oraz decyzje, jakie w tych stanach są podejmowane:

$$\mathbf{D} = \{d = (x_1,y_1,x_2,y_2,\ldots,x_T,y_T): \forall_{t\in\overline{1,T}}\, y_{t+1} = \grave{U}_t(y_t, x_t)\} \qquad (2)$$

Głównym celem, jaki formułowany jest przez decydentów rozważających podjęcie działań w celu podwyższenia zdolności produkcyjnej, jest maksymalizacja dochodów przedsiębiorstwa. Warto jednak zauważyć, że nie jest to jedyna przesłanka skłaniająca przedsiębiorców do podjęcia tego typu decyzji. Zdają sobie oni sprawę, że rezygnacja z ekspansji skutkuje brakiem możliwości pełnego zaspokojenia zapotrzebowania klientów, a w efekcie ryzyko przejęcia części rynku przez konkurentów. Z drugiej strony nadmierne zwiększenie zdolności produkcyjnej prowadzi do nieefektywnego wykorzystania zasobów, co negatywnie wpływa na wyniki finansowe firmy. W proponowanym modelu przyjmować zatem będziemy, że decydent zainteresowany jest maksymalizacją wartości trzech kryteriów:

1) f_1 – suma zdyskontowanych przepływów pieniężnych,
2) f_2 – średni poziom zaspokojenia zapotrzebowania zgłaszanego przez klientów w okresie objętym planowaniem,
3) f_3 – średni poziom wykorzystania zdolności produkcyjnej w okresie objętym planowaniem.

3. Procedura interaktywna

Przyjmijmy następujące oznaczenia:

K – liczba analizowanych kryteriów,

$\mathbf{D}^{(l)}$ – zbiór realizacji procesu rozważanych w iteracji l,

$d^{(l)}$ – realizacja proponowana decydentowi w iteracji l,

$\mathbf{f}^*$ – wektor idealny – jego składowe określają wartości optymalne poszczególnych kryteriów uzyskiwane jako rozwiązania jednokryterialnych zadań programowania dynamicznego,

$\overline{\mathbf{f}}$ – wektor optymistyczny – jego składowe określają najlepsze wartości poszczególnych kryteriów uzyskane w dotychczas rozważanych rozwiązaniach,

$\underline{\mathbf{f}}$ – wektor pesymistyczny – wektor grupujący najgorsze wartości poszczególnych kryteriów uzyskane w dotychczas rozważanych rozwiązaniach,

$\hat{\mathbf{f}}$ – wektor wartości satysfakcjonujących – wektor grupujący wartości kryteriów, które decydent określił jako satysfakcjonujące.

Rozwiązanie końcowe problemu jest wyznaczane w sposób następujący:

Faza wstępna:

1. Prosimy decydenta o zdefiniowanie hierarchii kryteriów.
2. Rozwiązujemy K jednokryterialnych zadań programowania dynamicznego, w których kolejno optymalizujemy wartości poszczególnych kryteriów i wyznaczamy wektor $\mathbf{f}^*$.
3. Wyznaczamy optymalne realizacje procesu ze względu na kryterium, które decydent umieścił na najwyższym poziomie hierarchii. Dla każdej z nich obliczamy wartości wszystkich kryteriów. Do zbioru $\mathbf{D}^{(1)}$ włączamy te spośród realizacji zapewniających uzyskanie wartości optymalnej najważniejszego kryterium, które są niezdominowane.
4. Jako wartości początkowe współrzędnych wektora $\hat{\mathbf{f}}$ przyjmujemy najgorsze wartości kryteriów uzyskiwane dla realizacji włączonych do zbioru $\mathbf{D}^{(1)}$.
5. Przyjmujemy $l = 1$.

Iteracja *l*:

1. Wyznaczamy wektory $\bar{\mathbf{f}}$ oraz $\underline{\mathbf{f}}$, przyjmując za ich współrzędne odpowiednio najlepsze i najgorsze wartości kryteriów uzyskiwane dla realizacji włączonych do zbioru $\mathbf{D}^{(l)}$.
2. Określamy realizację $d^{(l)}$, która zostanie zaproponowana decydentowi. W tym celu sprawdzamy, które spośród rozwiązań ze zbioru $\mathbf{D}^{(l)}$ gwarantują uzyskanie takich wartości kryteriów, które nie są gorsze niż wartości zapisane w wektorze $\hat{\mathbf{f}}$. Wyznaczając spośród nich kolejną propozycję dla decydenta, kierujemy się hierarchią kryteriów.
3. Przedstawiamy decydentowi realizację $d^{(l)}$ oraz wektory $\mathbf{f}^*$, $\bar{\mathbf{f}}$, $\underline{\mathbf{f}}$ oraz $\hat{\mathbf{f}}$. Pytamy decydenta, czy wyniki uzyskiwane dla proponowanej realizacji procesu uważa za satysfakcjonujące. Jeżeli odpowiedź brzmi *Tak*, kończymy procedurę, przyjmując za rozwiązanie końcowe realizację $d^{(l)}$.
4. Prosimy decydenta o ponowne zdefiniowanie wektora $\hat{\mathbf{f}}$.
5. Jeżeli wartość satysfakcjonująca dla kryterium uznawanego przez decydenta za najważniejsze nie jest gorsza niż dotychczasowa wartość pesymistyczna tego kryterium, to przechodzimy do kroku 7.
6. Korzystając z algorytmu wyznaczania rozwiązań prawie optymalnych, identyfikujemy wszystkie realizacje pozwalające na uzyskanie wartości satysfakcjonującej dla najważniejszego kryterium. Dla każdej z nich obliczamy wartości wszystkich kryteriów. Do zbioru $\mathbf{D}^{(l)}$ dodajemy te spośród nowo wyznaczonych realizacji, które nie są zdominowane przez jakąkolwiek inną spośród jak dotąd wyznaczonych.
7. Jeżeli w zbiorze $\mathbf{D}^{(l)}$ istnieją realizacje, dla których wartości wszystkich kryteriów nie są gorsze od wartości uznanych przez decydenta

za satysfakcjonujące, przyjmujemy $\mathbf{D}^{(l+1)} = \mathbf{D}^{(l)}$, $l = l + 1$ i przechodzimy do kolejnej iteracji (krok 1).

8. Informujemy decydenta, że nie ma możliwości wyznaczenia realizacji zapewniających uzyskanie wartości satysfakcjonujących dla wszystkich kryteriów. Następnie pytamy go, czy jest skłonny dokonać zmiany wartości satysfakcjonującej dla przynajmniej jednego kryterium. Jeżeli decydent udzielił odpowiedzi pozytywnej, wracamy do kroku (4), w przeciwnym wypadku kończymy rozwiązywanie problemu, przyjmując, że wyznaczenie rozwiązania spełniającego wymagania decydenta nie jest możliwe.

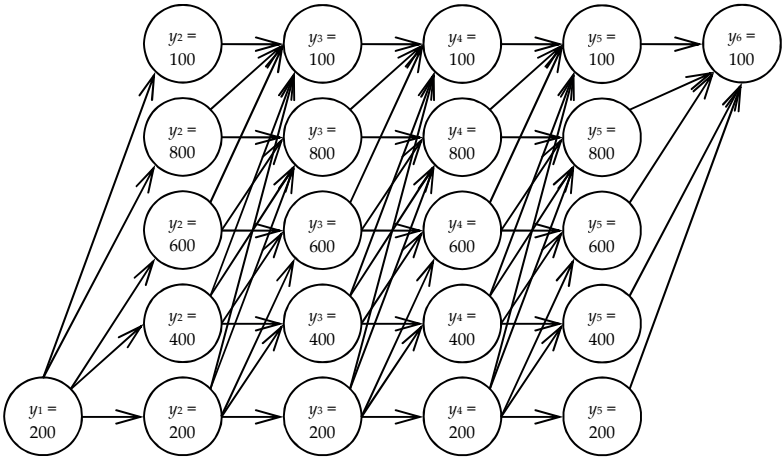
Sposób wykorzystania procedury ilustruje poniższy przykład liczbowy.

4. Przykład ilustracyjny

Przedsiębiorstwo X wprowadza na rynek nowy produkt. Zasoby, którymi aktualnie dysponuje, pozwalają na wytworzenie 200 jednostek produktu rocznie. Opracowane prognozy wskazują, że zapotrzebowanie na produkt w ciągu najbliższych pięciu lat będzie kształtowało się na następującym poziomie: rok 1 – 140 jednostek, rok 2 – 420 jednostek, rok 3 – 820 jednostek, rok 4 – 950 jednostek, rok 5 – 990 jednostek. Oczekuje się, że począwszy od roku 6 przez kolejnych 5 lat zapotrzebowanie będzie się utrzymywało na poziomie ok. 1000 jednostek. Firma zamierza osiągnąć pełną zdolność produkcyjną na koniec piątego roku. Ze względu na wykorzystywaną technologię wielkość, o jaką może być podwyższona zdolność produkcyjna, musi być wielokrotnością 200 jednostek.

Analizowany problem można przedstawić w postaci grafu, w którym węzły odpowiadają poszczególnym stanom procesu, zaś łuki – decyzjom, które mogą być podjęte w każdym z nich. Stan procesu decyzyjnego na początku etapu pierwszego jest określony przez zdolność produkcyjną, jaką firma aktualnie dysponuje. Jedynym stanem dopuszczalnym w etapie pierwszym jest stan $y_1 = 200$. Zakładamy, że na koniec etapu 5 osiągnięta zostanie pełna zdolność produkcyjna 1000 jednostek. Tym samym stan procesu na początek etapu 6 będzie wynosił $y_6 = 1000$. W etapach od 1 do 5 zdolność produkcyjna może być utrzymana na niezmienionym poziomie lub zwiększona o 200, 400, 600 lub 800 jednostek. Tym samym na początku etapów od 2 do 5 zmienna stanu y_t może przyjmować wartość 200, 400, 600, 800 lub 1000. Graf procesu przedstawia rysunek 1. Wierzchołki grafu odpowiadają stanom dopuszczalnym na początku każdego etapu, zaś łuki – decyzjom dopuszczalnym.

Rysunek 1. Graf procesu



Źródło: Opracowanie własne.

W tablicy 1 przedstawiono wartość zdyskontowanych przepływów pieniężnych uzyskiwanych w trakcie pierwszych pięciu lat realizacji projektu. Wyliczając je, uwzględniono:

- przychody ze sprzedaży,
- koszty produkcji, które zależą zarówno od liczby wyprodukowanych wyrobów, jak też zainstalowanej zdolności produkcyjnej,
- koszty związane z podwyższeniem zdolności produkcyjnej, na które składają się koszty stałe, niezależne od wielkości, o jaką zdolność jest podwyższana, oraz koszty zmienne.

Tablica 1. Wartości zdyskontowanych przepływów pieniężnych

Stan	Decyzja	Etap				
y_t	x_t	1	2	3	4	5
200	0	480	1 050	950	870	×
	200	–2 980	–2 100	–1 910	–1 730	×
	400	–5 290	–4 200	–3 820	–3 470	×
	600	–7 600	–6 300	–5 730	–5 200	×
	800	–9 910	–8 400	–7 630	–6 940	8640
400	0	×	2 100	1 910	1 730	×
	200	×	–1 050	–950	–870	×
	400	×	–3 150	–2 860	–2 600	×
	600	×	–5 250	–4 770	–4 340	11 000
600	0	×	1 320	2 860	2 600	×
	200	×	–1 830	0	0	×
	400	×	–3 930	–1 910	–1 730	13 370

Stan y_t	Decyzja x_t	Etap				
		1	2	3	4	5
800	0	×	270	3 820	3 470	×
	200	×	–2 880	950	870	15 740
1000	0	×	–780	3 030	3 940	18 870

Źródło: Opracowanie własne.

Szczegóły dotyczące wyliczania wartości zdyskontowanych przepływów pieniężnych dla analizowanego problemu znaleźć można w pracy [Nowak, 2015].

Biorąc pod uwagę zainstalowaną zdolność produkcyjną oraz przewidywany popyt dla poszczególnych etapów, obliczono wskaźniki poziomu zaspokojenia zapotrzebowania oraz poziomu wykorzystania zdolności produkcyjnej w latach 1–5 (tab. 2 i 3). Wyznaczając je, porównywano wielkość zdolności produkcyjnej na początku każdego etapu z wielkością prognozowanego zapotrzebowania.

Wskaźnik poziomu zaspokojenia zapotrzebowania można obliczyć za pomocą następującej formuły:

$$\delta_t = \frac{\min\{Z_t, y_t\}}{Z_t} \tag{3}$$

gdzie: Z_t – prognozowana wielkość zapotrzebowania w etapie t, y_t – zdolność produkcyjna pozostająca w dyspozycji przedsiębiorstwa w etapie t.

Z kolei do wyliczenia wskaźnika poziomu wykorzystania zdolności produkcyjnej wykorzystano wzór postaci:

$$\delta_t = \frac{\min\{Z_t, y_t\}}{y_t} \tag{4}$$

Jak łatwo zauważyć, wartości obu wskaźników zależą wyłącznie od stanu, w jakim znalazł się proces na początku danego etapu.

Obliczając wartość kryterium f_1 dla etapów od t_1 do t_2, sumujemy wartości uzyskiwane w poszczególnych etapach. Z kolei wartości kryteriów f_2 oraz f_3 dla tego samego okresu wyliczamy jako średnią z odpowiednich wartości etapowych.

Tablica 2. Wskaźniki poziomu zaspokojenia zapotrzebowania

Stan y_t	Etap				
	1	2	3	4	5
1000	1,0000	0,4762	0,2439	0,2105	0,2020
2000	×	0,9524	0,4878	0,4211	0,4040
3000	×	1,0000	0,7317	0,6316	0,6061
4000	×	1,0000	0,9756	0,8421	0,8081
5000	×	1,0000	1,0000	1,0000	1,0000

Źródło: Opracowanie własne.

Tablica 3. Wskaźniki poziomu wykorzystania zdolności produkcyjnej

Stan y_t	Etap				
	1	2	3	4	5
1000	0,7000	1,0000	1,0000	1,0000	1,0000
2000	×	1,0000	1,0000	1,0000	1,0000
3000	×	0,7000	1,0000	1,0000	1,0000
4000	×	0,5250	1,0000	1,0000	1,0000
5000	×	0,4200	0,8200	0,9500	0,9900

Źródło: Opracowanie własne.

Wyznaczenie rozwiązania problemu przy wykorzystaniu proponowanej procedury interaktywnej przebiega następująco:

Faza wstępna:

1. Decydent definiuje następującą hierarchię kryteriów: f_1, f_2, f_3. Oznacza to, że za najważniejszy cel uznał maksymalizację wartości bieżącej netto, na drugim miejscu umieścił maksymalizację średniego poziomu zaspokojenia zapotrzebowania, a za najmniej istotną uznał maksymalizację średniego poziomu wykorzystania zdolności produkcyjnej.
2. Rozwiązujemy trzy jednokryterialne zadania programowania dynamicznego, w których kolejno optymalizujemy wartości poszczególnych kryteriów. Wektor wartości idealnych ma postać: $\mathbf{f}^*$ = [17 940; 1,0000; 0,9400].
3. Wyznaczamy optymalne realizacje procesu ze względu na kryterium maksymalizacji wartości bieżącej netto. Jedyna realizacja zapewniająca uzyskanie wartości tego kryterium na poziomie 17 940 jest następująca:
   $y_1 = 200$, $x_1 = 0$, $y_2 = 200$, $x_2 = 600$, $y_3 = 800$, $x_3 = 0$, $y_4 = 800$, $x_4 = 0$, $y_5 = 800$, $x_5 = 200$, $y_6 = 1000$.
   Realizację tą oznaczymy jako d_1. Jej zastosowanie oznacza, że należy zwiększyć zdolność produkcyjną o 600 jednostek w roku drugim, a następnie o 200 jednostek w roku piątym. Wartości poszczególnych kryteriów uzyskiwanych dla tej realizacji są równe: $f_1(d_1)$ = 17 940; $f_2(d_1) = 0{,}8904$; $f_3(d_1) = 0{,}9280$. Ponieważ tylko jedna realizacja zapewnia uzyskanie wartości optymalnej dla najważniejszego kryterium (f_1), jest ona niezdominowana i zostaje jako jedyna włączona do zbioru $\mathbf{D}^{(1)}$.
4. Jako wartości początkowe współrzędnych wektora $\hat{\mathbf{f}}$ przyjmujemy najgorsze wartości kryteriów uzyskiwane dla realizacji włączonych do zbioru $\mathbf{D}^{(1)}$: $\hat{\mathbf{f}}$ [17 940; 0,8904; 0,9280].
5. Przyjmujemy $l = 1$.

Iteracja 1:

1. Wyznaczamy wektory $\bar{\mathbf{f}}$ oraz $\underline{\mathbf{f}}$:
   $\bar{\mathbf{f}}$ = [17 940; 0,8904; 0,9280, $\underline{\mathbf{f}}$ = [17 940; 0,8904; 0,9280].
2. Jako pierwszą propozycję dla decydenta $d^{(1)}$ przyjmujemy realizację d_1. Jest to jedyna realizacja należąca do zbioru $\mathbf{D}^{(1)}$, a wartości kryteriów, jakie są dla niej uzyskiwane, nie są gorsze niż wartości satysfakcjonujące zdefiniowane w fazie wstępnej procedury.
3. Przedstawiamy decydentowi realizację $d^{(1)}$ oraz wartości idealne (wektor $\mathbf{f}^*$), wartości optymistyczne (wektor $\bar{\mathbf{f}}$), wartości pesymistyczne (wektor $\underline{\mathbf{f}}$) oraz wartości satysfakcjonujące (wektor $\hat{\mathbf{f}}$) – tablica 4.

Tablica 4. Informacje prezentowane decydentowi w iteracji 1

Rodzaj informacji	Kryterium f_1	Kryterium f_2	Kryterium f_3
Proponowana realizacja	17 940	0,8904	0,9280
Wartości idealne	17 940	1,0000	0,9400
Wartości optymistyczne	17 940	0,8904	0,9280
Wartości pesymistyczne	17 940	0,8904	0,9280
Wartości satysfakcjonujące	17 940	0,8904	0,9280

Źródło: Opracowanie własne.

Ponieważ decydent uznaje, że proponowane rozwiązanie nie jest satysfakcjonujące, przechodzimy do następnego kroku.

4. Decydent poproszony o ponowne zdefiniowanie wektora $\hat{\mathbf{f}}$ podaje następujące wartości satysfakcjonujące:
   $\hat{\mathbf{f}}$ = [17 700; 0,9500; 0,9000].
5. Ponieważ wartość satysfakcjonująca, którą decydent podał dla kryterium f_1, jest gorsza od dotychczasowo przyjmowanej, przechodzimy do kroku 6.
6. Korzystając z algorytmu wyznaczania rozwiązań prawie optymalnych, identyfikujemy kolejne realizacje zapewniające uzyskanie wartości satysfakcjonującej dla kryterium f_1. W efekcie uzyskujemy dwie kolejne realizacje d_2 oraz d_3. Wartości kryteriów uzyskiwane dla tych realizacji są następujące:
   $f_1(d_2)$ = 17 920; $f_2(d_2)$ = 0,8952; $f_3(d_2)$ = 0,8920,
   $f_1(d_3)$ = 17 740; $f_2(d_3)$ = 0,8588; $f_3(d_3)$ = 0,9380.
   Ponieważ żadna z poniższych realizacji nie jest zdominowana przez którąkolwiek z pozostałych dotąd wyznaczonych, dodajemy je do zbioru $\mathbf{D}^{(1)}$.
7. Analizując wyniki uzyskiwane dla realizacji ze zbioru $\mathbf{D}^{(2)}$, stwierdzamy, że żadna z nich nie zapewnia uzyskania wartości satysfakcjonujących dla wszystkich kryteriów. Wobec powyższego przechodzimy do kroku 8.

8. Decydent poinformowany, że nie jest możliwe wyznaczenie realizacji zapewniającej uzyskanie wartości satysfakcjonujących dla wszystkich kryteriów decyduje się dokonać modyfikacji wartości satysfakcjonujących. Przechodzimy do kroku 4.
4. Decydent poproszony o ponowne zdefiniowanie wektora $\hat{\mathbf{f}}$ podaje następujące wartości satysfakcjonujące:
   $\hat{\mathbf{f}}$ = [17 700; 0,9500; 0,9000].
5. Ponieważ wartość satysfakcjonująca, którą decydent podał dla kryterium f_1 jest gorsza od dotychczasowo przyjmowanej, przechodzimy do kroku 6.
6. Korzystając z algorytmu wyznaczania rozwiązań prawie optymalnych, identyfikujemy pięć kolejnych realizacji zapewniających uzyskanie wartości kryterium f_1 nie niższej niż 17 000. Dwie z nich są zdominowane. W efekcie do zbioru $\mathbf{D}^{(1)}$ dodajemy cztery nowe realizacje d_4, d_5, d_6 oraz d_7. Wartości kryteriów uzyskiwane dla tych realizacji są następujące:
   $f_1(d_4)$ = 17 630; $f_2(d_4)$ = 0,9856; $f_3(d_4)$ = 0,9280,
   $f_1(d_5)$ = 17 610; $f_2(d_5)$ = 0,9905; $f_3(d_5)$ = 0,8920,
   $f_1(d_6)$ = 17 430; $f_2(d_6)$ = 0,9540; $f_3(d_6)$ = 0,9380,
   $f_1(d_7)$ = 17 210; $f_2(d_7)$ = 0,8204; $f_3(d_7)$ = 0,9400.
7. Analizując uzyskane wyniki, stwierdzamy, że dwie realizacje: d_4 oraz d_6 zapewniają uzyskanie wartości satysfakcjonujących dla wszystkich kryteriów. Wobec powyższego przyjmujemy $\mathbf{D}^{(2)} = \mathbf{D}^{(1)}$, $l = 2$ i przechodzimy do kolejnej iteracji.

Iteracja 2:

1. Wyznaczamy wektory $\bar{\mathbf{f}}$ oraz $\underline{\mathbf{f}}$:
   $\bar{\mathbf{f}}$ = [17 940; 0,9905; 0,9400], $\underline{\mathbf{f}}$ = [17 210; 0,8204; 0,8920].
2. Jako kolejną propozycję dla decydenta $d^{(2)}$ przyjmujemy realizację d_4, która jest jedną z dwóch realizacji zapewniających uzyskanie wartości satysfakcjonujących, a jednocześnie pozwala na uzyskanie wyższej wartości kryterium f_1 niż realizacja d_6. Wartości zmiennych stanu oraz zmiennych decyzyjnych dla realizacji d_4 są następujące:
   $y_1 = 200$, $x_1 = 200$, $y_2 = 400$, $x_2 = 400$, $y_3 = 800$, $x_3 = 200$, $y_4 = 1000$, $x_4 = 0$, $y_5 = 1000$, $x_5 = 0$, $y_6 = 1000$.
   Zastosowanie tej realizacji oznacza, że należy zwiększyć zdolność produkcyjną trzykrotnie: o 200 jednostek w roku pierwszym, o 400 jednostek w roku drugim i ponownie o 200 jednostek w roku trzecim. Na koniec roku trzeciego uzyskiwana jest już docelowa zdolność produkcyjna.
3. Przedstawiamy decydentowi realizację $d^{(2)}$ oraz wartości idealne (wektor $\mathbf{f}^*$), wartości optymistyczne (wektor $\bar{\mathbf{f}}$), wartości pesymistyczne (wektor $\underline{\mathbf{f}}$) oraz wartości satysfakcjonujące (wektor $\hat{\mathbf{f}}$) – tablica 5.

Tablica 5. Informacje prezentowane decydentowi w iteracji 2

Rodzaj informacji	Kryterium f_1	Kryterium f_2	Kryterium f_3
Proponowana realizacja	17 630	0,9856	0,9280
Wartości idealne	17 940	1,0000	0,9400
Wartości optymistyczne	17 940	0,9905	0,9400
Wartości pesymistyczne	17 210	0,8204	0,8920
Wartości satysfakcjonujące	17 000	0,9500	0,9000

Źródło: Opracowanie własne.

Decydent może w tym momencie zaakceptować realizację d_4 jako rozwiązanie końcowe lub kontynuować poszukiwanie rozwiązania jeszcze lepiej dopasowanego do jego oczekiwań w sposób analogiczny do opisanego powyżej.

Zakończenie

W rozpatrywanym w niniejszej pracy zagadnieniu dynamicznego planowania zdolności produkcyjnej założono, że dostępne są precyzyjne prognozy zapotrzebowania zdolności produkcyjnej w przyszłych etapach, co umożliwia zastosowania przedstawionego w pracy modelu deterministycznego. Odejście od tego założenia uniemożliwia tego typu modelowanie. Interesujące byłoby rozpatrzenie problemu planowania zdolności produkcyjnej w sytuacji, gdy znamy rozkłady prawdopodobieństwa zapotrzebowania na zdolność produkcyjną w kolejnych etapach lub też gdy zapotrzebowanie to przedstawione byłoby za pomocą liczb rozmytych (trójkątnych lub trapezowych). Innym kierunkiem rozszerzenia przedstawionym w niniejszej pracy rozważań byłoby uwzględnienie metodologii planowania kroczącego. Zagadnienia te warte są podjęcia w przyszłości.

Literatura

Bellman R. (1957), *Dynamic Programming*, Princenton University Press.

Benayoun R., De Montgolfier J., Tergny J., Laritchev O. (1971), *Linear programming with multiple objective functions: Step method* (*STEM*), „Mathematical Programming", Vol. 1, No. 1.

Cheng L., Subrahmanian E., Westerberg A.W. (2004), *Multi-objective Decisions on Capacity Planning and Production–Inventory Control under Uncertainty*, „Industrial & Engineering Chemistry Research", Vol. 43, No. 9.

Elmaghraby S.E. (1970), *The Theory of Networks and Management Science*. *Part 1*, „Management Science", Vol. 17, No. 1.

Fengqi Y., Grossmann I.E., Wassick J.M. (2011), *Multisite Capacity, Production, and Distribution Planning with Reactor Modifications: MILP Model, Bilevel Decomposition Algorithm versus Lagrangean Decomposition Scheme*, „Industrial & Engineering Chemistry Research", Vol. 50, No. 9.

Ganguly S., Sahoo N. C., Das D. (2013), *Multi-objective planning of electrical distribution systems using dynamic programming*, „International Journal of Electrical Power & Energy Systems", Vol. 46.

Geng N., Jiang Z. (2009), *A review on strategic capacity planning for the semiconductor manufacturing industry*, „International Journal of Production Research", Vol. 47, No. 13.

Huang L., Suh I. H., Abraham A. (2011), *Dynamic multi-objective optimization based on membrane computing for control of time-varying unstable plants*. „Information Sciences", Vol. 181, No. 11.

Jafarian-Moghaddam A. R., Ghoseiri K. (2011), *Fuzzy dynamic multi-objective Data Envelopment Analysis model*, „Expert Systems with Applications", Vol. 38, No. 1.

Kannan D., Khodaverdi R., Olfat L., Jafarian A., Diabat A. (2013), *Integrated fuzzy multi criteria decision making method and multi-objective programming approach for supplier selection and order allocation in a green supply chain*, „Journal of Cleaner Production", Vol. 47.

Korhonen P. J., Laakso J. (1986), *A visual interactive method for solving the multiple criteria problem*, „European Journal of Operational Research", Vol. 24, No. 2.

Li W., Guo G., Yue C., Zhao Y. (2010), *Dynamic programming methodology for multi-criteria group decision-making under ordinal preferences*, „Journal of Systems Engineering and Electronics", Vol. 21, No. 6.

Mafakheri F., Breton M., Ghoniem A. (2011), *Supplier selection-order allocation: A two--stage multiple criteria dynamic programming approach*, „International Journal of Production Economics", Vol. 132, No. 1.

Miettinen K., Makela M. M. (2000), *Interactive multiobjective optimization system WWW-NIMBUS on the Internet*, „Computers & Operations Research", Vol. 27, No. 7–8.

Nguyen S., Zhang M., Johnston M., Tan K. C. (2013), *Dynamic multi-objective job shop scheduling: a genetic programming approach*, w: A.S. Uyar, E. Ozcan, N. Urquhart (eds.), *Automated Scheduling and Planning*, Springer, Berlin, Heidelberg.

Nowak M. (2008), *Interaktywne wielokryterialne wspomaganie decyzji w warunkach ryzyka. Metody i zastosowania*, Wydawnictwo Akademii Ekonomicznej im. Karola Adamieckiego w Katowicach, Katowice.

Nowak M. (2015), *Modelowanie decyzji w zarządzaniu operacyjnym*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.

Özpeynirci Ö., Özpeynirci S., Kaya A. (2017), *An interactive approach for multiple criteria selection problem*, „Computers & Operations Research", Vol. 78.

Soroudi A., Caire R., Hadjsaid N., Ehsan M. (2011), *Probabilistic dynamic multi-objective model for renewable and non-renewable distributed generation planning*, „IET Generation, Transmission & Distribution", Vol. 5, No. 11.

Steuer R.E. (1977), *An Interactive Muliple Objective Linear Programming procedure*, „TIMS Studies in the Management Sciences", Vol. 6.

Trzaskalik T. (2016), *MCDM applications of near optimal solutions in dynamic programming*, „Multiple Criteria Decision Making", Vol. 10.

Trzaskalik T. (1998), *Multiobjective Analysis in Dynamic Environmnent*, The Karol Adamiecki University of Economics, Katowice.

Trzaskalik T. (1990), *Wielokryterialne dyskretne programowanie dynamiczne. teoria i zastosowania w praktyce gospodarczej*, Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice.
Wager H.M. (1980), *Badania operacyjne. Zastosowania w zarządzaniu*, PWE, Warszawa.

Streszczenie

W pracy rozważany jest problem zarządzania zdolnością produkcyjną, rozumianą jako maksymalna ilość produktów lub usług, którą organizacja może wytworzyć w określonej jednostce czasu. Celem jest określenie strategii, jaką powinna zastosować firma w warunkach wzrastającego popytu. Rozważany problem sformułowano jako zadanie wielokryterialnego, dyskretnego programowania dynamicznego, przyjmując, że decydent jest zainteresowany jednoczesną maksymalizacją trzech kryteriów: 1) sumy zdyskontowanych przepływów pieniężnych, 2) średniego poziomu zaspokojenia zapotrzebowania zgłaszanego przez klientów, 3) średniego poziomu wykorzystania zdolności produkcyjnej. Zaproponowana w pracy procedura rozwiązania problemu wykorzystuje metodę quasi-hierarchiczną oraz podejście interaktywne. W każdej iteracji decydentowi przedstawiana jest nowa propozycja rozwiązania, który może ją zaakceptować jako rozwiązanie końcowe lub określić, w jaki sposób powinna ona być zmodyfikowana. W pracy zamieszczono przykład liczbowy ilustrujący sposób wykorzystania prezentowanej procedury.

Słowa kluczowe

zarządzanie zdolnością produkcyjną, analiza wielokryterialna, programowanie dynamiczne, podejście interaktywne, metoda quasi-hierarchiczna

Interactive dynamic programming application to capacity production management (Summary)

In the paper capacity planning problem is considered, where capacity is defined as a maximum number of goods or services that the organization can produce during a particular period. The main goal is to determine the strategy that should be applied by organization faced with growing demand. The problem is defined as a multicriteria, discrete dynamic programming model, assuming that the decision maker is going to maximize three criteria: 1) net present value, 2) average customer service level, 3) average capacity utilization. The procedure proposed in the paper uses quasi-hierarchical method and interactive approach. In each solution a new proposal is proposed to the decision maker, who can either accept it as a final solution or specify in what way it should be improved. A numerical example is presented to illustrate the applicability of the procedure.

Keywords

capacity planning, multicriteria analysis, dynamic programming, interactive approach, quasi-hierarchical method

Beata Jackowska*

Metody analizy przeżycia w badaniu zróżnicowania rozkładów czasu istnienia przedsiębiorstw ze względu na rodzaj działalności

Wstęp

Badania rozkładów czasu istnienia przedsiębiorstw oraz poszukiwania czynników determinujących długość okresu funkcjonowania firm za pomocą metod analizy przeżycia sięgają na świecie początków lat 90. [Manjon-Antolin, Arauso-Carod, 2008]. W literaturze polskiej znaleźć można wyniki badań wykazujących, że rodzaj działalności określany według Polskiej Klasyfikacji Działalności (PKD) różnicuje rozkłady czasu istnienia przedsiębiorstw [Gurgul, Zając, 2016; Markowicz, 2012; Ptak-Chmielewska, 2016; Szymański, 2011]. Rodzaj działalności traktowany jest jako zmienna grupująca, a grupy przedsiębiorstw uznane w wyniku przeprowadzonych badań za podobne pod względem ryzyka zaprzestania działalności można połączyć w większe grupy. Jednakże przy grupowaniu sekcji PKD różne zbiory danych, jak też różne algorytmy postępowania, mogą prowadzić do różniących się wyników. Zagadnienie to nie ma jednoznacznego rozwiązania, a w literaturze można znaleźć różne sposoby łączenia wariantów rozważanej cechy (niekiedy w głównej mierze arbitralne) prowadzące do niejednoznacznych wyników.

W niniejszym artykule, na przykładzie danych empirycznych dla przedsiębiorstw z Gdańska, został przedstawiony sposób postępowania w celu uniknięcia przypadkowych decyzji przy wyodrębnianiu grup przedsiębiorstw o zbliżonym ryzyku zaprzestania działalności. Głównym celem badania jest wyłonienie grup przedsiębiorstw, które są zarazem wewnętrznie jednorodne i jednocześnie różnią się między sobą statystycznie istotnie ze względu na rozkład czasu trwania. Uogólniając, rozważany problem badawczy sprowadza się do znalezienia takiego sposobu rekategoryzacji, czy inaczej redukcji liczby wariantów cechy jakościowej, który pozwoli na odpowiednie pogrupowanie jednostek pod względem rozkładu czasu trwania. Statystycznej weryfikacji tej hipotezy można dokonać za pomocą nieparametrycznych testów jednorodności należących do grupy metod analizy przeżycia.

* Prof. UG dr hab., Katedra Statystyki, Wydział Zarządzania, Uniwersytet Gdański, ul. Armii Krajowej 101, 81-824 Sopot, beata.jackowska@ug.edu.pl

Analiza przeżycia jest dziedziną nauki zajmującą się badaniem rozkładów czasu trwania[1]. Do analizy czasu upływającego od zdarzenia początkowego do wystąpienia zdarzenia końcowego rozwija się odrębne metody statystyczne z powodu występowania obserwacji cenzurowanych, jak też zazwyczaj niespełnienia założenia normalności rozkładu. W badaniach empirycznych zwykle nie dla wszystkich jednostek udaje się odnotować zdarzenie końcowe. Mamy wówczas do czynienia z cenzurowaniem prawostronnym[2]. Jednak dla jednostek, których obserwacja została przerwana przed wystąpieniem badanego zdarzenia, posiadamy częściową informację, że istniały dłużej od okresu znajdowania się w polu obserwacji. Tę częściową informację można wykorzystać przy określaniu rozkładu czasu trwania.

1. Opis badania czasu istnienia przedsiębiorstw w Gdańsku

Niniejsze badanie jest kontynuacją przedstawionego w artykule [Jackowska, 2015] badania czasu istnienia przedsiębiorstw, w którym wykorzystano analizę kohortową. Analizie podlega kohorta przedsiębiorstw powstałych w Gdańsku w 2009 r. Dane pochodzą z Urzędu Statystycznego w Gdańsku z rejestru REGON. Okres badania obejmuje lata 2009–2014, tzn. przedsiębiorstwa były śledzone w rejestrze od momentu rozpoczęcia działalności w 2009 r. do momentu zakończenia działalności lub zakończenia badania w dniu 31.12.2014. Dla każdego przedsiębiorstwa dostępne są informacje:

- dokładne daty zarejestrowania i rozpoczęcia działalności,
- dokładne daty wyrejestrowania i zakończenia działalności,
- forma prawna,
- przewidywana w momencie rejestracji liczba pracujących w czterech kategoriach: do 9, 10–49, 50–249, 250 i więcej osób,
- rodzaj głównej działalności według PKD.

W tablicy 1 przedstawiono strukturę przedsiębiorstw, które rozpoczęły działalność w Gdańsku w 2009 r. według wielkości oraz formy prawnej. Wielkość mierzona jest liczbą pracowników i przyjęto zgodnie z klasyfikacją GUS rozróżnienie na przedsiębiorstwa: małe – poniżej 50 pracowników, średnie – 50–249 oraz duże – zatrudniające 250 i więcej. Według tej klasyfikacji mikroprzedsiębiorstwa liczą poniżej 10 pracowników i zawierają się w grupie małych przedsiębiorstw [*Warunki…*, 2015, s. 16]. Jednak coraz

[1] W zależności od obszaru zastosowań analiza przeżycia może być nazywana analizą historii zdarzeń, analizą trwania, analizą przejścia czy też analizą niezawodności.

[2] Cenzurowanie prawostronne występuje najczęściej, gdy badanie dobiegło końca zanim nastąpiła realizacja zdarzenia dla wszystkich jednostek lub niektóre jednostki zostały utracone w trakcie obserwacji, lub czas trwania pewnych jednostek zakończył się z powodu realizacja ryzyka konkurującego [Rossa, 2005, s. 13–15].

częściej można spotkać w literaturze podział rozłączny na mikroprzedsiębiorstwa poniżej 10 pracowników oraz małe przedsiębiorstwa od 10 do 49, które nazywane są przez GUS „pozostałymi małymi".

W Gdańsku w 2009 r. działalność rozpoczęły 4783 przedsiębiorstwa. Dominowały mikroprzedsiębiorstwa (98,7%), nie powstało żadne duże przedsiębiorstwo, a niewielka liczba pozostałych małych oraz średnich przedsiębiorstw nie pozwoliła na analizę wpływu wielkości firmy na ryzyko zakończenia działalności. Osoby fizyczne prowadziły 88,0% działalności gospodarczych.

Tablica 1. Struktura przedsiębiorstw, które rozpoczęły działalność w Gdańsku w 2009 r. według przewidywanej liczby pracujących oraz formy prawnej

Forma prawna	Przewidywana liczba pracujących			
	do 9 osób	10–49 osób	50–249 osób	Ogółem
osoba fizyczna prowadząca działalność gospodarczą	4174	36	1	4211
jednostka organizacyjna niemająca osobowości prawnej	326	10	0	336
osoba prawna	220	14	2	236
Ogółem	4720	60	3	4783

Źródło: Badania własne na podstawie bazy REGON.

Przedmiotem niniejszego badania są małe przedsiębiorstwa (w tym mikroprzedsiębiorstwa) osób fizycznych powstałe w Gdańsku w 2009 r. Ze względu na jakość danych w REGON badanie zostało ograniczone do przedsiębiorstw osób fizycznych. Od 31 marca 2009 r. wprowadzono bowiem dla osób fizycznych jeden formularz rejestracji (wyrejestrowania) w REGON, ZUS oraz urzędzie skarbowym. Wcześniej nie wszystkie podmioty wywiązywały się z ustawowego obowiązku wyrejestrowania z REGON. Należy oczekiwać poprawy jakości bazy REGON odnośnie do rejestracji zdarzeń od 2009 r.[3]

W badaniu zdarzeniem początkowym jest rozpoczęcie działalności gospodarczej, natomiast zdarzeniem końcowym – zakończenie działalności. Obserwacją cenzurowaną jest przedsiębiorstwo aktywne w dniu zakończenia obserwacji, tj. 31.12.2014. W Gdańsku w 2009 r. działalność rozpoczęło 4210 małych (w tym mikro) przedsiębiorstw osób fizycznych, 5 lat przetrwało 2030 (48%), a w momencie zakończenia obserwacji w dniu 31.12.2014 aktywnych było jeszcze 1898 przedsiębiorstw, czyli w badaniu

[3] Więcej na ten temat w: [Jackowska, 2015].

45% obserwacji było cenzurowanych. Najkrótszy czas cenzurowany wynosił 5 lat (dla firm powstałych w grudniu 2009 r.), a najdłuższy 6 lat (dla firm powstałych w styczniu 2009 r.).

Podstawową charakterystykę badanych przedsiębiorstw według sekcji PKD przedstawiono w tablicy 2. W badaniu nie wystąpiły trzy sekcje (D, T, U), a cztery (A, B, E, O) miały niewielką liczebność. Wstępna analiza pokazuje już, że występują znaczne różnice pomiędzy sekcjami pod względem odsetka zlikwidowanych przedsiębiorstw oraz pod względem mediany czasu trwania. Najmniejszy odsetek likwidacji w okresie badania wystąpił w sekcji Q „Opieka zdrowotna i pomoc społeczna" (31,4%), a największy w sekcji K „Działalność finansowa i ubezpieczeniowa" (75,1%)[4]. Medianę czasu trwania można było wyznaczyć w tych grupach, w których co najmniej 50% czasów nie było cenzurowanych. W najbardziej narażonej sekcji K mediana wynosiła 2,6 roku, natomiast aby wyznaczyć medianę w najmniej narażonej sekcji Q, należałoby poprowadzić badanie przypuszczalnie jeszcze przez kilka lat, do czasu aż połowa podmiotów zakończy działalność.

2. Wstępna rekategoryzacja jakościowej zmiennej objaśniającej „rodzaj działalności"

Sekcja PKD zostanie potraktowana jako zmienna jakościowa objaśniająca podatność na ryzyko zakończenia działalności. Celem badania jest wyodrębnienie grup przedsiębiorstw o zbliżonym ryzyku zaprzestania działalności, co jest równoznaczne z połączeniem odpowiednich wariantów, tzw. rekategoryzacją, grupującej cechy jakościowej „sekcja PKD".

Tablica 2. Charakterystyka przedsiębiorstw, które rozpoczęły działalność w Gdańsku w 2009 r. według rodzaju działalności

Lp.	Sekcja PKD*		Liczba rozpoczynających w 2009 r.	Odsetek zlikwidowanych do 31.12.2014	Mediana czasu trwania w latach
1	A	Rolnictwo, leśnictwo, łowiectwo i rybactwo	7	42,86	–
2	B	Górnictwo i wydobywanie	4	25,00	–
3	C	Przetwórstwo przemysłowe	444	61,71	4,26
4	E	Dostawa wody; gospodarowanie ściekami i odpadami oraz działalność związana z rekultywacją	6	33,33	–

[4] Nie biorąc pod uwagę sekcji A, B, E, O o zbyt małych liczebnościach.

Lp.	Sekcja PKD*		Liczba rozpoczynających w 2009 r.	Odsetek zlikwidowanych do 31.12.2014	Mediana czasu trwania w latach
5	F	Budownictwo	466	59,87	4,62
6	G	Handel hurtowy i detaliczny; naprawa pojazdów samochodowych, włączając motocykle	929	61,89	3,52
7	H	Transport i gospodarka magazynowa	289	46,37	–
8	I	Działalność związana z zakwaterowaniem i usługami gastronomicznymi	145	66,21	3,11
9	J	Informacja i komunikacja	158	44,94	–
10	K	Działalność finansowa i ubezpieczeniowa	245	75,10	2,63
11	L	Działalność związana z obsługą rynku nieruchomości	58	34,48	–
12	M	Działalność profesjonalna, naukowa i techniczna	545	38,72	–
13	N	Działalność w zakresie usług administrowania i działalność wspierająca	144	61,81	4,13
14	O	Administracja publiczna i obrona narodowa; obowiązkowe zabezpieczenia społeczne	1	100,00	–
15	P	Edukacja	168	54,76	4,93
16	Q	Opieka zdrowotna i pomoc społeczna	290	31,38	–
17	R	Działalność związana z kulturą, rozrywką i rekreacją	64	54,69	4,72
18	S	Pozostała działalność usługowa	247	62,35	4,07
		Ogółem	**4210**	**54,92**	**4,84**

* W badaniu nie wystąpiły sekcje: D – Wytwarzanie i zaopatrywanie w energię elektryczną, gaz, parę wodną, gorącą wodę i powietrze do układów klimatyzacyjnych; T – Gospodarstwa domowe zatrudniające pracowników; gospodarstwa domowe produkujące wyroby i świadczące usługi na własne potrzeby; U – Organizacje i zespoły eksterytorialne.
Źródło: Badania własne na podstawie bazy REGON.

Testy statystyczne, które zostaną wykorzystane do porównania rozkładów czasu trwania, wymagają odpowiednio dużych liczebności grup, powyżej 50 jednostek. We wstępnym etapie rekategoryzacji grupy o zbyt małej liczebności zostaną dołączane do podobnych grup. Decyzję można podjąć subiektywnie, kierując się podobieństwem rodzaju działalności lub za pomocą miar jakości podziału.

Przykładowo można tu wykorzystać miarę *weight of ewidence* (*WoE*), informującą o sile predykcyjnej poszczególnych wariantów cechy (atrybutów), obliczaną dla każdego wariantu oddzielnie według wzoru:

$$WoE_i = \left[\ln\left(\frac{p_i}{q_i}\right) \cdot 100 \right] \tag{1}$$

oraz miarę *information value* (*IV*), informującą o sile predykcyjnej całej zmiennej, obliczaną jako suma po wszystkich wariantach cechy:

$$IV = \sum_{i=1}^{k} \left(p_i - q_i\right) \cdot \ln\left(\frac{p_i}{q_i}\right) \tag{2}$$

gdzie p_i – prawdopodobieństwo wystąpienia i-tego wariantu cechy w grupie jednostek, które nie doznały zdarzenia, q_i – prawdopodobieństwo wystąpienia i-tego wariantu cechy w grupie jednostek, które doznały zdarzenia, k – liczba wariantów cechy.

Miara *IV* umożliwia ocenę jakości kategoryzacji zmiennej poprzez porównanie rozkładów danej cechy wśród jednostek, które doznały zdarzenia oraz wśród tych, które nie doznały zdarzenia. W tym przypadku jej wartość wyniosła *IV* = 0,236 (dla k = 18) i wskazuje na dosyć wysoką zdolność predykcyjną zmiennej[5] „Sekcja PKD".

W tablicy 3 posortowano sekcje PKD rosnąco według wartości miary *WoE*. Ujemne wartości świadczą o podwyższonym ryzyku likwidacji, dodatnie o obniżonym w stosunku do średniego ryzyka, dla którego *WoE* = 0. Kierując się podobieństwem działalności, dołączono sekcję O reprezentowaną przez jedną firmę do sekcji N, a następnie, kierując się kryterium *WoE*, sekcję A do J, sekcję E do L oraz sekcję B do Q. Siła predykcyjna zmiennej „Sekcja PKD" po wstępnej rekategoryzacji *IV* = 0,236 (dla k = 14) nie zmieniła się. Miara *IV* została wykorzystana tylko pomocniczo, gdyż nie bierze pod uwagę rozkładu zdarzeń w czasie. Po ustaleniu

[5] Przyjmuje się, że wartość miary *IV*: poniżej 0,02 świadczy o braku zdolności predykcyjnej danej zmiennej, z przedziału 0,02–0,10 o słabej zdolności predykcyjnej, z przedziału 0,10–0,30 o średniej, a powyżej 0,30 silnej. Wartość *IV* powyżej 0,50 powoduje, że dana zmienna dominuje w modelu nad innymi zmiennymi – zaleca się wówczas zbudowanie odrębnych modeli, wykorzystując dominującą zmienną do podziału zbioru danych [Siddiqi, 2006, s. 81–82].

sposobu postępowania ze zbyt mało licznymi grupami, można już przejść do porównania rozkładów czasu trwania w grupach.

Tablica 3. Ryzyko likwidacji według rodzaju działalności przedsiębiorstw, które rozpoczęły działalność w Gdańsku w 2009 r. (sekcje posortowane według miary *WoE*)

Lp.	Sekcja PKD	Liczba rozpoczy--nających w 2009 r.	Liczba aktyw-nych w dniu 31.12. 2014	Liczba zlikwido--wanych do 31.12. 2014	Odsetek zlikwido-wanych do 31.12. 2014	Odsetek aktyw-nych w dniu 31.12. 2014	*WoE*
1	O	1	0	1	0,04	0,00	–
2	K	245	61	184	7,96	3,21	–90,67
3	I	145	49	96	4,15	2,58	–47,52
4	S	247	93	154	6,66	4,90	–30,70
5	G	929	354	575	24,87	18,65	–28,78
6	N	144	55	89	3,85	2,90	–28,40
7	C	444	170	274	11,85	8,96	–28,00
8	F	466	187	279	12,07	9,85	–20,28
9	P	168	76	92	3,98	4,00	0,63
10	R	64	29	35	1,51	1,53	0,93
11	H	289	155	134	5,80	8,17	34,29
12	J	158	87	71	3,07	4,58	40,05
13	A	7	4	3	0,13	0,21	48,50
14	M	545	334	211	9,13	17,60	65,66
15	L	58	38	20	0,87	2,00	83,92
16	E	6	4	2	0,09	0,21	89,05
17	Q	290	199	91	3,94	10,48	97,98
18	B	4	3	1	0,04	0,16	129,59
Ogółem		**4210**	**1898**	**2312**	**100,00**	**100,00**	**0,00**

Źródło: Badania własne na podstawie bazy REGON.

3. Porównanie krzywych przetrwania dla przedsiębiorstw pogrupowanych według rodzaju działalności

Do określenia rozkładów czasu trwania w poszczególnych grupach przedsiębiorstw została wykorzystana funkcja przetrwania[6] $S(x) = P(X > x)$, mówiąca, jakie jest prawdopodobieństwo, że jednostka będzie trwać przez czas dłuższy niż x. Prawdopodobieństwo zdarzenia przeciwnego

[6] Inaczej nazywana funkcją trwania lub w przypadku obiektów ożywionych – funkcją dożycia lub przeżycia.

$F(x) = P(X \leq x)$ nazywane jest dystrybuantą rozkładu. W analizie przeżycia, odmiennie niż w innych dziedzinach, zwyczajowo przyjmuje się do określania rozkładu zamiast dystrybuanty funkcję przetrwania.

Funkcja przetrwania została oszacowana za pomocą estymatora Kaplana-Meiera[7]. W tym celu czas trwania zjawiska dzielony jest losowo na przedziały w momentach wystąpienia badanego zdarzenia. W każdym z przedziałów czasu odrębnie estymuje się wartość funkcji przetrwania $\hat{S}(x) = \hat{S}_m$ dla $x \in [x_m, x_{m+1})$ za pomocą wzoru:

$$\hat{S}_m = \begin{cases} 1 & \text{dla} \quad m = 0 \\ \prod_{j=1}^{m} \left(\frac{N_j - D_j}{N_j} \right) & \text{dla} \quad m = 1, 2, \ldots, \omega - 1 \\ 0 & \text{dla} \quad m = \omega, \text{ gdy w ostatnim przedziale nie ma czasów cenzurowanych} \\ \text{nieokreślony} & \text{dla} \quad m = \omega, \text{ gdy w ostatnim przedziale są czasy cenzurowane} \end{cases} \tag{3}$$

gdzie D_j jest liczbą zdarzeń w przedziale $[x_j, x_{j+1})$; N_j jest liczbą jednostek, które nie doznały zdarzenia przed momentem x_j oraz x_ω jest maksymalnym odnotowanym czasem trwania.

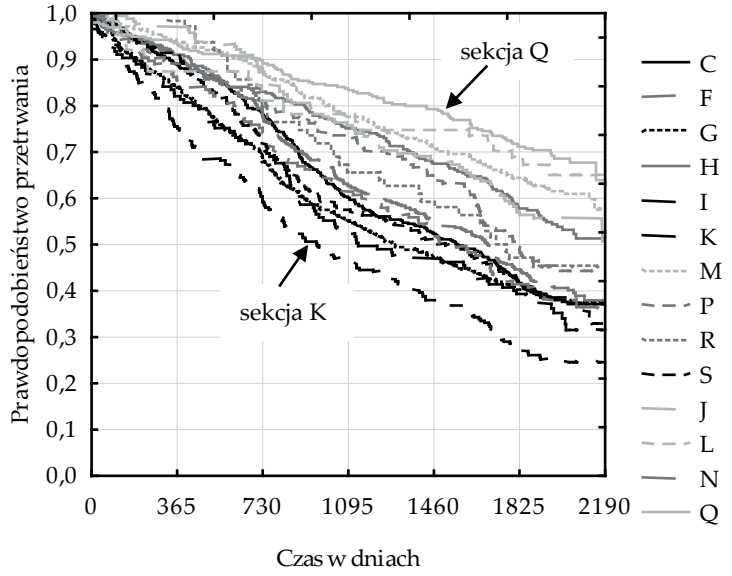
W efekcie zastosowania powyższego estymatora uzyskuje się funkcję nierosnącą, przedziałami stałą, ze skokami w losowych punktach określonych przez każdą z obserwacji niecenzurowanych [Kaplan, Meier, 1958].

Na wykresie (rys. 1) przedstawiono krzywe przetrwania Kaplana-Meiera dla 14 grup przedsiębiorstw otrzymanych po wstępnej rekategoryzacji zmiennej „sekcja PKD". Z wykresu można odczytać, że najkrócej funkcjonowały podmioty prowadzące działalność finansową i ubezpieczeniową (sekcja K – najszybciej malejąca krzywa), a najdłużej oferujące opiekę zdrowotną i pomoc społeczną (sekcja Q – najwolniej malejąca krzywa).

W literaturze z dziedziny analizy przeżycia można znaleźć wiele nieparametrycznych testów jednorodności dla par krzywych przetrwania. Własności tych testów przy różnych założeniach są tematem nadal otwartej dyskusji [Klein, Moeschberger, 1997, s. 191–198; Suciu i inni, 2004, s. 251–262; Jurkiewicz, Wycinka, 2011; Karadeniz, Ercan, 2017]. W praktycznych zastosowaniach szczególnie uciążliwa jest możliwość utraty mocy testu w przypadku krzyżowania się krzywych przetrwania. W artykule [Jurkiewicz, Wycinka, 2011] wykazano za pomocą symulacji, że utrata mocy zależy od rodzaju testu i miejsca przecięcia krzywych przetrwania (na początku, w środku, czy też na końcu rozkładu).

[7] Analiza przebiegu krzywych Kaplana-Meiera w celu redukcji liczby wariantów cechy jakościowej, a także cechy ilościowej znalazła zastosowanie także w scoringu kredytowym [Wycinka, 2015].

Rysunek 1. Funkcje przetrwania przedsiębiorstw pogrupowanych według 14 rodzajów działalności (estymator Kaplana-Meiera)



Źródło: Badania własne na podstawie bazy REGON (obliczenia za pomocą pakietu STATISTICA 12).

W niniejszym badaniu do porównań wykorzystano uogólnienie Gehana testu Wilcoxona (test Wilcoxona według Gehana), inaczej nazywanego testem Breslowa. Test ten jest stosowany do porównywania grup o podobnych rozkładach czasów cenzurowanych (w tym badaniu czasy cenzurowane znajdują się w ostatnim roku obserwacji) i nie wymaga założenia proporcjonalności funkcji intensywności zdarzeń w grupach. W teście tym większa waga jest przykładana do początkowej części rozkładu. Z tego powodu, gdy krzywe przetrwania przecinają się na początku rozkładu, spada moc tego testu i rośnie szansa, że różne, lecz przecinające się krzywe, zostaną uznane za podobne. Aby uniknąć takiej sytuacji, w przypadku gdy test nie wykazał statystycznie istotnych różnic, zastosowano pomocniczo metodę wizualną porównania krzywych na wykresie[8] oraz sprawdzono wyniki w innych tego typu testach (log-rank, Peto-Peto).

Do sprawdzenia jednorodności wielu rozkładów został wykorzystany test statystyczny, będący rozwinięciem uogólnienia Gehana testu Wilcoxona. Test wykazał, że cecha grupująca o $k = 14$ wariantach (atrybutach) jest czynnikiem ryzyka, czyli rozkłady czasu istnienia przedsiębiorstw różnią się ze względu na rodzaj prowadzonej działalności (wartość *p-value*

[8] Należy sprawdzić, czy krzywe przetrwania krzyżują się (w szczególności na początku rozkładu). Jeżeli tak, to czy wynika to z bliskiego położenia i przeplatania się, czy też po skrzyżowaniu krzywe rozbiegają się.

dla testu $p < 0{,}00001$). W następnym etapie badania należy wykonać testy dla każdej pary krzywych przetrwania w celu sprawdzenia, czy występują krzywe nieróżniące się statystycznie istotnie.

W tablicy 4 przedstawiono krytyczne poziomy istotności (*p-value*) w testach dla wszystkich par krzywych Kaplana-Meiera. Wśród sprawdzonych 91 różnych par znalazły się krzywe, które nie różniły się statystycznie istotnie. Przy łączeniu grup bierze się pod uwagę te grupy, dla których nie odrzucono testowanej hipotezy o jednakowych rozkładach, więc nie narażamy się na popełnienie błędu I rodzaju. Z tego powodu poprawkę na wielokrotne testowanie (zwiększającą wartość *p-value*) wystarczy wprowadzić w ostatnim etapie, aby sprawdzić, czy po przeprowadzonej rekategoryzacji zmiennej grupy różnią się statystycznie istotnie we wszystkich parach. Jeżeli po wprowadzeniu poprawki znajdą się krzywe, które nie różnią się statystycznie istotnie, to należy dokonać następnych połączeń.

Ponieważ kolejność łączenia krzywych w pary wpływa na ostateczny efekt grupowania, należy przyjąć możliwie najbardziej obiektywny algorytm postępowania. W niniejszym badaniu algorytm postępowania oparto na wartości *p-value* w testach jednorodności. Na podstawie przeprowadzonych testów wyodrębniono pary rozkładów nieróżniące się statystycznie istotnie i w kolejnych krokach łączono krzywe w pary według wartości *p-value*. Należy mieć na uwadze, że nie można jednocześnie połączyć więcej niż dwóch grup, gdyż relacja podobieństwa rozkładów nie jest przechodnia, tzn. jeżeli rozkład X jest podobny do rozkładu Y oraz rozkład Y jest podobny do rozkładu Z, to rozkład X nie musi być podobny do rozkładu Z. Kolejno przeprowadzane połączenia wyszczególniono w tablicy 5. W przypadku łączenia dwóch grup powstałych w wyniku wcześniejszego połączenia, należy sprawdzić czy powstanie grupa wewnętrznie jednorodna. W tym celu wykorzystano test dla wielu krzywych Kaplana-Meiera wyodrębnionych na podstawie pierwotnej zmiennej o $k = 14$ wariantach. Jeżeli wszystkie pierwotne krzywe wchodzące w skład dwóch wybranych grup nie różniły się statystycznie istotnie, to łączono te grupy. Na etapie łączenia grup powstałych z wcześniejszego połączenia jako kryterium rozstrzygające o tym, czy dokonać połączenia i w jakiej kolejności, przyjęto wyniki testu dla wielu pierwotnych grup. W ten sposób algorytm ten przedkłada wewnętrzną jednorodność grup nad minimalizację liczby grup. Przedstawiony powyżej algorytm wykonujemy, dopóki nie otrzymamy możliwie najmniejszej liczby grup wewnętrznie jednorodnych.

Tablica 4. Krytyczny poziom istotności (*p-value*) w teście Wilcoxona według Gehana dla par krzywych Kaplana-Meiera dla przedsiębiorstw pogrupowanych według 14 rodzajów działalności

Sekcja PKD	C	F	G	H	I	J	K	L	M	N	P	Q	R	S
C	1,0000													
F	0,4864	1,0000												
G	0,0223	0,0024	1,0000											
H	0,0000	0,0004	0,0000	1,0000										
I	0,0744	0,0261	0,6571	0,0000	1,0000									
J	0,0001	0,0005	0,0000	0,7082	0,0000	1,0000								
K	0,0000	0,0000	0,0004	0,0000	0,0402	0,0000	1,0000							
L	0,0011	0,0028	0,0001	0,2128	0,0002	0,3681	0,0000	1,0000						
M	0,0000	0,0000	0,0000	0,0590	0,0000	0,2362	0,0000	0,7867	1,0000					
N	0,6965	0,3848	0,3134	0,0008	0,2745	0,0005	0,0008	0,0018	0,0000	1,0000				
P	0,0513	0,1640	0,0009	0,1118	0,0048	0,0843	0,0000	0,0289	0,0008	0,0730	1,0000			
Q	0,0000	0,0000	0,0000	0,0006	0,0000	0,0070	0,0000	0,4676	0,0384	0,0000	0,0000	1,0000		
R	0,2089	0,3767	0,0224	0,2492	0,0264	0,1403	0,0001	0,0704	0,0173	0,1787	0,9271	0,0005	1,0000	
S	0,4609	0,2081	0,2636	0,0001	0,2769	0,0000	0,0001	0,0009	0,0000	0,9088	0,0314	0,0000	0,1153	1,0000

Źródło: Badania własne na podstawie bazy REGON (obliczenia za pomocą pakietu STATISTICA 12).

Tablica 5. Kolejne kroki w grupowaniu przedsiębiorstw według sekcji PKD

Lp.	Połączenie	*p**	Nazwy łączonych sekcji PKD	
1	P + R	0,9271	P: Edukacja	R: Działalność związana z kulturą, rozrywką i rekreacją
2	N + S	0,9088	N: Działalność w zakresie usług administrowania i działalność wspierająca	S: Pozostała działalność usługowa
3	L + M	0,7867	L: Działalność związana z obsługą rynku nieruchomości	M: Działalność profesjonalna, naukowa i techniczna
4	H + J	0,7082	H: Transport i gospodarka magazynowa	J: Informacja i komunikacja
5	G + I	0,6571	G: Handel hurtowy i detaliczny; naprawa pojazdów samochodowych, włączając motocykle	I: Działalność związana z zakwaterowaniem i usługami gastronomicznymi
6	C + F	0,4864	C: Przetwórstwo przemysłowe	F: Budownictwo
7	[C+F]+ [N+S]	0,2311	C: Przetwórstwo przemysłowe F: Budownictwo	N: Działalność w zakresie usług administrowania i działalność wspierająca S: Pozostała działalność usługowa
8	[H+J]+ [L+M]	0,0376	H: Transport i gospodarka magazynowa J: Informacja i komunikacja	L: Działalność związana z obsługą rynku nieruchomości M: Działalność profesjonalna, naukowa i techniczna

* *p-value* w teście jednorodności Wilcoxona według Gehana dla dwóch grup
Źródło: Badania własne na podstawie bazy REGON (obliczenia za pomocą pakietu STATISTICA 12).

W siódmym (przedostatnim) kroku procedury po raz pierwszy pod uwagę były brane dwie grupy powstałe w wyniku wcześniejszego połączenia: testy parami wskazały na podobieństwa sekcji [N+S] do sekcji [C+F] oraz do sekcji [G+I]. Warunek wewnętrznej jednorodności był spełniony dla dwóch potencjalnych połączeń. Zgodnie z przyjętym powyżej kryterium wybrano połączenie grupy [N+S] z grupą [C+F]. Nie można połączyć jednocześnie rozważanych 3 grup, gdyż powstałaby grupa wewnętrznie niejednorodna. Test dla 6 pierwotnych krzywych, które weszłyby w skład tej grupy, wykazał istotne różnice w ich przebiegu.

Po siódmym połączeniu wartości *p-value* we wszystkich testach dla par krzywych przetrwania były poniżej poziomu istotności $\alpha = 0{,}05$. Ze względu na wielokrotne testowanie i wzrost prawdopodobieństwa błędnego odrzucenia testowanej hipotezy zastosowano korektę wartości *p-value* według procedury Holma [Holm, 1979]. W efekcie okazało się, że w czterech przypadkach nie ma podstaw do odrzucenia hipotezy mówiącej o jednakowych funkcjach przetrwania w dwóch grupach. Jedno z czterech potencjalnych połączeń nie spełniało warunku wewnętrznej jednorodności, a spośród pozostałych trzech wybrano połączenie [H+J]+[L+M], które uznano za najbardziej jednorodne (tab. 6).

Tablica 6. Krytyczny poziom istotności (p-value) w teście dla krzywych Kaplana-Meiera przy wyborze połączenia w ostatnim kroku procedury

Połączenie	*p-value* w teście jednorodności	
	dla 2 grup	dla *k* podgrup
[L+M] + Q	$p = 0{,}0422$ (0,1023)*	$p = 0{,}1206$ ($k = 3$)
[H+J] + [L+M]	$p = 0{,}0376$ (0,1023)*	$p = 0{,}2102$ ($k = 4$)
[H+J] + [P+R]	$p = 0{,}0341$ (0,1023)*	$p = 0{,}1993$ ($k = 4$)
{[C+F]+[N+S]} + [P+R]	$p = 0{,}0191$ (0,0764)*	$p < 0{,}0001$ ($k = 6$)

* W nawiasie podano wartość *p-value* skorygowaną z powodu testowania wielokrotnego za pomocą procedury Holma

Źródło: Badania własne na podstawie bazy REGON (obliczenia za pomocą pakietu STATISTICA 12).

Po ósmym połączeniu testy parami pokazały, że krzywe Kaplana-Meira różnią się parami statystycznie istotnie (tab. 7), także po korekcie Holma uwzględniającej liczbę przeprowadzonych testów. Ostatecznie powstało sześć grup przedsiębiorstw, które są wewnętrznie jednorodne i jednocześnie różnią się między sobą pod względem rozkładu czasu trwania. Wśród tych grup dwie pozostały niepołączone – są to zauważone we wstępnej analizie sekcja Q „Opieka zdrowotna i pomoc społeczna" o najniższym ryzyku zakończenia działalności oraz sekcja K „Działalność finansowa i ubezpieczeniowa" o najwyższym ryzyku. Dwie grupy składały się z dwóch sekcji, a pozostałe dwie z czterech sekcji.

Tablica 7. Krytyczny poziom istotności (*p-value*) w teście dla par krzywych Kaplana-Meiera dla przedsiębiorstw pogrupowanych według 6 grup rodzajów działalności

Sekcja PKD	C+F+ N+S	G+I	H+J+ L+M	K	Q	P+R
C+F+N+S	1,0000					
G+I	0,0009	1,0000				
H+J+ L+M	0,0000	0,0000	1,0000			
K	0,0000	0,0004	0,0000	1,0000		
Q	0,0000	0,0000	0,0028	0,0000	1,0000	
P+R	0,0191	0,0001	0,0006	0,0000	0,0000	1,0000

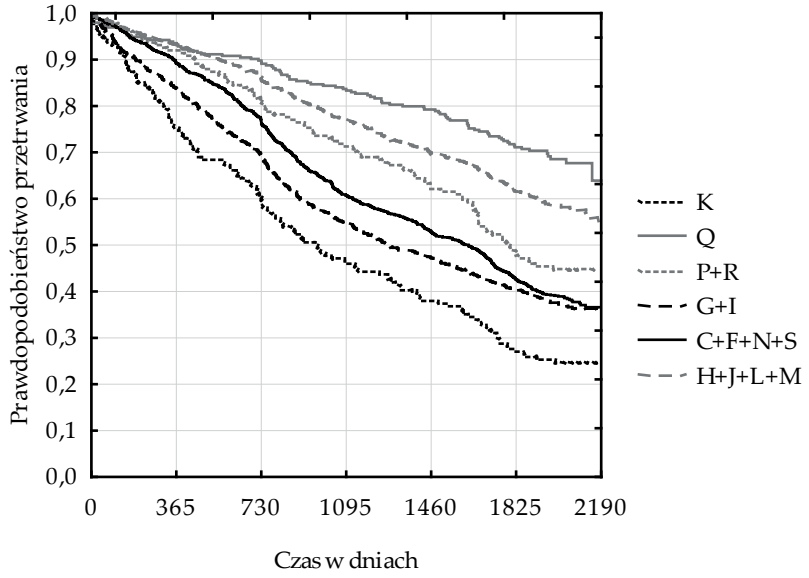
Źródło: Badania własne na podstawie bazy REGON (obliczenia za pomocą pakietu STATISTICA 12).

Analizując tempo spadku krzywych Kaplana-Meiera, można uporządkować grupy według ryzyka likwidacji (rys. 2). Kolejność grup według rosnącego ryzyka była następująca: Q, H+J+L+M, P+R, C+F+N+S, I+G, K (nazwy pogrupowanych sekcji – patrz tab. 5). Pomocniczo stosowana w tym badaniu miara jakości kategoryzacji $IV = 0{,}236$ dla pierwotnych $k = 14$ wariantów cechy, po redukcji do $k = 6$ wariantów, zmalała niewiele[9] do wartości $IV = 0{,}227$. Zmienna „sekcja PKD" po rekategoryzacji pozostała nadal dosyć silnym predyktorem ryzyka zaprzestania działalności.

Wyniki analogicznych badań, w których porównywano krzywe przetrwania dla różnych rodzajów działalności, znaleźć można w publikacjach: [Markowicz, 2012; Ptak-Chmielewska, 2016; Szymański, 2011]. Badania te różniły się nie tylko bazami danych, lecz także sposobem postępowania w łączeniu sekcji PKD w grupy. I. Markowicz oparła swoje badania na bazie danych z REGON, otrzymując dla przedsiębiorstw powstałych w Szczecinie w latach 2001–2006 trzy grupy sekcji PKD: usługowe (5 sekcji), przemysłowo-budowlane (2 sekcje) i handlowo-finansowe (2 sekcje) oraz „niestowarzyszoną" sekcję „ochrona zdrowia i opieka społeczna". Autorka wykazała, że w ramach grup sekcje parami nie różniły się statystycznie istotnie [Markowicz, 2012, s. 129–130].

[9] Przy łączeniu wariantów cechy wartość miary *IV* maleje tym bardziej, im bardziej różni się moc predykcyjna łączonych wariantów cechy (przy łączeniu atrybutów o identycznej sile predykcyjnej, wartość miary *IV* nie zmienia się).

Rysunek 2. Funkcje przetrwania przedsiębiorstw pogrupowanych według 6 grup rodzajów działalności (estymator Kaplana-Meiera)



Źródło: Badania własne na podstawie bazy REGON (obliczenia za pomocą pakietu STATISTICA 12).

Badania zawarte w [Ptak-Chmielewska, 2016] oraz [Szymański, 2011] zostały oparte na danych indywidualnych zgromadzonych przez GUS w reprezentacyjnych badaniach panelowych przeprowadzanych w 5-letnich cyklach dla mikro, a od 2005 r. także małych przedsiębiorstw w Polsce [*Warunki…*, 2015]. W obu pracach wybór sposobu grupowania według rodzaju działalności wyjaśniano za pomocą teorii ekonomicznych[10], a dopiero po wyodrębnieniu grup dokonywano weryfikacji statystycznej[11].

Natomiast A. Ptak-Chmielewska mikro i małe przedsiębiorstwa zarejestrowane w Polsce w 2004 r. podzieliła na pięć grup: budownictwo, handel, produkcję, usługi, pozostałą działalność. Testy parami dla funkcji

[10] Postawiono hipotezy, że przedsiębiorstwa z branży przemysłowej (produkcyjnej), ze względu na większe nakłady początkowe, mają większe bariery wejścia niż przedsiębiorstwa z branży usługowej i handlowej. Przedsiębiorstwa poddane większej selekcji na wejściu mają większe szanse przetrwania na rynku.

[11] W badaniach panelowych GUS wylosowana próba przedsiębiorstw zarejestrowanych w danym roku była poddawana każdej jesieni w kolejnych latach badaniom ankietowym. Informacja o rodzaju działalności pochodziła z badania ankietowego. Dla przedsiębiorstw, które zlikwidowano przed pierwszym badaniem ankietowym, nie posiadano informacji o rodzaju działalności. W celu oszacowania krzywych Kaplana-Meira według rodzaju działalności w pracach [Ptak-Chmielewska, 2016] oraz [Szymański, 2011] usunięto obserwacje z brakującymi informacjami o rodzaju działalności, otrzymując w ten sposób rozkłady warunkowe (pod warunkiem dotrwania do pierwszej ankiety).

przetrwania pokazały tylko różnice pomiędzy grupą „pozostała działalność" a pozostałymi czterema grupami [Ptak-Chmielewska, 2016, s. 126–127, 135–136]. Natomiast po wyodrębnieniu przez Autorkę dwóch grup: produkcyjnej i nieprodukcyjnej w celu konstrukcji modelu hazardu Coxa, okazało się, że przedsiębiorstwa produkcyjne mają statystycznie istotnie niższe ryzyko likwidacji niż pozostałe przedsiębiorstwa [tamże, s. 172–177].

W swojej pracy D. Szymański zaproponował podział mikroprzedsiębiorstw powstałych w Polsce w latach 2002–2003 na trzy grupy: przemysł, budownictwo oraz handel i usługi. Zastosowane różne testy statystyczne wykazały, że rozkłady czasu trwania różnią się statystycznie istotnie w tych trzech grupach, lecz nie podano wyników testów parami [Szymański, 2011, s. 116–118]. Tak skategoryzowana zmienna okazała się nieistotna w modelu hazardu Coxa, co Autor tłumaczył związkiem tej zmiennej z wielkością przedsiębiorstwa (mierzoną liczbą pracowników), będącą także zmienną objaśniającą w tym modelu [tamże, s. 134–136].

Zakończenie

W przedstawionym badaniu cecha jakościowa o wielu wariantach została wykorzystana do grupowania jednostek badania. Dokonano rekategoryzacji (redukcji liczby wariantów) rozważanej cechy w taki sposób, aby połączyć w grupy jednostki podobne pod względem rozkładu czasu trwania. Zagadnienie to nie ma jednoznacznego rozwiązania, a w literaturze można znaleźć różne sposoby łączenia wariantów rozważanej cechy, prowadzące do niejednoznacznych wyników. W artykule przedstawiono algorytm postępowania w celu wyodrębnienia grup jednostek, które są zarazem wewnętrznie jednorodne i jednocześnie różnią się między sobą statystycznie istotnie ze względu na rozkład czasu trwania. Wykorzystano w tym celu testy jednorodności dla krzywych przetrwania dla dwóch oraz większej liczby grup, przy czym kryterium wewnętrznej jednorodności było nadrzędne nad minimalizacją liczby grup. Ponieważ testy te wymagają odpowiednio dużej liczebności grup, zaproponowano wykorzystanie miary *information value* (*IV*) na wstępnym etapie rekategoryzacji cechy jakościowej. Miarę tę wykorzystano także pomocniczo do śledzenia zmian siły predykcyjnej zmiennej w kolejnych etapach rekategoryzacji.

W badaniach tego typu należy mieć na uwadze, że sposób łączenia grup może wpłynąć na wyniki badań. Tworząc niejednorodne grupy, osłabia się moc predykcyjną zmiennej grupującej, a nawet można doprowadzić do braku istotnego wpływu źle skategoryzowanej zmiennej na ryzyko wystąpienia badanego zdarzenia. W niniejszym badaniu zmienna „sekcja PKD" okazała się dosyć silnym predyktorem (według miary *IV*) ryzyka

zaprzestania działalności, nawet po rekategoryzacji według przedstawionego algorytmu i redukcji liczby wariantów z 14 do 6.

Kryterium ustalania kolejności łączenia grup powinno być możliwie najbardziej obiektywne, lecz w pewnych sytuacjach dobrym rozwiązaniem może wydawać się zastosowanie wiedzy eksperckiej, tzn. niepozbawionej subiektywizmu merytorycznej oceny zjawiska, jednak pod warunkiem, że w teście jednorodności dla łączonych rozkładów wartość *p-value* jest większa od przyjętego poziomu istotności ($p > \alpha$).

Literatura

Gurgul H., Zając P. (2016), *Modelowanie nieparametryczne czasu przeżycia przedsiębiorstw Małopolsce*, „Wiadomości Statystyczne", nr 12 (667).

Holm S. (1979), *A Simple Sequentially Rejective Multiple Test Procedure*, „Scandinavian Journal of Statistics", Vol. 6, No. 2.

Jackowska B. (2015), *Analiza kohortowa czasu istnienia mikroprzedsiębiorstw w Gdańsku*, „Zarządzanie i Finanse", nr 4, cz. 2.

Jurkiewicz T., Wycinka E. (2011), *Significance tests of differences between two crossing survival curves for small samples*, „Acta Universitatis Lodziensis. Folia Oeconomica", Vol. 255.

Kaplan E. L., Meier P. (1958), *Nonparametric Estimation from Incomplete Observations*, „Journal of the American Statistical Association", Vol. 53, No. 282.

Karadeniz P.G., Ercan I. (2017), *Examining Tests for Comparing Survival Curves with Right Censored Data*, „Statistics in Transition", Vol. 18, No. 2.

Klein J., Moeschberger M. (1997), *Survival Analysis. Techniques for Censored and Truncated Data*, Springer Science + Business Media, New York.

Manjon-Antolin M.C., Arauso-Carod J.M. (2008), *Firm Survival: Methods and Evidence*, „Empirica", Vol. 35, No. 1.

Markowicz I. (2012), *Statystyczna analiza żywotności firm*, „Rozprawy i Studia", nr 835, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Ptak-Chmielewska A. (2016), *Determinanty przeżywalności mikro i małych przedsiębiorstw w Polsce*, Oficyna Wydawnicza SGH, Warszawa.

Rossa A. (2005), *Metody estymacji rozkładu czasu trwania zjawisk dla danych cenzurowanych oraz ich zastosowania*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Siddiqi N. (2006), *Credit Risk Scorecards. Developing and Implementing Intelligent Credit Scoring*, John Wiley & Sons, New Jersey.

Suciu G.P., Lemeshow S., Moeschberger M. (2004), *Statistical Tests of the Equality of Survival Curves: Reconsidering the Option*, w: N. Balakrishnan, C.R. Rao (eds.), *Advances in Survival Analysis*, Elsevier, Amsterdam.

Szymański D. (2011), *Badanie żywotności nowo powstałych mikroprzedsiębiorstw w Polsce w latach 2002–2007*, rozprawa doktorska, https://depotuw.ceon.pl/bitstream/handle/item/177/Dariusz%20Szyma%c5%84ski%20doktorat.pdf?sequence=1, dostęp: 6.04.2017.

Warunki powstania i działania oraz perspektywy rozwojowe polskich przedsiębiorstw powstałych w latach 2009–2013 (2015), „Informacje i Opracowania Statystyczne", GUS, Warszawa.

Wycinka E. (2015), *Time to default analysis in personal credit scoring*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", nr 381.

Streszczenie

W literaturze można znaleźć wyniki badań wskazujące na to, że rodzaj działalności różnicuje rozkład czasu istnienia przedsiębiorstw. Statystycznej weryfikacji tej hipotezy można dokonać za pomocą testów jednorodności dla krzywych przetrwania. Testy te należą do grupy metod analizy przeżycia. Głównym celem badania było wyłonienie grup przedsiębiorstw, które są zarazem wewnętrznie jednorodne i jednocześnie różnią się między sobą statystycznie istotnie ze względu na rozkład czasu trwania. Problem badawczy sprowadza się do znalezienia takiego sposobu redukcji liczby wariantów cechy jakościowej, który pozwoli na odpowiednie pogrupowanie jednostek. Problem ten nie ma jednoznacznego rozwiązania: różne zbiory danych, jak też różne algorytmy postępowania mogą prowadzić do różniących się wyników. W artykule, na podstawie danych empirycznych dla przedsiębiorstw z Gdańska, przedstawiono sposób postępowania w celu uniknięcia przypadkowych decyzji przy wyodrębnianiu grup przedsiębiorstw o zbliżonym ryzyku zaprzestania działalności.

Słowa kluczowe

analiza przeżycia, krzywa przetrwania Kaplana-Meiera, mikro i małe przedsiębiorstwa, rodzaj działalności gospodarczej

Survival analysis methods in the study of the differentiation of distributions of business duration times by type of activity (Summary)

In the literature, research results can be found to indicate that the type of business activity differentiates the time distribution of business duration. A statistical verification of this hypothesis can be made by tests of equality of survival curves. These tests belong to the group of survival analysis methods. The main objective of the study was to identify groups of enterprises that are both internally homogeneous and significantly differ from each other with respect to distribution of survival time. The research problem is to find a way to reduce the number of variants of the qualitative variable which will allow for proper grouping of units. This problem does not have a unequivocal solution: different data sets and different algorithms may lead to different results. In the article, based on empirical data for enterprises from Gdańsk, a procedure was introduced to avoid accidental decisions in the grouping of enterprises with similar risk of terminating business activity.

Keywords

survival analysis, Kaplan-Meier survival curve, micro and small enterprises, type of business activity

Iwona Markowicz*

Modele trwania firm w województwie zachodniopomorskim według grup sekcji PKD

Wstęp

Prezentowany artykuł wpisuje się w nurt badań żywotności firm i dotyczy analizy czasu trwania podmiotów gospodarczych z wykorzystaniem metod wywodzących się z demografii.

Analiza przeżycia jest coraz częściej stosowana w demografii przedsiębiorstw. Przyjęta metodyka jest definiowana przez takie dyscypliny naukowe, jak statystyka, ekonometria, demografia lub analiza przetrwania. W literaturze przedmiotu można znaleźć prace prezentujące wyniki analizy przedsiębiorstw oparte na metodach modelowania zdarzeń. Przykładem może być publikacja zatytułowana *Business Demography in Spain: Determinants of Survival*. Jego autorzy przyjmują procedury analizy zdarzeń w celu zbadania firm działających w Hiszpanii [López-Garcia, Puente, 2006, s. 1–44]. Wyniki wielu badań [Hannan, Freeman, 1989, s. 3–27; Carroll, Hannan, 2000; Caves, 1998, s. 1947–1982] sugerują, że wiek, wielkość, powstawanie i likwidacja przedsiębiorstw są ze sobą powiązane. Santarelli [2000, s. 315–325] wykorzystał model regresji Coxa do analizy czasu trwania nowych firm. Analiza miała na celu określenie związku między czasem trwania i wielkością nowo utworzonych przedsiębiorstw zajmujących się pośrednictwem finansowym we Włoszech. W swoich badaniach nad bankructwem spółek notowanych w Wielkiej Brytanii w latach 1965–2002 Bhattacharjee [2005, s. 1–29] sugeruje uwzględnienie początkowej i aktualnej wielkości firm w modelu ryzyka. Udowodnił, że wpływ początkowej wielkości danej firmy na jej żywotność ulega zmianie w miarę, jak przedsiębiorstwo „dorasta". Geroski, Mata i Portugal [2007, s. 1–38] przeprowadzili analizę przeżycia w odniesieniu do nowych firm założonych w Portugalii. Oparto ją na danych zebranych przez portugalskie Ministerstwo Pracy w latach 1982–1995. Obejmowała ona obserwacje panelowe oparte na rocznym cyklu (trwanie firm było analizowane w ciągu jednego roku). Na podstawie szacunkowych modeli semiparametrycznych stwierdzono, że duże przedsiębiorstwa działały dłużej. Kaniovski i Peneder [2008, s. 41–58] wykorzystali analizę parametryczną do określenia czasu trwania firm

* Prof. US dr hab., Instytut Ekonometrii i Statystyki, Wydział Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński, ul. Mickiewicza 64, 71-101 Szczecin, iwona.markowicz@usz.edu.pl

austriackich w latach 1975–2004. Badacze udowodnili, że ryzyko likwidacji było większe w przypadku przedsiębiorstw działających w sektorze usług (w porównaniu z sektorem produkcyjnym) oraz że większe przedsiębiorstwa miały większe szanse na przetrwanie. Nunes i de Morais Sarmento [2012, s. 260–272] określili funkcję przetrwania dla firm założonych w Portugalii w latach 1987–2005; 86% przedsiębiorstw przetrwało pierwszy rok, a tylko 22% działało przez 18 lat.

W artykule zaprezentowane zostały wyniki analizy kohortowej. Kohorty stanowią firmy danej sekcji PKD powstałe w województwie zachodniopomorskim w latach 2009–2011 (rejestr REGON). Obserwacja trwała do końca 2013 r. (59 587 firm). Analizie poddano 13 sekcji spośród 21. Pominięto sekcje o stosunkowo małej liczbie podmiotów powstających i likwidowanych.

Celem artykułu jest budowa modeli trwania firm dla grup sekcji PKD w województwie zachodniopomorskim i ich porównanie. Można postawić hipotezę, że model trwania firm jest uzależniony od rodzaju jej działalności. Zatem podobieństwo procesu trwania firm z określonych sekcji PKD daje możliwość połączenia sekcji w grupy. Badania przeprowadzono etapami. Pierwszy z nich to wyodrębnienie grup sekcji o podobnych modelach czasu trwania firm (estymator Kaplana-Meiera i test Gehana). Drugi etap to analiza przebiegu funkcji przetrwania dla tych grup. Kolejny etap obejmuje analizę przebiegu funkcji intensywności likwidacji firm w grupach (skonstruowanie tablic trwania firm). Następnie porównano odsetek firm zlikwidowanych w pierwszym i drugim roku działalności w poszczególnych grupach. Wyniki badań wskazujące na rodzaje działalności wpływające na zwiększone ryzyko likwidacji mogą stanowić wkład w literaturę przedmiotu.

Do wyznaczenia modeli trwania firm według sekcji wykorzystano estymator Kaplana-Meiera (model nieparametryczny). Metoda ta jest stosowana przy założeniu występowania obserwacji cenzurowanych. Firmy zlikwidowane do końca 2013 r. przyjęto jako obserwacje pełne, a pozostałe jako cenzurowane (firmy, które przetrwały poza okres obserwacji). Za pomocą testu Gehana zweryfikowano hipotezę o podobieństwie tych modeli. Na tej podstawie wyznaczono grupy sekcji o podobnych modelach trwania.

Dla uzyskanych grup wyznaczono funkcje intensywności likwidowania firm (funkcje hazardu). Wiele badań potwierdza, że funkcja ta ma kształt odwróconej litery *U* z określonym maksimum, co jest zgodne z teoretycznym modelem uczenia się [Wagner, 1994, s. 141–154; Audretsch i inni, 1999, s. 965–983; Bartelsmann i inni, 2005, s. 365–391; Markowicz, 2012; 2016].

Należy podkreślić, że zastosowane w badaniu metody analizy trwania (przeżycia) wywodzą się z demografii [Frątczak i inni, 2005], ale coraz

częściej są stosowane w badaniach trwania firm [Nehrebecka, Dzik, 2013; Jackowska, 2015; Mikulec, 2017], a także w badaniach z innych dziedzin nauki [Bieszk-Stolorz, 2013; Landmesser, 2013; Sączewska-Piotrowska, 2016]. O stosowaniu innych metod w analizie trwania firm przeczytać można np. w pracach [Dehnel, 2010; Pociecha, Pawełek, 2011].

1. Dane statystyczne

W artykule zaprezentowane zostały wyniki analizy kohortowej. Kohorty stanowią firmy danej sekcji PKD2007 powstałe w województwie zachodniopomorskim w latach 2009–2011 i obserwowane do końca 2013 r. Informacje o liczbie powstałych i zlikwidowanych firm według sekcji zawarto w tablicy 1. Natomiast na rysunku 1 przedstawiono liczbę firm zlikwidowanych i cenzurowanych (niezlikwidowanych do końca okresu obserwacji) według sekcji. W latach 2009–2011 w województwie zachodniopomorskim powstało 59 587 firm. Spośród 21 sekcji analizie poddano 13. Pominięto sekcje o stosunkowo małej liczbie podmiotów[1]. Zarówno wśród firm powstających, jak i likwidowanych przeważały firmy handlowe i budowlane.

Tablica 1. Liczba firm powstałych w latach 2009–2011 i zlikwidowanych do końca 2013 r. w województwie zachodniopomorskim według sekcji PKD2007

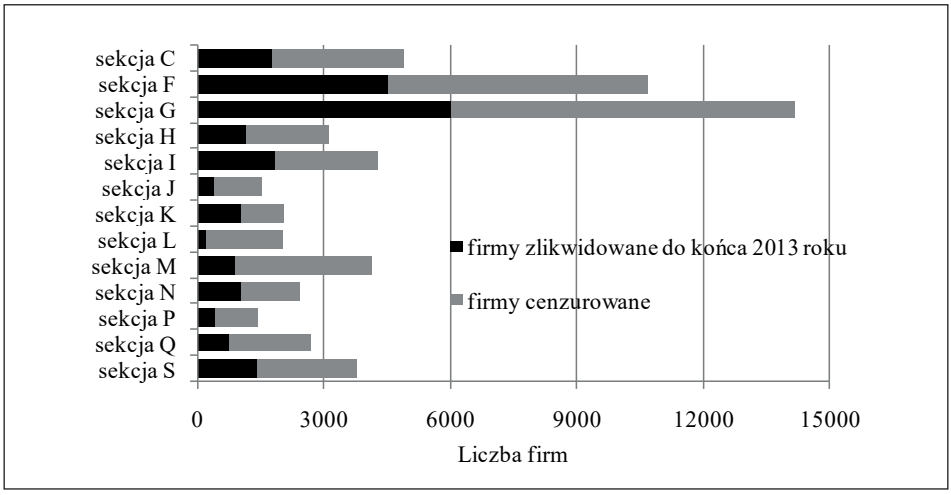
Sekcje PKD2007	Liczba firm	
	powstałych w latach 2009–2011	zlikwidowanych do końca 2013 r. (%)
Ogółem 21 sekcji	59 587	22 234 (37,3)
Razem 13 sekcji badanych	57 209	21 531 (37,6)
C – przetwórstwo przemysłowe	4 885	1 767 (36,2)
F – budownictwo	10 696	4 557 (42,6)
G – handel hurtowy i detaliczny; naprawa pojazdów samochodowych, włączając motocykle	14 182	6 018 (42,4)
H – transport i gospodarka magazynowa	3 129	1 155 (36,9)
I – działalność związana z zakwaterowaniem i usługami gastronomicznymi	4 292	1 848 (43,1)
J – informacja i komunikacja	1 514	391 (25,8)
K – działalność finansowa i ubezpieczeniowa	2 056	1 044 (50,8)
L – działalność związana z obsługą rynku nieruchomości	2 020	198 (9,8)

[1] Sekcje pominięte: A – rolnictwo, leśnictwo, łowiectwo i rybactwo, B – górnictwo i wydobywanie, D – wytwarzanie i zaopatrywanie w energię elektryczną, gaz…, E – dostawa wody…, O – administracja publiczna i obrona narodowa…, R – działalność związana z kulturą, rozrywką i rekreacją, T – gospodarstwa domowe zatrudniające pracowników…, U – organizacje i zespoły eksterytorialne.

Sekcje PKD2007	Liczba firm	
	powstałych w latach 2009–2011	zlikwidowanych do końca 2013 r. (%)
M – działalność profesjonalna, naukowa i techniczna	4 124	915 (22,2)
N – działalność w zakresie usług administrowania i działalność wspierająca	2 412	1 028 (42,6)
P – edukacja	1 452	442 (30,4)
Q – opieka zdrowotna i pomoc społeczna	2 687	768 (28,6)
S – pozostała działalność usługowa	3 760	1 400 (37,2)

Źródło: Opracowanie własne na podstawie danych z rejestru REGON.

Rysunek 1. Liczba firm powstałych w latach 2009–2011 z podziałem na zlikwidowane i cenzurowane do końca 2013 r. w województwie zachodniopomorskim według sekcji PKD



Źródło: Opracowanie własne.

2. Model czasu trwania firm – wyodrębnienie grup sekcji

Do budowy nieparametrycznego modelu czasu trwania firm według rodzaju działalności można wykorzystać estymator Kaplana-Meiera (funkcja przetrwania), przy założeniu występowania obserwacji cenzurowanych. Metoda ta, w odróżnieniu od tablic trwania, nie wymaga arbitralnego grupowania czasu obserwacji w przedziały klasowe. Czas trwania firmy jest realizacją zmiennej losowej (T; δ), przy czym:

$$T = \begin{cases} T_z & \text{dla } \delta = 1 \\ T_c & \text{dla } \delta = 0 \end{cases} \tag{1}$$

gdzie: T_z – czas trwania firmy zlikwidowanej, T_c – czas trwania firmy cenzurowanej, δ – zmienna losowa przyjmująca wartość 1 dla obserwacji pełnej i wartość 0 dla obserwacji cenzurowanej.

Estymator Kaplana-Meiera można zapisać w postaci [Kaplan, Meier, 1958; Markowicz, 2012; Bieszk-Stolorz, Markowicz, 2015]:

$$\hat{S}(t_i) = \prod_{t \le t_i} \left(1 - \frac{z_i}{n_i}\right) \qquad \text{dla } i = 1, ..., k \tag{2}$$

gdzie: t_i – punkt czasu, w którym wystąpiło co najmniej jedno zdarzenie (została zlikwidowana firma), z_i – liczba zdarzeń w czasie t_i (obserwacji pełnych), n_i – liczba jednostek objętych obserwacją w czasie t_i.

Jest to funkcja nierosnąca, przedziałami stała, o skokach w losowych punktach czasu, określonych przez obserwacje pełne (likwidacja co najmniej jednej firmy). Estymator przyjmuje wartości:

$$\hat{S}(t_i) = \begin{cases} 1 & \text{dla} \quad t_0 \\ \prod_{t \le t_i} \left(1 - \frac{z_i}{n_i}\right) & \text{dla} \quad t_1 \le t \le t_k \\ 0 & \text{dla} \quad t > t_k \quad \text{gdy} \quad \delta_n = 1 \\ \text{niezdefiniowany} & \text{dla} \quad t > t_k \quad \text{gdy} \quad \delta_n = 0 \end{cases} \tag{3}$$

Początkowa wartość funkcji czasu trwania wynosi 1 i maleje w kolejnych punktach czasu t_i (t_1, t_2, …, t_k), w których zaszło przynajmniej jedno analizowane zdarzenie. Własności statystyczne estymatora Kaplana-Meiera są w literaturze uznane za dobre w przypadku dużych prób [Rossa, 2005]. Stosowanie tego estymatora, jak wspomniano, nie wymaga określenia przedziałów czasu trwania, ponieważ prawdopodobieństwo przetrwania można oszacować w dowolnym momencie. Modele przetrwania odzwierciedlające procesy trwania w wyodrębnionych grupach można porównywać i aby stwierdzić statystyczną istotność różnic, przeprowadza się odpowiedni test statystyczny. Należy zastosować test nieparametryczny (nieznany jest rozkład czasu trwania), uwzględniający istnienie danych cenzurowanych. W literaturze proponowane są różne testy, brakuje natomiast jednolitych sposobów wyboru testu w danej sytuacji [Gehan, 1965; Klainbaum, Klein, 2005]. Podkreśla się jednak, że testy te dają rzetelne wyniki dla dużych prób. Sprawdzeniu podlega hipoteza o równości funkcji przeżycia dla badanych podgrup. Do jej zweryfikowania wykorzystano test Gehana (uogólnienie testu Wilcoxona), którego statystykę można zapisać jako [Namboodiri, Suchindran, 1987]:

$$G = \frac{W}{\sqrt{V}} \tag{4}$$

przy czym [Domański, Pruska, 2000; Domański i inni, 2014]:

$$W = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} U_{ij} \tag{5}$$

$$U_{ij} = \begin{cases} -1 & \text{dla} \quad t_i < t_j \text{ lub } t_i \leq t_j^+ \\ 0 & \text{dla} \quad t_i = t_j \text{ lub } t_i^+ = t_j^+ \text{ lub } t_i^+ < t_j \text{ lub } t_j^+ < t_i \\ 1 & \text{dla} \quad t_i > t_j \text{ lub } t_i^+ \geq t_j \end{cases} \tag{6}$$

gdzie: t_i – obserwacje pełne pierwszej grupy, t_j – obserwacje pełne drugiej grupy, t_i^+ – obserwacje cenzurowane pierwszej grupy, t_j^+ – obserwacje cenzurowane drugiej grupy,

$$V = \frac{n_1 n_2 \sum_{i=1}^{n_1+n_2} U_i^2}{(n_1 + n_2)(n_1 + n_2 - 1)} \tag{7}$$

W celu wyodrębnienia grup sekcji o podobnych modelach czasu trwania firm zbadano istotność różnic funkcji przeżycia firm powstałych w latach 2009–2011. Oszacowano funkcje przetrwania (estymatory Kaplana-Meiera) dla każdej z 13 sekcji i porównano je parami. Grupy wyodrębniono w taki sposób, aby w każdej z nich znajdowały się sekcje, których modele czasu trwania nie różnią się statystycznie istotnie. Dla każdej pary sekcji sprawdzono hipotezę o postaci: $H_0 : S_1(t) = S_2(t)$, dla wszystkich t. Różnice przyjęto za istotne dla $p \leq 0{,}05$. Zatem w danej grupie znajdują się wyłącznie sekcje z podobnymi funkcjami przeżycia. Funkcje te dla sekcji z danej grupy mogą być istotnie lub nieistotnie różne od funkcji dla sekcji z innych grup. Ponadto grupy uszeregowano według zmniejszania się prawdopodobieństwa przetrwania firm wraz z czasem. Wyodrębnione grupy przedstawiono w tablicy 2. W wyniku zastosowanego schematu podziału ustalono osiem grup o podobnych modelach czasu trwania firm. Grupa pierwsza to sekcja L – działalność związana z obsługą rynku nieruchomości, w której prawdopodobieństwo przetrwania firm w kolejnych miesiącach było największe. Kolejne grupy 2 i 3 są także jednoelementowe i są to odpowiednio: sekcja M – działalność profesjonalna, naukowa i techniczna i sekcja J – informacja i komunikacja. Grupę czwartą stanowią dwie sekcje Q – opieka zdrowotna i pomoc społeczna oraz P – edukacja. Grupa piąta obejmuje przetwórstwo przemysłowe, transport i pozostałe usługi (C, H i S). W grupie 6 znalazły się dwie sekcje z największą liczbą powstałych firm: budownictwo i handel (F i G) oraz działalność w zakresie usług administrowania i działalność wspierająca (N). Jak już wspomniano, kolejne grupy charakteryzują się coraz szybciej malejącymi funkcjami trwania. Dwie ostatnie grupy (7 i 8) to sekcje: I – działalność związana z zakwaterowaniem i usługami gastronomicznymi

oraz K – działalność finansowa i ubezpieczeniowa. Charakteryzują się one najszybszym ubywaniem firm z kohorty. Należy zaznaczyć, że jedyne podobieństwo dla pary sekcji spoza jednej grupy dotyczy pary I-N, poza tym przypadkiem wszystkie grupy są rozłączne. W ostatniej kolumnie tablicy 2 podano wartości testu dla wielu prób wyznaczonego dla grup wieloelementowych. Wyniki wskazują na brak istotnych różnic w przebiegu funkcji przeżycia w obrębie wydzielonych grup.

Tablica 2. Wyodrębnione grupy sekcji PKD w województwie zachodniopomorskim o podobnych modelach trwania firm

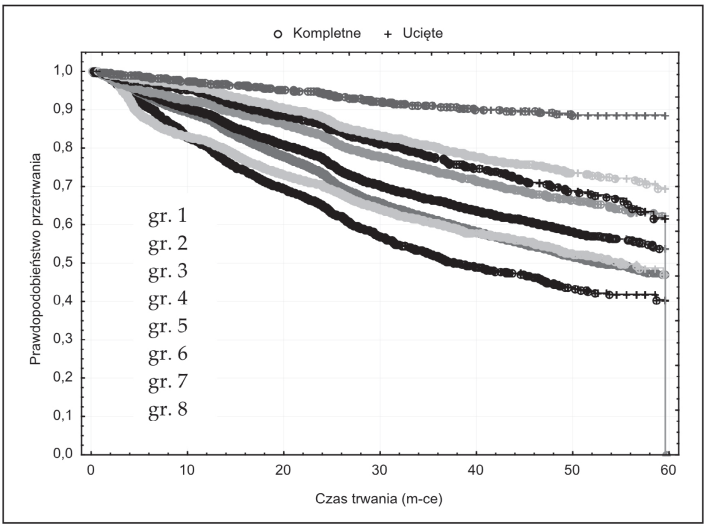
Grupa	Sekcja PKD		Wartość statystyki testu Gehana (wartość *p*)	Wartość statystyki testu dla wielu prób (wartość *p*)
1	L	działalność związana z obsługą rynku nieruchomości	0 podobieństw	–
2	M	działalność profesjonalna, naukowa i techniczna	0 podobieństw	–
3	J	informacja i komunikacja	0 podobieństw	–
4	Q	opieka zdrowotna i pomoc społeczna	Q-P: 0,0748 (0,9404)	Q-P: 0,0748 (0,9404)
	P	edukacja		
5	C	przetwórstwo przemysłowe	C-H: 1,6140 (0,1065) C-S: 1,9275 (0,0539) H-S: 0,1772 (0,8595)	C-H-S: 4,5050 (0,1052)
	H	transport i gospodarka magazynowa		
	S	pozostała działalność usługowa		
6	F	budownictwo	F-G: 0,6905 (0,4899) F-N: 1,2187 (0,2230) G-N: 0,7957 (0,4262)	F-G-N: 1,5309 (0,4651)
	G	handel hurtowy i detaliczny; naprawa pojazdów samochodowych, włączając motocykle		
	N	działalność w zakresie usług administrowania i działalność wspierająca		
7	I	działalność związana z zakwaterowaniem i usługami gastronomicznymi	1 podobieństwo I-N: –1,8247 (0,0681)	–
8	K	działalność finansowa i ubezpieczeniowa	0 podobieństw	–

Źródło: Opracowanie własne.

3. Analiza przebiegu funkcji przetrwania dla grup

W drugim etapie badań wyznaczono ponownie funkcje przetrwania w oparciu o estymator Kaplana-Meiera, ale dla wyodrębnionych grup sekcji PKD. Funkcje te przedstawiono na rysunku 2.

Rysunek 2. Modele trwania (estymator Kaplana-Meiera) dla grup sekcji PKD (grupy od 1 do 8)



Źródło: Opracowanie własne.

Grupy sekcji od 1 do 8 charakteryzują się coraz krótszym czasem do likwidacji firm (coraz szybciej malejące funkcje przetrwania). Wartość kwartyla pierwszego (likwidacja 25% firm z danej kohorty) osiągnęły grupy od 2 do 8 po czasie odpowiednio: 47,2; 39,6; 34,4; 25,3; 22,8; 17,9; i 15,1 miesiąca. Natomiast wartość mediany osiągnęły trzy grupy z najszybciej malejącymi funkcjami przetrwania (grupy 6, 7 i 8) odpowiednio po czasie: 53,2; 55,0 i 38,2 miesiąca. Zatem 50% firm powstałych w latach 2009–2011, zajmujących się działalnością finansową i ubezpieczeniową, zostało zlikwidowanych w ciągu 38,2 miesięcy od rozpoczęcia funkcjonowania.

4. Intensywność likwidacji firm w grupach

Kolejnym etapem analizy była konstrukcja kohortowych tablic trwania firm (w ujęciu aktuarialnym) dla ośmiu grup sekcji. Model tabelaryczny zbudowano dla trzymiesięcznych okresów [Markowicz, 2015]. Firmy niezlikwidowane do końca 2013 r. stanowią obserwacje cenzurowane. Spośród elementów tablicy analizie poddano funkcję intensywności likwidowania firm (funkcja hazardu). Funkcja hazardu określa ryzyko wystąpienia określonego zdarzenia w krótkim przedziale czasu (t; $t + \Delta t$), pod warunkiem że

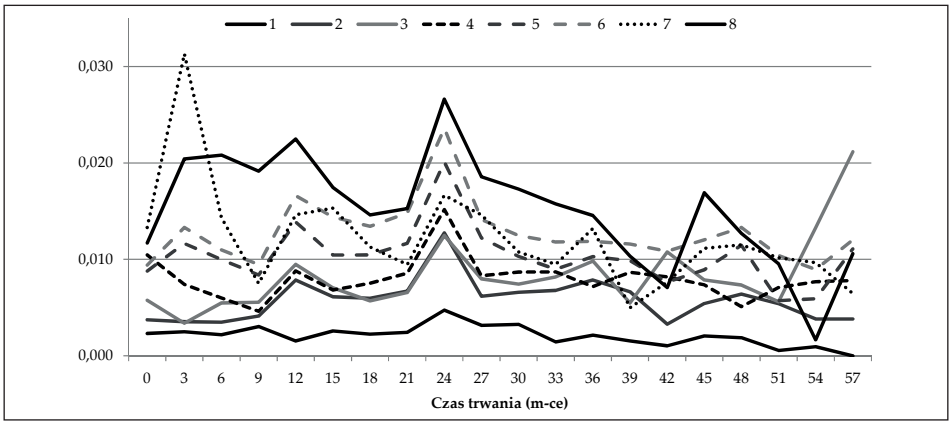
nie wystąpiło ono do czasu t [Bieszk-Stolorz, Markowicz, 2012]. Estymator $\hat{h}_t^*$ funkcji intensywności jest wyznaczany jako [Balicki, 2006; Landmesser, 2013; Markowicz, 2017]:

$$\hat{h}_t^* = \frac{z_t}{\left(n_t^* - \frac{z_t}{2}\right)a_t} \tag{8}$$

gdzie: t – początek przedziału czasu trwania $\langle t, t+1)$, n^*_t – liczba firm narażonych w przedziale, z_t – liczba firm zlikwidowanych w przedziale, a_t – długość przedziału trwania

Wiele badań potwierdza, że funkcja intensywności likwidowania przedsiębiorstw ma kształt odwróconej litery *U* z określonym maksimum, co jest zgodne z teoretycznym modelem uczenia się. Funkcje hazardu dla ośmiu grup przedstawiono na rysunku 3.

Rysunek 3. Ocena intensywności likwidacji firm w grupach sekcji PKD



Źródło: Opracowanie własne.

Funkcje poszczególnych grup mają charakterystyczny przebieg. Czas trwania analizowany jest w przedziałach (rys. 3 – początki przedziałów czasu). Analizując kształt funkcji intensywności likwidacji firm w poszczególnych grupach sekcji utworzonych według modeli trwania, można sformułować następujące spostrzeżenia:

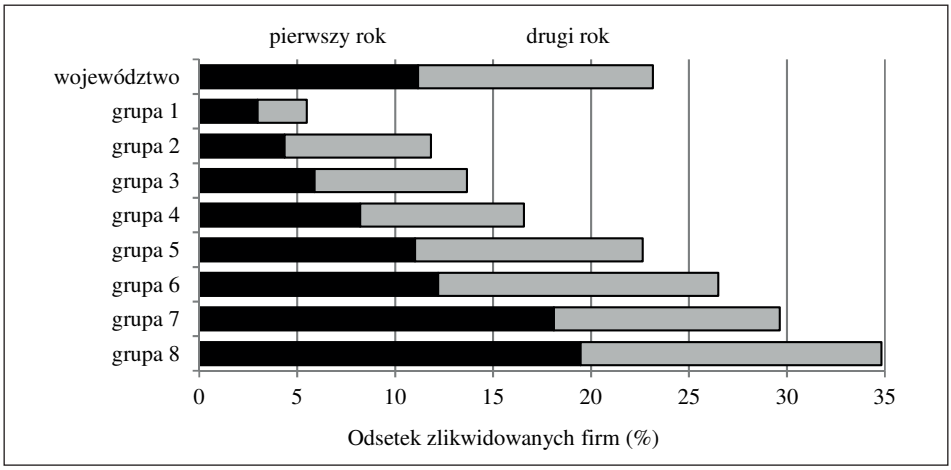
- funkcja intensywności likwidacji firm dla grup ma przeważnie przybliżony kształt odwróconej litery *U* z zaznaczonym maksimum w przedziale 24–27 miesięcy,
- przechodząc od grupy 1 do grupy 8, zauważa się coraz większe intensywności likwidacji firm oraz coraz większe jej wahania w czasie,
- funkcja dla grupy 1 (sekcja L – obsługa rynku nieruchomości) przyjmuje niskie wartości; wyrównana intensywność likwidacji,

- dla grupy 7 (sekcja I – zakwaterowanie i gastronomia) funkcja przyjmuje maksimum po 3 miesiącach działalności (przyczyną jest tu częsta działalność sezonowa)
- dla firm grupy 8 (sekcja K – działalność finansowa i ubezpieczeniowa) zauważa się wyższą intensywność likwidacji w pierwszych 2 latach funkcjonowania.

5. Likwidacja firm w pierwszych dwóch latach działalności

W województwie zachodniopomorskim w latach 2009–2011 powstało 59 587 firm. Do końca 2013 r. zostało zlikwidowanych 37,3% z nich. Odsetek ten dla poszczególnych grup sekcji wynosi: 9,8% (gr. 1, sekcja L), 22,2% (gr. 2, sekcja M), 25,8% (gr. 3, sekcja J), 29,2% (gr. 4, sekcja P i Q), 36,7% (gr. 5, sekcja C, H i S), 42,5% (gr. 6, sekcja F, G i N), 43,1% (gr. 7, sekcja I), 50,8% (gr. 8, sekcja K). Zatem przechodząc od 1 do 8 grupy, odsetek likwidowanych firm był coraz większy. Wcześniejsza analiza funkcji intensywności wykazała, że krytycznym momentem dla firm jest 24 miesiąc działalności. Po tym czasie intensywność likwidacji malała. Wyznaczony został więc wskaźnik likwidacji po 24 miesiącach funkcjonowania firm. Wyniki prezentuje rysunek 4 (podano też wskaźnik po 1 roku). W województwie po dwóch latach działalności zlikwidowano 23,2% badanych firm. Na rysunku 4 sekcje PKD przedstawiono według ustalonych grup. Odsetek likwidowanych firm jest coraz większy; w kolejnych grupach odpowiednio: 55% (gr. 1), 11,8% (gr. 2), 13,7% (gr. 3), 16,6% (gr. 4), 22,6% (gr. 5), 26,5% (gr. 6), 29,6% (gr. 7) i 34,8% (gr. 8).

Rysunek 4. Odsetek zlikwidowanych firm według grup sekcji PKD w pierwszym i drugim roku działalności



Źródło: Obliczenia własne.

Zakończenie

Wyniki badań zaprezentowane w niniejszym artykule pozwoliły wskazać na zróżnicowanie modeli trwania firm województwa zachodniopomorskiego według rodzaju działalności (PKD2007). Utworzono osiem grup. Grupy te (od 1 do 8) można scharakteryzować następująco:

1) coraz krótszy czas do likwidacji firm (coraz szybciej malejące funkcje trwania),
2) funkcja intensywności likwidacji firm dla grup ma przeważnie przybliżony kształt odwróconej litery *U* z zaznaczonym maksimum w przedziale 24–27 miesięcy,
3) przechodząc od grupy 1 do grupy 8, zauważa się coraz większe intensywności likwidacji firm oraz coraz większe jej wahania w czasie,
4) zauważono pewne wyjątki:
   - funkcja dla grupy 1 (sekcja L – obsługa rynku nieruchomości) przyjmuje niskie wartości; wyrównana intensywność likwidacji,
   - dla grupy 7 (sekcja I – zakwaterowanie i gastronomia) funkcja przyjmuje maksimum po 3 miesiącach działalności (przyczyną jest tu częsta działalność sezonowa),
   - dla firm grupy 8 (sekcja K – działalność finansowa i ubezpieczeniowa) zauważa sie wyższą intensywność likwidacji w pierwszych 2 latach funkcjonowania,
5) coraz większy odsetek firm likwidowanych po dwóch latach działalności.

Badania wskazały, że krytycznym momentem jest 24 miesiąc działalności. Jednym z powodów likwidowania firmy może być zakończenie okresu obowiązywania tzw. małego ZUS-u [ZUS, 2014]. Od 2005 r. właściciele nowych firm mogą płacić dużo niższe składki przez pierwsze 24 miesiące, co z pewnością ułatwia prowadzenie działalności. Można jednak stwierdzić, że likwidowanie firm po zakończeniu obowiązywania preferencyjnych składek sugeruje słabą pozycję tych firm.

Jak już wspomniano, wyniki badań zespołu Kaniovski i Peneder [2008], wskazują, że ryzyko likwidacji było większe w przypadku przedsiębiorstw działających w sektorze usług niż w sektorze produkcyjnym. Z badań omówionych w niniejszym artykule wynika, że dla niektórych rodzajów usług to ryzyko było większe (np. działalność finansowa i ubezpieczeniowa), a dla niektórych mniejsze (np. działalność związana z obsługą nieruchomości).

Literatura

Audretsch D.B., Santarelli E., Vivarelli M. (1999), *Start-up Size and Industrial Dynamics: Some Evidence from Italian Manufacturing*, „International Journal of Industrial Organization", Vol. 17, No. 7.

Balicki A. (2006), *Analiza przeżycia i tablice wymieralności*, PWE, Warszawa.

Bartelsmann E., Scarpetta S., Schivardi F. (2005), *Comparative Analysis of Firm Demographics and Survival: Micro-Level Evidence for the OECD Countries*, „Industrial and Corporate Change", Vol. 14, No. 3.

Bieszk-Stolorz B. (2013), *Analiza historii zdarzeń w badaniu bezrobocia*, Volumina.pl, Szczecin.

Bieszk-Stolorz B., Markowicz I. (2012), *Modele regresji Coxa w analizie bezrobocia*, CeDeWu, Warszawa.

Bieszk-Stolorz B., Markowicz I. (2015), *Influence of Unemployment Benefit on Duration of Registered Unemployment Spells*, „Equilibrium. Quarterly Journal of Economics and Economic Policy", Vol. 10, No. 3.

Bhattacharjee A. (2005), *Models of Firm Dynamics and the Hazard Rate of Exits: Reconciling Theory and Evidence using Hazard Regression Models*, CRIEFF Discussion Papers 0502, Centre for Research into Industry, Enterprise, Finance and the Firm, University of St. Andrews.

Carroll G.R., Hannan M.T. (2000), *The Demography of Corporations and Industries*, Princeton University Press.

Caves R.E. (1998), *Industrial Organization and New Findings on the Turnover and Mobility of Firms*, „Journal of Economic Literature", Vol. 36, No. 4.

Dehnel G. (2010), *Rozwój mikroprzedsiębiorczości w Polsce w świetle estymacji dla małych domen*, Wydawnictwo UE w Poznaniu, Poznań.

Domański Cz., Pruska K. (2000), *Nieklasyczne metody statystyczne*, PWE, Warszawa.

Domański Cz., Pekasiewicz D., Baszczyńska A., Witaszczyk A. (2014), *Testy statystyczne w procesie podejmowania decyzji*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Frątczak E., Gach-Ciepiela U., Babiker H. (2005), *Analiza historii zdarzeń. Elementy teorii, wybrane przykłady zastosowań*, SGH, Warszawa.

Gehan E.A. (1965), *A Generalized Two-Sample Wilcoxon Test for Double-Censored Data*, „Biometrika", Vol. 52, No. 3–4.

Geroski P.A., Mata J., Portugal P. (2010), *Founding Conditions and the Survival of New Firms*, „Strategic Management Journal", Vol. 31, No. 5.

Hannan M.T., Freeman J. (1989), *Organizational Ecology*, Harvard University Press, Cambridge.

Jackowska B. (2015), *Analiza kohortowa czasu istnienia mikroprzedsiębiorstw w Gdańsku*, „Zarządzanie i Finanse", nr 4, cz. 2.

Kaniovski S., Peneder M. (2008), *Determinants of Firm Survival: A Duration Analysis using the Generalized Gamma Distribution*, „Empirica", Vol. 35, No. 1.

Kaplan E. L., Meier P. (1958), *Nonparametric Estimation from Incomplete Observations*, „Journal of the American Statistical Association", Vol. 53, No. 282.

Kleinbaum D.G., Klein M. (2005), *Survival Analysis*, Second Edit., Springer, New York.

Landmesser J.M. (2013), *Wykorzystanie metod analizy czasu trwania do badania aktywności ekonomicznej ludności w Polsce*, Wydawnictwo SGGW, Warszawa.

López-Garcia P., Puente S. (2006), *Business Demography in Spain: Determinants of Firm Survival*, Documentos de Trabajo No 0608, Madrid, Banco de España.
Markowicz I. (2012), *Statystyczna analiza żywotności firm*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.
Markowicz I. (2015), *Duration Analysis of Firms – Cohort Tables and Hazard Function*, „International Journal of Business and Social Research", Vol. 5, No. 11.
Markowicz I. (2016), *Tablice trwania firm w województwie zachodniopomorskim według rodzaju działalności*, „Taksonomia", nr 26, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu.
Markowicz I. (2017), *Duration Model of Enterprises – Analysis of Territorial Groups*, Institute of Economic Research Working Papers, No. 71, Toruń.
Mikulec A. (2017), *Kohortowe tablice trwania Przedsiębiorstw w województwie łódzkim – ujęcie kwartalne*, „Taksonomia", nr 28, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu.
Namboodiri K., Suchindran C.M. (1987), *Life Table Techniques and Their Applications*, Academic Press Inc., New York.
Nehrebecka N., Dzik A.M. (2013), *Zdolność przetrwania przedsiębiorstw w Polsce*, „Wiadomości Statystyczne", nr 5.
Nunes A., de Morais Sarmento E. (2012), *Business Demography Dynamics in Portugal: A Non-Parametric Survival Analysis*, w: J. Bonnet, M. Dejardin, A. Madrid-Guijarro (eds.), *The Shift to the Entrepreneurial Society. A Built Economy in Education, Sustainability and Regulation*, Edward Elgar Publishing, Cheltenham.
Pociecha J., Pawełek B. (2011), *Prognozowanie bankructwa a koniunktura gospodarcza*, „Metody analizy danych", Zeszyty Naukowe UE w Krakowie, nr 873, Kraków.
Rossa A. (2005), *Metody estymacji rozkładu czasu trwania zjawisk dla danych cenzurowanych oraz ich zastosowania*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
Santarelli E. (2000), *The duration of new firms in banking: an application of Cox regression analysis*, „Empirical Economics", Vol. 25, No. 2.
Sączewska-Piotrowska A. (2016), *Badanie dynamiki ubóstwa gospodarstw domowych z wykorzystaniem wybranych modeli analizy historii zdarzeń*, „Collegium of Economic Analysis Annals", nr 41.
Wagner J. (1994), *The Post-Entry Performance of New Small Firms in German Manufacturing Industries*, „Journal of Industrial Economics", Vol. 42, No. 2.
ZUS (2014), *Ubezpieczenia społeczne i ubezpieczenie zdrowotne osób prowadzących pozarolniczą działalność i osób z nimi współpracujących*, http://www.zus.pl/pliki/poradniki/porad25.pdf.pdf.

Streszczenie

W artykule zaprezentowane zostały wyniki analizy kohortowej. Kohorty stanowią firmy danej sekcji PKD powstałe w województwie zachodniopomorskim w latach 2009-2011 (rejestr REGON). Obserwacja trwała do końca 2013 roku (59587 firm). Analizie poddano 13 sekcji spośród 21. Pominięto sekcje o stosunkowo małej liczbie podmiotów powstających i likwidowanych. Do wyznaczenia modeli trwania firm według sekcji wykorzystano estymator Kaplana-Meiera (model nieparametryczny). Metoda ta jest stosowana przy założeniu występowania obserwacji cenzurowanych. W odróżnieniu od tablic trwania, nie wymaga ona arbitralnego

grupowania czasu obserwacji w przedziały klasowe. Firmy zlikwidowane do końca 2013 roku przyjęto jako obserwacje pełne a pozostałe jako cenzurowane (firmy, które przetrwały poza okres obserwacji). Za pomocą testu Gehana zweryfikowano hipotezę o podobieństwie tych modeli. Jest to test nieparametryczny – nieznany jest rozkład czasu trwania. Na tej podstawie wyznaczono grupy sekcji o podobnych modelach trwania. W literaturze przedmiotu podkreśla się, że własności statystyczne estymatora Kaplana-Meiera są dobre a wyniki testu rzetelne tylko w przypadku dużych prób.

Efektem analizy jest podział sekcji na osiem grup. Wśród nich są trzy grupy wieloelementowe i aż pięć grup jednoelementowych. Świadczy to o indywidualnym przebiegu procesu trwania firm dla wielu sekcji PKD.

Słowa kluczowe
model trwania, kohorty firm, funkcja intensywności likwidacji firm, sekcje PKD

Duration models of firms in the Zachodniopomorskie voivodeship by NACE section groups (Summary)

The paper presents results of a cohort analysis. The cohorts under consideration consist of firms within a NACE section established in Zachodniopomorskie voivodeship between 2009-2011 (according to REGON). The entities (59.587 firms) were observed until the end of 2013. There were 13 out of 21 sections analysed. Sections that were omitted consisted of relatively small amounts of firms – both established and liquidated. To estimate firm duration models within sections Kaplan-Meier estimator was used (i.e. a non-parametric model). This method is usually applied with the assumption of censored observations in place. Unlike the survival tables, it does not depend on arbitrary division of observation times into classes. Businesses that were liquidated before the end of 2013 have been considered full observations while the rest of them being censored (i.e. all the firms that survived beyond the end of the observation period). Gehan test was employed to verify the hypothesis about similarity of duration models. In literature many authors emphasize the facts that Kaplan-Meier estimator has good statistical properties and the test is reliable only when there is a large sample involved.

The effect of the study is a partition of NACE sections into eight groups. Three of these groups contain multiple elements while five are one-element only. It might be the evidence that there exist individual survival processes in many NACE sections.

Keywords
duration models, cohorts of firms, intensity function of firms liquidation, NACE sections

Tomasz Pisula*

Zastosowanie ensemble klasyfikatorów do oceny ryzyka upadłości przedsiębiorstw na przykładzie firm sektora produkcyjnego działających na Podkarpaciu

Wstęp

Udzielanie kredytów przedsiębiorstwom, czyli klientom korporacyjnym, jest jednym z podstawowych rodzajów działalności banków i instytucji kredytowych. Działalność ta, jak każda inna forma działalności banków związana z pożyczaniem pieniędzy, obarczona jest dużym ryzykiem, wynikającym z tego, że przedsiębiorstwo w przyszłości może nie być w stanie zrealizować zaciągniętych zobowiązań kredytowych. Istniejące realne potencjalne ryzyko bankructwa przedsiębiorstwa może spowodować sytuację, gdy udzielony kredyt zostanie z punktu widzenia banku stracony. Z tych powodów zostały opracowane, i w dalszym ciągu są udoskonalane, liczne metody i modele przewidywania oraz wczesnego ostrzegania przed potencjalnym ryzykiem upadłości przedsiębiorstw (klientów korporacyjnych banków).

W literaturze przedmiotu, zarówno polskiej, jak i światowej, można znaleźć wiele publikacji poświęconych tej problematyce. Pierwsze prekursorskie wyniki badań nad ryzykiem upadłości przedsiębiorstw w warunkach polskich omawiają prace [Mączyńska, 1994; Pogodzińska, Sojak, 1995, s. 53–61]. Później wielu innych autorów przyczyniło się do rozwoju modeli oraz popularyzacji wyników badań z tej tematyki w Polsce. Istotny wkład w tym zakresie wniosła monografia [Hadasik, 1998].

Do analizy i oceny ryzyka upadłości przedsiębiorstw stosuje się obecnie liczne klasyczne modele oparte o metody statystyczne, a także metody niestatystyczne tzw. sztucznej inteligencji. Do najczęściej stosowanych klasycznych metod przewidywania zagrożenia upadłością przedsiębiorstw należy wymienić: analizę dyskryminacyjną [Gajdka, Stos, 1996, s. 59–63], uogólnione modele liniowe (*Generalized Linear Models* – GLM), w tym: regresję logitową oraz probitową [Brożyna i inni, 2016, s. 80–102], drzewa klasyfikacyjne [Lasek, 2007, s. 59–69], metodę k-najbliższych sąsiadów, sieci neuronowe: np. perceptrony wielowarstwowe MLP czy też sieci Kohonena [Korol, 2013, s. 22–30], metodę wektorów nośnych: *Support Vector*

* Dr, Katedra Metod Ilościowych, Wydział Zarządzania, Politechnika Rzeszowska, al. Powstańców Warszawy 10, 35-959 Rzeszów, tpisula@prz.edu.pl

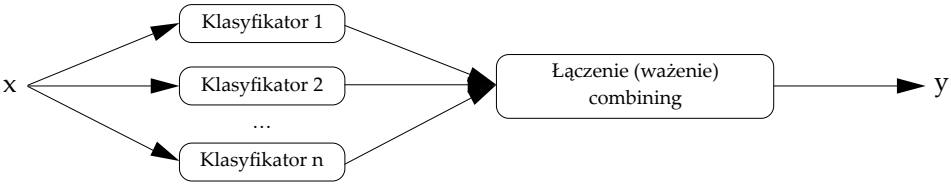
Machines – SVM [Pisula i inni, 2015, s. 7–21], Naive-Bayes klasyfikatory i inne. Z punktu widzenia zastosowań praktycznych bardzo ważne jest, aby stosowane model posiadały jak najlepsze własności klasyfikacyjne, a tym samym posiadały jak najlepszą efektywność i skuteczność w rozpoznawaniu przedsiębiorstw zagrożonych upadłością. W celu poprawy skuteczności klasyfikacyjnej stosowanych modeli z dużym powodzeniem stosuje się ostatnio tzw. zespołowe klasyfikatory typu ensemble. Są to techniki mające na celu łączenie różnych indywidualnych autonomicznych modeli klasyfikacyjnych w celu znacznej poprawy ich skuteczności. W literaturze światowej można znaleźć bardzo dużo publikacji opisujących teorię i praktykę zastosowania tego typu podejścia np. [Fedorova i inni, 2013, s. 7285–7293], jednak brak jest publikacji z tego zakresu w literaturze polskiej.

W artykule przedstawiono wyniki autorskich badań dotyczących możliwości wykorzystania technik ensemble klasyfikatorów do oceny ryzyka upadłości firm z sektora produkcyjnego działających na Podkarpaciu. Analizowane modele klasyfikatorów typu ensemble zostały wykorzystane do oszacowania prawdopodobieństwa upadłości w rocznym horyzoncie czasu (dla roku 2016) dla 144 przedsiębiorstw sektora produkcyjnego z województwa podkarpackiego i oceny ryzyka ich upadłości.

1. Klasyfikatory ensemble i ich zastosowanie do oceny ryzyka upadłości przedsiębiorstw

Główna idea metodologii klasyfikatorów zespołowych (tzw. *ensembles of classifiers*) polega na zwiększaniu efektywności klasyfikacyjnej (zdolności dyskryminacyjnej w poprawnym rozróżnianiu prognozowanych klas) dla stosowanych bazowych klasyfikatorów indywidualnych, poprzez odpowiednie agregowanie lub ważenie wyników klasyfikacji dla poszczególnych modeli składowych, w celu uzyskania klasyfikatorów wynikowych o jak najlepszych własnościach (przewyższających wszystkie klasyfikatory bazowe). Model funkcjonowania klasyfikatorów grupowych typu ensemble przedstawia rysunek 1.

Rysunek 1. Schemat funkcjonowania klasyfikatorów grupowych typu ensemble



Źródło: Opracowanie własne.

Wyczerpujący opis metodologii oraz charakterystykę zespołowych klasyfikatorów ensemble, a także ich klasyfikację i możliwości praktycznych zastosowań można znaleźć w monografiach: [Zhang, Ma, 2012; Zhou, 2012]. W praktyce stosowania tej metodologii stosuje się najczęściej trzy typy podejść: *boosting*, *bagging* oraz *stacking*. Terminologia klasyfikatorów ensemble typu *boosting* odnosi się do szerokiej klasy metod i algorytmów, które potrafią wzmocnić „słabe klasyfikatory" i zamienić je na „klasyfikatory silne" (o bardzo dobrych własnościach klasyfikacyjnych zbliżonych do perfekcyjnych). Przykładem tego typu podejścia jest algorytm adaptacyjnego boostingu *AdaBoost* [Freund, Schapire, 1997, s. 119–139]. Jako klasyfikatory składowe stosuje się tutaj klasyfikatory tego samego typu, np. wzmacniane drzewa klasyfikacyjne. Określenie kolejnej klasy metod ensemble, a więc *bagging*, pochodzi od angielskiego skrótu Bootstrap AGGregatING [Breiman, 1996, s. 123–140]. Metodologia ta wykorzystuje techniki bootstrapowego próbkowania do uzyskania podzbiorów uczących dla składowych klasyfikatorów oraz stosuje najczęściej tzw. strategie głosowania (*voting*) dla agregowania ich wyjściowych klasyfikacji: np. *majority voting*, *plurality voting*, *weighted voting* lub *soft voting*. Jako klasyfikatory składowe stosuje się tutaj także klasyfikatory tego samego typu, np. losowy las (*random forest*). Trzecią grupę metod ensemble stanowią tzw. metody kombinowane dla wyników klasyfikacji klasyfikatorów składowych. Do grupy tych metod zalicza się: podejście typu *averaging*, czyli proste lub ważone uśrednianie wyników, podejście typu *voting* z wykorzystaniem strategii głosowania oraz podejście typu *stacking*. W metodologii ensemble typu *stacking* [Wolpert, 1992, s. 241–259] stosuje się podejście kombinacyjne, w którym bazowe klasyfikatory (tzw. 1 poziomu) są trenowane na próbach losowych, a następnie ich wyniki klasyfikacji (ich funkcje klasyfikacyjne) są wykorzystywane jako próby uczące dla meta-klasyfikatora (tzw. 2 poziomu) i agregowane w klasyfikacje wynikowe.

2. Metody doboru zmiennych diagnostycznych w zagadnieniach klasyfikacyjnych

Bardzo istotnym problemem występującym w zagadnieniach klasyfikacyjnych, w tym w szczególności w zagadnieniach przewidywania zagrożenia upadłością (zagrożenia niedopełnienia zobowiązań spłaty kredytowej) dla przedsiębiorstw, jest problem wyboru właściwego podzbioru zmiennych predykcyjnych, tzw. *feature selection problem* [John i inni, 1994, s. 121–129; Jović i inni, 2015, s. 1200–1205]. Metody doboru najlepszego podzbioru potencjalnych zmiennych diagnostycznych dla zagadnień klasyfikacyjnych można ogólnie podzielić na dwa podstawowe typy: metody filtracyjne (*filter methods*) oraz tzw. metody typu wrapper (*wrapper methods*). Metody

filtracyjne polegają na wyborze zmiennych o jak najlepszych zdolnościach predykcyjnych, niezależnie od klasyfikatora, pomijając tym samym ich wzajemne oddziaływanie z klasyfikatorem. Są to najczęściej techniki jednowymiarowe (*univariate filters*), w których każda zmienna jest rozpatrywana oddzielnie i rangowana względem ustalonej miary efektywności (jakości) klasyfikacyjnej, pomija się tym samym możliwe współzależności z innymi predyktorami, co może prowadzić do gorszych wyników. Pozbawione tej niedogodności są metody filtrowania wielowymiarowego (*multivariate filters*), które oszacowują potencjalną zdolność predykcyjną całych wejściowych podzbiorów czynników. Metody typu *wrapper* są technikami, które analizują możliwe podzbiory predyktorów oraz ustalają ich efektywność oddziaływania na zmienną zależną modelu na podstawie: przyjętego algorytmu poszukiwania (*search algorithm*), najlepszego podzbioru zmiennych i zastosowanej metody klasyfikacji. Aby przeszukać wszystkie podzbiory predyktorów, algorytm poszukiwania jest jakby owijany (*wrapped*) wokół modelu klasyfikacyjnego, stąd też pochodzi nazwa tej grupy metod.

Spośród metod filtracyjnych z grupy metod jednowymiarowych technik doboru zmiennych predykcyjnych można wyróżnić metody doboru czynników wykorzystujące:

- miary oparte na entropii (zdolności informacyjnej), są to metody wykorzystujące takie miary, jak: *information gain*, *gain ratio* lub *symmetrical uncertainty*,
- miary wykorzystujące współzależności statystyczne, np. współczynnik V-Cramera, współczynnik Giniego,
- miary bazujące na wartości informacyjnej zmiennych, np. współczynnik *Information Value* (*IV*) lub *Weight of Evidence* (WoE),
- miary bazujące na odległości obiektów w przestrzeni, np. metoda *ReliefF*, która nadaje zmiennym odpowiednie wagi, według lokalnie określanej zdolności dyskryminacyjnej dla pewnego podzbioru zmiennych.

Spośród metod filtracyjnych wielowymiarowych można stosować np. następujące metody:

- rekurencyjnej eliminacji zmiennych (*Recursive Feature Elimination* – RFE),
- analizy spektralnej doboru zmiennych (*Spectral Feature Selection* – SFS).

W metodach doboru zmiennych typu wrapper stosowane są różne metody poszukiwania optymalnego podzbioru predyktorów. Metody te można podzielić na dwie zasadnicze grupy: metody deterministyczne (*deterministic*) i losowe (*randomized*). W grupie metod deterministycznych stosuje się różnego rodzaju algorytmy sekwencyjne, np. selekcji krokowej postępującej lub eliminacji krokowej wstecznej. Najczęściej stosowane algorytmy losowe przy wyborze zmiennych w metodach wrapper to: algorytm symulowanego

wyżarzania (*simulated annealing*), algorytmy genetyczne (*genetic algorithm*) lub algorytmy „mrówkowe" (*ant colony optimization algorithms*).

3. Charakterystyka firm z sektora produkcyjnego działających w województwie podkarpackim

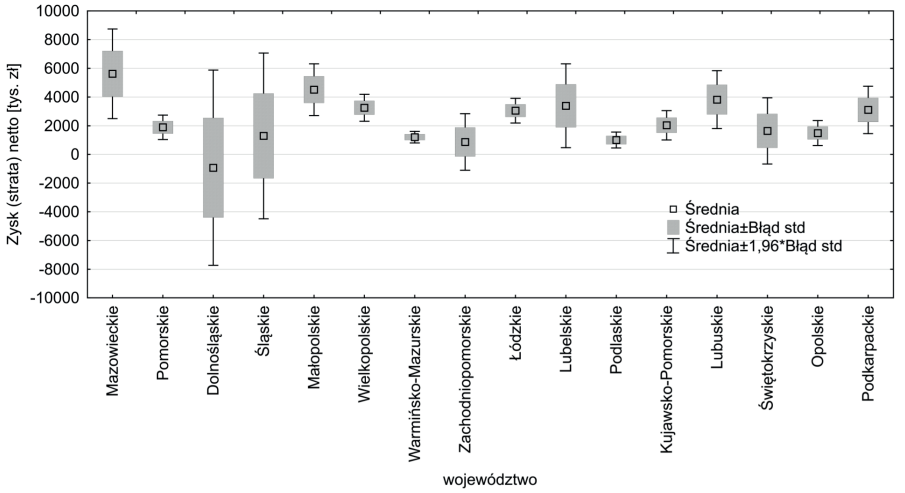
Przedmiotem analiz w badaniach możliwości zastosowania ensemble klasyfikatorów do oceny ryzyka upadłości przedsiębiorstw są firmy sektora produkcyjnego działające w regionie Podkarpacia. Sektor przedsiębiorstw produkcyjnych jest jednym z ważniejszych sektorów w województwie podkarpackim. Na podstawie danych statystycznych z systemu EMIS[1] liczba sklasyfikowanych firm z tego sektora w województwie podkarpackim w 2016 r. wynosiła 762, co stanowiło ponad 19% wszystkich firm działających w województwie. Są to firmy, które przyczyniają się w znaczący sposób do rozwoju regionu, podnoszenia jego prestiżu oraz wzrostu zatrudnienia. W firmach tego sektora było zatrudnionych 85 550 osób (co daje średnią wielkość zatrudnienia na poziomie 112 osób/firmę). Wiele firm z tego sektora działa w ramach Stowarzyszenia Dolina Lotnicza. Jest to stowarzyszenie rozwoju regionu skupiające konsorcjum firm wytwarzających części i komponenty do samolotów i silników lotniczych (np. Pratt & Whitney Rzeszów, Goodrich Aerospace Poland, Safran Transmission Systems Poland, MTU Aero Engines Polska). Największy odsetek, bo aż 68%, stanowiły firmy funkcjonujące na rynku już od dłuższego czasu, 7–16 lat, 31% – to firmy stosunkowo młode, funkcjonujące nie dłużej niż 7 lat, zaś 1% – to firmy funkcjonujące dłużej niż 16 lat. W latach 2015–2016 rozpoczęło działalność ponad 70 nowych firm tego sektora. Jednak równocześnie w kilku ostatnich latach ponad 40 firm z tego sektora zostało zlikwidowanych albo zamkniętych w wyniku upadłości albo restrukturyzacji. Rysunek 2 przedstawia osiągany średni zysk (stratę) netto przez firmy z sektora produkcyjnego w województwach Polski w 2015 r. Największą średnią zyskownością charakteryzowały się firmy z tego sektora w województwach mazowieckim i małopolskim (4–6 mln zł). Województwo podkarpackie wraz z czterema innymi województwami znalazło się w gronie województw z dość dużym średnim zyskiem netto (od 2–4 mln zł), co świadczy o dużej zyskowności firm tego sektora w regionie i ich dobrej kondycji finansowej.

Rysunek 3 przedstawia średnią wartość zobowiązań krótkoterminowych, czyli zadłużenia krótkoterminowego, dla firm sektora produkcyjnego w województwach Polski. Najbardziej zadłużone firmy prowadzą swoją działalność w województwie śląskim, zachodniopomorskim i mazowieckim (średnie zadłużenie krótkookresowe wynosi 30–40 mln zł). Firmy

[1] Emerging Markets Information Service (http://www.emis.com/pl).

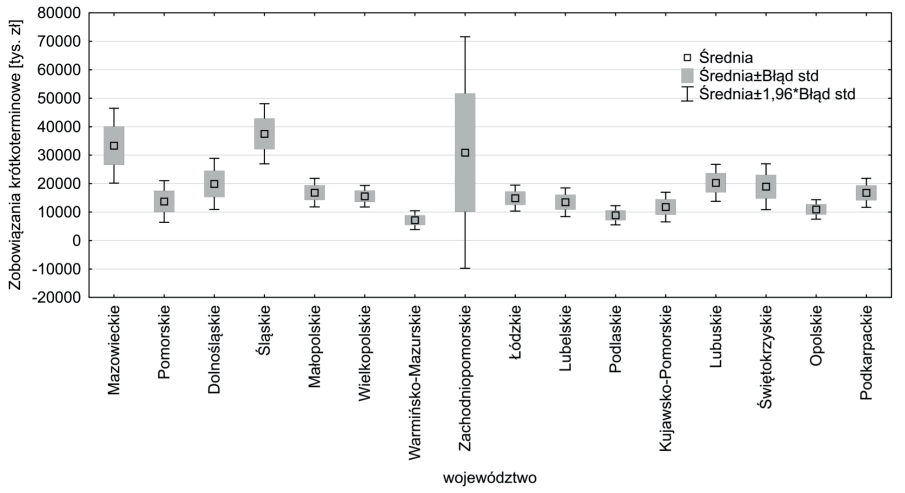
w województwie podkarpackim ze średnim zadłużeniem krótkookresowym na poziomie ponad 16 mln zł należą do grupy województw o stosunkowo niskim poziomie zadłużenia, co również ma wpływ na ich dość dobrą kondycję i stabilność finansową.

Rysunek 2. Średni zysk (strata) netto firm z sektora produkcyjnego w województwach Polski w 2015 r.



Źródło: Opracowanie własne.

Rysunek 3. Średnia wartość zobowiązań krótkoterminowych w firmach z sektora produkcyjnego w województwach Polski w 2015 r.



Źródło: Opracowanie własne.

4. Dobór zmiennych diagnostycznych do modeli klasyfikatorów ensemble

Do badań nad oceną ryzyka upadłości dla firm produkcyjnych z regionu Podkarpacia jako zmienne diagnostyczne wykorzystano 11 wskaźników finansowych. Wykorzystane wskaźniki wybrano metodą filtracji spośród wielu wstępnie wytypowanych wskaźników opisujących: płynność finansową, zyskowność, zadłużenie oraz sprawność działania analizowanych przedsiębiorstw. Ostateczny zbiór predyktorów zawiera tylko te wskaźniki, które posiadały dobre zdolności dyskryminacyjne w rozróżnianiu przedsiębiorstw zagrożonych i niezagrożonych upadłością. Jako miary zdolności dyskryminacyjnych dla zmiennych diagnostycznych wybrano: miarę Symmetrical Uncertainty, współczynnik V-Cramera oraz współczynnik Information Vaulue (IV). Tablica 1 przedstawia zbiór 11 ostatecznie wybranych wskaźników oraz wartości miar określających ich zdolności predykcyjne. Można zauważyć duże wartości tych miar, co świadczy o bardzo dobrych własnościach klasyfikacyjnych wybranych predyktorów.

Tablica 1. Zmienne diagnostyczne wybrane metodą filtracji uporządkowane według malejących wartości wskaźnika Symmetrical Uncertainty

Zmienna (wskaźnik) opis wskaźnika	Miara zdolności dyskryminacyjnej		
	Symmetrical Uncertainty	V-Cramer	Information Value (IV)
X1 – zadłużenia: Zysk brutto / Zobowiązania krótkoterminowe	0,781	0,927	6,292
X2 – ROS (*return on sales*) Rentowność sprzedaży	0,719	0,899	5,402
X3 – ROA (*return on asets*) Rentowność aktywów	0,685	0,918	5,431
X4 – Marża zysku brutto	0,613	0,878	4,596
X5 – dźwignia finansowa: Aktywa / Kapitał własny	0,470	0,848	4,134
X6 – płynności finansowej szybkiej, tzw. Quick Ratio	0,462	0,788	3,394
X7 – płynności: Kapitał pracujący obrotowy / Suma bilansowa	0,397	0,763	2,789
X8 – sprawności: Rotacja zobowiązań [dni]	0,362	0,670	2,307
X9 – zwrot z kapitału: Zysk netto / (Aktywa – Zobowiązania krótkoterminowe)	0,356	0,693	2,486
X10 – rentowności: Zysk netto / (Kapitał własny – Zysk netto)	0,281	0,595	1,728
X11 – sprawności: Cykl środków pieniężnych [dni]	0,226	0,605	1,470

Źródło: Opracowanie własne.

Wartości wskaźników obliczono na podstawie danych ze sprawozdań finansowych firm, które pochodziły z bazy systemu EMIS i obejmowały roczne wskaźniki finansowe przedsiębiorstw z sektora produkcyjnego w okresie 2008–2014 obliczone na 1 rok przed okresem ogłoszenia ich upadłości. Informacje o zdarzeniach upadłości przedsiębiorstw pochodziły z *Ogólnopolskiego Informatora Upadłościowego* systemu EMIS. Zdarzenie ogłoszenia upadłości było podstawą do przyjęcia dla takiego przedsiębiorstwa dla zmiennej zależnej w modelu klasyfikacyjnym (Y=1) – przynależność do klasy przedsiębiorstw upadłych. Jeżeli takie zdarzenie nie występowało, to zmienna zależna w modelu klasyfikacyjnym przyjmowała wartość (Y=0) – przynależność przedsiębiorstwa do klasy przedsiębiorstw zdrowych, czyli niezagrożonych upadłością.

W celu kalibracji zastosowanych modeli klasyfikatorów ensemble oraz do oceny ich efektywności (w rozpoznawaniu przedsiębiorstw zdrowych i upadłych) wykorzystano dwa warianty prób uczących: próba zbalansowana obejmowała jednakową liczbę przedsiębiorstw w obu klasach, natomiast w próbie niezbalansowanej liczebność klasy przedsiębiorstw niezagrożonych ryzykiem upadłości w stosunku do bankrutów była dwa razy większa. W obu próbach liczba przedsiębiorstw upadłych wynosiła 81. Ponadto 70% w obu próbach stanowiła próba ucząco-walidacyjna, zaś 30% próba testowa, która nie była brana pod uwagę przy trenowaniu i kalibracji modeli klasyfikacyjnych. Osobną ostateczną próbę badawczą stanowiły przedsiębiorstwa z województwa podkarpackiego (144 przedsiębiorstwa), dla których dostępne były pełne dane finansowe z 2015 r. w systemie EMIS dla 11 wybranych wskaźników będących zmiennymi diagnostycznymi. Na podstawie tej próby oraz skalibrowanych modeli klasyfikatorów ensemble zostały oszacowane prawdopodobieństwa upadłości dla firm sektora produkcyjnego z Podkarpacia dla 2016 r.

5. Analiza ryzyka upadłości dla podkarpackich firm z sektora produkcyjnego

Jako klasyfikatory bazowe zastosowano pięć znanych i często stosowanych modeli klasyfikacyjnych: liniową analizę dyskryminacyjną (*LDA*), drzewa decyzyjne klasyfikacyjno-regresyjne (*CaRT*), metodę k-najbliższych sąsiadów (*k-NN*), metodę wektorów nośnych z radialnymi funkcjami bazowymi (*SVM Radial*) oraz naiwną metodę Bayesa (*Naive Bayes*). Są to metody dokładnie opisane w literaturze przedmiotu, dlatego szczegóły techniczne tych metod zostaną w pracy pominięte. Jako klasyfikatory ensemble wykorzystano trzy warianty klasyfikatorów zespołowych. W pierwszym wariancie zastosowano metodologię *stacking*, gdzie jako metaklasyfikator 2 poziomu wykorzystano sieć neuronową *NNet* z jedną warstwą neuronów

ukrytych. W drugim wariancie zastosowano jedną z technik wzmacniania klasyfikatorów (*boosting*), jaką jest metoda *GBM* (*stochastic gradiend boosting*). W trzecim wariancie wykorzystano natomiast technikę bootstrapowego agregowania (*bagging*), gdzie wykorzystano jako technikę klasyfikacji losowy las (*random forest*). Do kalibracji rozpatrywanych modeli wykorzystano pakiet obliczeniowy „*R*" oraz specjalistyczne biblioteki *caret* i *caret ensemble*. W celu ustalenia najlepszych wartości parametrów dla trenowanych modeli zastosowano metodę poszukiwania sieciowego (*grid search*) oraz technikę próbkowania: (5-folds) CV (*cross-validation*). Dla potrzeb walidacji oraz oceny dokładności klasyfikacyjnej zastosowanych modeli ensemble oraz modeli bazowych wykorzystano znane miary zgodności:

- pole pod krzywą ROC (*Area Under Curve Receiver Operating Characteristic – AUC ROC*),
- miarę ogólnej dokładności klasyfikacyjnej *Accuracy*:

$$AC = \frac{a+d}{(a+b+c+d)} \tag{1}$$

  gdzie: a – liczba poprawnie zakwalifikowanych przez model rzeczywistych epizodów upadłości, b – liczba zakwalifikowanych niepoprawnie przez model (jako upadłość) rzeczywistych epizodów braku upadłości, c – liczba niepoprawnie zakwalifikowanych przez model (jako zdarzenie braku upadłości) rzeczywistych epizodów upadłości, zaś d – liczba poprawnie zakwalifikowanych przez model rzeczywistych epizodów upadłości,
- statystykę zgodności klasyfikacyjnej *Kappa*:

$$Kappa = \frac{(AC - p_e)}{(1 - p_e)} \tag{2}$$

  gdzie: $p_e = \frac{a+c}{a+b+c+d} \cdot \frac{a+b}{a+b+c+d} + \frac{b+d}{a+b+c+d} \cdot \frac{c+d}{a+b+c+d}$ – jest tzw. zgodnością oczekiwaną.

Tablice 2 i 3 przedstawiają uśrednione wartości statystyk walidacyjnych dla przyjętych klasyfikatorów bazowych oraz dla metaklasyfikatora grupowego ensemble *NNet* w przypadku wykorzystania zbalansowanej i niezbalansowanej próby uczącej (na podstawie której modele były trenowane) z wykorzystaniem techniki (5-folds) *cross validation*. Można zaobserwować, że w obu przypadkach (niezależnie od próby) klasyfikator ensemble posiadał takie same lub lepsze własności klasyfikacyjne (wyższe wartości statystyk walidacyjnych), a tym samym lepiej prognozował upadłość przedsiębiorstw niż modele bazowe.

Tablica 2. Statystyki walidacyjne dla klasyfikatorów bazowych oraz metaklasyfikatora ensemble dla zbalansowanej próby uczącej

Klasyfikator bazowy (jednostkowy)	Accuracy (AC)	Kappa	ROC (AUC)
LDA	0,88	0,76	0,97
CaRT	0,93	0,85	0,96
k-NN	0,92	0,83	0,98
SVM Radial	0,85	0,71	0,99
Naive Bayes	0,90	0,80	0,99
Klasyfikator ensemble metaklasyfikator NNet	0,98	0,96	0,99

Źródło: Opracowanie własne.

Tablica 3. Statystyki walidacyjne dla klasyfikatorów bazowych oraz metaklasyfikatora ensemble dla niezbalansowanej próby uczącej

Klasyfikator bazowy (jednostkowy)	Accuracy (AC)	Kappa	ROC (AUC)
LDA	0,89	0,74	0,98
CaRT	0,94	0,86	0,94
k-NN	0,91	0,79	0,99
SVM Radial	0,90	0,78	0,99
Naive Bayes	0,93	0,84	0,98
Klasyfikator ensemble metaklasyfikator NNet	0,98	0,96	0,99

Źródło: Opracowanie własne.

Tablica 4 przedstawia natomiast wartości tych samych statystyk walidacyjnych, lecz dla zbalansowanej próby testowej. Można również zauważyć, że klasyfikator ensemble NNet przewyższa zdolnościami klasyfikacyjnymi większość zastosowanych modeli bazowych. Porównanie efektywności klasyfikacyjnej wszystkich trzech zastosowanych wariantów klasyfikatorów ensemble dla nierównolicznej próby niezbalansowanej, zarówno uczącej, jak i testowej, przedstawia tablica 5. W przypadku próby testowej najlepsze wyniki uzyskano w przypadku metody losowego lasu. Jednakże wszystkie modele posiadają bardzo wysokie wartości statystyk walidacyjnych, np. dla próby uczącej model GBM w 100% poprawnie sklasyfikował wszystkie przedsiębiorstwa upadłe.

Tablica 4. Statystyki walidacyjne dla klasyfikatorów bazowych oraz meta-klasyfikatora ensemble dla zbalansowanej próby testowej

Klasyfikator bazowy	Accuracy (AC)	Kappa	AC (upadłe)	AC (zdrowe)	ROC (AUC)
LDA	0,96	0,92	1,00	0,92	0,98
CaRT	1,00	1,00	1,00	1,00	1,00
KNN	0,90	0,81	0,85	0,96	0,96
SVM Radial	0,90	0,81	1,00	0,81	1,00
Naive Bayes	0,96	0,92	1,00	0,92	0,97
Klasyfikator ensemble (stacking)					
metaklasyfikator NNet	0,98	0,96	1,00	0,96	1,00

Źródło: Opracowanie własne.

Tablica 5. Statystyki walidacyjne modeli ensemble dla niezbalansowanej próby testowej i uczącej

Klasyfikator ensemble	Acurracy (AC)	Kappa	AC (upadłe)	AC (zdrowe)	ROC (AUC)
Próba testowa					
GBM (boosting)	0,94	0,86	0,89	0,96	0,92
Random Forest (bagging)	0,96	0,92	0,96	0,96	0,99
Klasyfikator ensemble (stacking) meta-klasyfikator NNet	0,94	0,86	0,93	0,95	0,98
Próba ucząca					
GBM (boosting)	0,99	0,98	1,00	0,99	0,99
Random Forest (bagging)	0,96	0,90	0,91	0,98	0,99
Klasyfikator ensemble (stacking) Metaklasyfikator NNet	0,98	0,96	0,96	0,99	0,99

Źródło: Opracowanie własne.

Skalibrowane modele klasyfikatorów ensemble zastosowano do prognozy upadłości firm sektora produkcyjnego z województwa podkarpackiego. Wartości wskaźników obliczono na podstawie danych z 2015 r., gdyż tylko dla tego roku dostępne były pełne dane dla dostatecznie dużej liczby 144 przedsiębiorstw, zaś prognozę wyznaczono na 2016 rok. W populacji 144 firm było tylko 3 firmy, które w rzeczywistości ogłosiły w 2016 r. upadłość. W tablicy 6 przedstawiono macierz klasyfikacji dla badanych przedsiębiorstw przy wykorzystaniu rozpatrywanych modeli klasyfikatorów ensemble. Warto zauważyć, że każdy z modeli poprawnie zakwalifikował

je jako prognozowane upadłe. Ponadto w zależności od zastosowanego modelu 20–38 firm było klasyfikowanych przez modele jako zagrożone potencjalną upadłością.

Tablica 6. Macierz poprawnych klasyfikacji dla zastosowanych trzech typów klasyfikatorów ensemble skalibrowanych na podstawie próby uczącej zbalansowanej dla badanych firm produkcyjnych z regionu Podkarpacia

Klasyfikator ensemble skalibrowany na podstawie próby uczącej zbalansowanej	Macierz poprawnych klasyfikacji		
	Przewidywane	Obserwowane	
		upadłe	zdrowe
GBM (boosting)	upadłe	3	38
	zdrowe	0	103
RF – Losowy Las (bagging)	upadłe	3	32
	zdrowe	0	109
Ensemble metaklasyfikator NNet (stacking)	upadłe	3	20
	zdrowe	0	121

Źródło: Opracowanie własne.

Na podstawie wszystkich modeli ensemble wyznaczono przewidywane ryzyko upadłości firm w rocznym horyzoncie czasu, wyrażone prawdopodobieństwem upadłości (PD – *probability of default*) i obliczono jego wartości średnie. Tablica 7 przedstawia stablicowane klasy ryzyka upadłości dla firm produkcyjnych z Podkarpacia względem średniej wartości rocznego prawdopodobieństwa upadłości. Spośród badanych 144 firm 46 zostało zakwalifikowanych jako firmy o bardzo niskim ryzyku upadłości (poniżej 2%), 54 firmy zostały zakwalifikowane jako posiadające średnie ryzyko upadłości (2%–20%). Natomiast 12 firm to przedsiębiorstwa o dużym ryzyku upadłości (20%–50%), czyli są to już potencjalni bankruci, a 32 firmy charakteryzuje nawet bardzo duże ryzyko upadłości w skali roku (powyżej 50%). Spośród nich 3 firmy już ogłosiły upadłość, a pozostałe mogą to zrobić w najbliższej przyszłości.

Tablica 7. Średnie roczne prawdopodobieństwo upadłości

PD (T=1)	Liczba firm	Charakterystyka – ocena zagrożenia upadłością
<2%	46	niezagrożone upadłością
[2,11]%	36	niezagrożone – ale narażone na zmiany kondycji finansowej
(11–20]%	18	aktualnie niezagrożone upadłością – jednak istnieje duże ryzyko zmiany sytuacji (zależnie od warunków ekonomicznych)

PD (T=1)	Liczba firm	Charakterystyka – ocena zagrożenia upadłością
(20–50]%	12	potencjalny bankrut – duże ryzyko niewypłacalności (występują trudności w spłacie pewnych zobowiązań)
>50%	32	upadłość finansowa – bardzo duże ryzyko niewypłacalności

Źródło: Opracowanie własne.

Zakończenie

Przeprowadzone w pracy eksperymenty badawcze z wykorzystaniem prób testowych pokazały, że klasyfikatory ensemble charakteryzują się lepszymi zdolnościami klasyfikacyjnymi niż klasyfikatory bazowe (autonomiczne). Statystyki walidacyjne otrzymane metodą próbkowania CV (*cross validation*) dla klasyfikatorów ensemble znacznie przewyższają efektywność klasyfikacyjną większości modeli bazowych dla prób uczących oraz wszystkich modeli dla prób testowych. Oszacowane średnie prawdopodobieństwa upadłości (dla horyzontu rocznego T=1) pokazują, że firmy z sektora produkcyjnego działające na Podkarpaciu są w przeważającej większości stabilne finansowo i niezagrożone ryzykiem upadłości (aż 100 na 144 badane firmy). Tylko 44 z nich charakteryzuje znaczne ryzyko upadłości, a modele ensemble przewidują w przyszłości dla nich albo bankructwo, albo występowanie aktualnych trudności ze spłatą swoich zobowiązań finansowych i kredytowych. Badania zostaną rozszerzone w przyszłości na inne sektory, a także na inne województwa Polski.

Literatura

Breiman L. (1996), *Bagging predictors*, „Machine Learning", Vol. 24, No. 2.

Brożyna J., Mentel G., Pisula T. (2016), *Statistical methods of the bankruptcy prediction in the logistics sector in Poland and Slovakia*, „Transformations in Business & Economics", Vol. 15, No. 1 (37).

Fedorova E., Gilenko E., Dovzhenko S. (2013), *Bankruptcy prediction for Russian companies: Application of combined classifiers*, „Expert Systems with Applications", Vol. 40, No. 18.

Freund Y., Schapire R.E. (1997), *A decision theoretic generalization of on-line learning and an application to boosting*, „Journal of Computer and System Sciences", Vol. 55, No. 1.

Gajdka J., Stos D. (1996), *Wykorzystanie analizy dyskryminacyjnej w ocenie kondycji finansowej przedsiębiorstw*, w: R. Borowiecki (red.), *Restrukturyzacja w procesie przekształceń i rozwoju przedsiębiorstw*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

Hadasik D. (1998), *Upadłość przedsiębiorstw w Polsce i metody jej prognozowania*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.

John G.H., Kohavi R., Pfleger K. (1994), *Irrelevant Features and the Subset Selection Problem*, w: W.W. Cohen, H. Hirsh (eds.), *Machine Learning. Proceedings of the Eleventh International Conference*, Morgan Kaufmann Publishers, San Francisco.

Jović A., Brkić K., Bogunović N. (2015), *A review of feature selection methods with applications*, w: *Proceedings of 38th International Convention on Information and Communication Technology, Electronics and Microelectronics* (MIPRO), IEEE Xplore.

Korol T. (2013), *Early warning models against bankruptcy risk for Central European and Latin American enterprises*, „Economic Modelling", Vol. 31.

Lasek M. (2007), *Metody Data Mining w analizowaniu i prognozowaniu kondycji ekonomicznej przedsiębiorstw*, Difin, Warszawa.

Mączyńska E. (1994), *Ocena kondycji przedsiębiorstwa*, „Życie Gospodarcze", nr 38.

Pisula T., Mentel G., Brożyna J. (2015), *Non-statistical methods of analysing of bankruptcy risk*, „Folia Oeconomica Stetinensia", Vol. 15, No. 1.

Pogodzińska M., Sojak S. (1995), *Wykorzystanie analizy dyskryminacyjnej w przewidywaniu bankructwa przedsiębiorstw*, „Acta Universitatis Nicolai Copernici – Ekonomia XXV", z. 299.

Wolpert D.H. (1992), *Stacked generalization*, „Neural Networks", Vol. 5, No. 2.

Zhang C., Ma Y. (2012), *Ensemble Machine Learning. Methods and applications*, Springer, New York.

Zhou Z.-H. (2012), *Ensemble Methods. Foundations and algorithms*, CRC Press, Boca Raton.

Streszczenie

Jednym z głównych rodzajów działalności banków oraz innych instytucji kredytowych jest udzielanie i obsługa kredytów dla klientów indywidualnych oraz korporacyjnych. Jest to jedno z podstawowych źródeł dochodów dla banków, jednak wiążące się często z dużym ryzykiem. Banki stosują różnego rodzaju metody oceny ryzyka kredytowego, w tym dla klientów korporacyjnych. Umożliwiają one z dostatecznie dużą dokładnością prognozowanie (w zadanym horyzoncie czasu) prawdopodobieństwa upadłości, a tym samym oszacowanie potencjalnego ryzyka poniesienia strat dla banku, wynikającego z niespłacenia części lub całości kredytu. W praktyce stosowane są różnego rodzaju modele oceny potencjalnego ryzyka kredytowego, wynikającego z możliwości bankructwa przedsiębiorstw. Są to zarówno modele wykorzystujące metody statystyczne, jak i metody niestatystyczne. W celu zwiększenia skuteczności klasyfikacyjnej stosowanych modeli w ostatnich latach z dużym powodzeniem wykorzystuje się tzw. grupowe klasyfikatory typu ensemble. Są to techniki polegające na łączeniu różnych metod klasyfikacyjnych w celu znacznej poprawy skuteczności tych modeli składowych i zwiększenia ich efektywności klasyfikacyjnej. W artykule przedstawiono wyniki autorskich badań nad możliwością zastosowania technik ensemble classifiers do oceny ryzyka upadłości dla firm z sektora produkcyjnego działających na Podkarpaciu. Skuteczność metod ensemble została porównana z wybranymi klasycznymi modelami oceny ryzyka upadłości przedsiębiorstw. Praktycznym aspektem pracy było zastosowanie analizowanych modeli do

oszacowania prawdopodobieństwa upadłości w rocznym horyzoncie czasu dla 144 przedsiębiorstw sektora produkcyjnego z województwa podkarpackiego oraz zbadanie ich potencjalnego zagrożenia upadłością w 2016 r.

Słowa kluczowe
bankructwo przedsiębiorstw, prognozowanie upadłości, prawdopodobieństwo upadłości, klasyfikatory grupowe ensemble

An application of ensemble classifiers to assess the bankruptcy risk of enterprises on the example of manufacturing sector companies operating in the Podkarpacie region (Summary)

One of the main activities of banks and credit institutions is lending and servicing activities for individuals and corporate clients. This is one of the main sources of income for the banking sector, but it sometimes involves very high risks. Contemporary banks use different types of credit risk assessment methods for corporate clients, with a view to reasonably assess the probability of bankruptcy in a given time horizon, and thus evaluate the risk of potential bank losses due to non-payment of a part or the whole loan. In practice, different types of bankruptcy risk assessment models are used successfully (both statistical models and also non-statistical methods). In order to increase the classification efficiency of used models in recent years, the so-called group ensemble classifiers have been used successfully. These are techniques to combine different individual classification methods to significantly improve the effectiveness of models used to predict bankruptcy occurrence in customer credit repayments. The article presents the results of own research on the possibility of using ensemble classifiers to assess the risk of bankruptcy for companies operating in the Podkarpacie region. The effectiveness of the ensemble methods was compared to some classic bankruptcy risk assessment methods. The practical aspect of the work was to use the analysed models to estimate the probability of bankruptcy in the annual time horizon for 144 manufacturing sector companies in the Podkarpacie Voivodeship and to examine their bankruptcy risk for 2016.

Keywords
corporate bankruptcy, bankruptcy prediction, probability of default, ensemble classifiers

Kamila Migdał-Najman*
Sylwia Badowska**

Wykorzystanie samouczących się sieci neuronowych w analizie zachowań zakupowych i identyfikacji ich wzorców wśród konsumentów w wieku 60 lat i więcej

Wstęp

Z początkiem XXI wieku marketingowcy na całym świecie stanęli przed kolejnym wyzwaniem. Związane jest ono z wyłanianiem się nowego segmentu rynkowego, zwanego „srebrnym rynkiem" (*silver market*). Jego kształtowanie jest wynikiem sukcesywnego przebudowywania się struktury demograficznej zarówno krajów rozwiniętych, jak i rozwijających się. Ten uniwersalny i globalny proces polega na dynamicznym wzroście liczby osób w wieku 60 i więcej. Spowodowany został wydłużeniem się przeciętnej długości życia ludności w wyniku m.in. poprawy warunków bytowych i postępu medycznego, końcowo prowadząc do wyewoluowania kolejnego segmentu konsumenckiego. Wstępne badania na świecie (szczególnie w USA i Europie Zachodniej) sugerują, iż segment ten cechuje swoista specyfika i autonomia w obszarze potrzeb i zachowań nabywczych. Zrozumienie rynku srebrnych osób (*silver citizen market*) wymaga pogłębionej wiedzy o postawach i stylu życia seniorów, aczkolwiek marketingowcy wciąż zajmują się tym lukratywnym rynkiem nieefektywnie [de Asis, 2007], jeśli w ogóle jest on pod ich obserwacją czy świadomą eksploracją. Rynek seniorów rozpatrywany jest najczęściej przez producentów kosmetyków, farmaceutyków oraz dostawców usług medycznych [Badowska, Rogala, 2016]. Zdecydowanie rzadziej (jeśli kiedykolwiek) jest rynkiem docelowym dla dóbr i usług technologicznych. Jednakże w świetle koncepcji srebrnej gospodarki (*silver economy*) akcentowanej przez Komisję Europejską [2015], w której rozwój wysokich technologii stanowi nieodzowny element, problematyka wzrostu potrzeb osób starszych może być pozytywnym impulsem do dalszego rozwoju przedsiębiorstw w skali lokalnej, jak i globalnej. Tym samym może stać się ona nowym źródłem inspiracji dla przedsiębiorstw dotąd niedostrzegających rosnącego popytu ze strony starszych konsumentów.

* Prof. UG dr hab., Katedra Statystyki, Wydział Zarządzania, Uniwersytet Gdański, ul. Armii Krajowej 101, 81-824 Sopot, kamila.migdal-najman@ug.edu.pl
** Dr inż., Katedra Marketingu, Wydział Zarządzania, Uniwersytet Gdański, ul. Armii Krajowej 101, 81-824 Sopot, sylwia.badowska@ug.edu.pl

Jednakże nadal w teorii i praktyce brakuje wystarczającej wiedzy o zachowaniach konsumentów w wieku 60 i więcej, co stanowi istotną barierę dla zrozumienia tej specyficznej nowo kształtującej się grupy nabywczej. Celem niniejszego artykułu jest identyfikacja wzorców zachowań zakupowych osób w wieku 60 i więcej w obszarze produktu technologicznego na przykładzie telefonu komórkowego typu smartfon. Do identyfikacji wzorców zakupowych wykorzystano sztuczną sieć neuronową typu GNG [Fritzke, 1994]. Sieć GNG jest efektywnym narzędziem analizy wielowymiarowych zbiorów danych, która znajduje szerokie zastosowanie w grupowaniu, w klasyfikacji i w poszukiwaniu wzorców. W badaniu wykorzystano dane pierwotne, pozyskane w ramach badań sondażowych zrealizowanych w latach 2014–2015 wśród uczestników Uniwersytetu Trzeciego Wieku w województwie pomorskim[1].

Niniejszy artykuł stanowi wprowadzenie do cyklu poświęconego badaniom wzorców zachowań konsumentów w wieku 60 i więcej przy wykorzystaniu sieci neuronowych. Składa się on z czterech części, z których dwie pierwsze posłużyły prezentacji przeglądu literatury przedmiotu, następne zaś analizie wyników badania i podsumowaniu rozważań.

1. Konsumenci 60 i więcej w świetle kształtowania się srebrnego rynku

Proces starzenia się społeczeństw jest wyzwaniem ogólnoświatowym, szczególnie zauważalnym w Europie XXI wieku. Stary Kontynent stanął przed zjawiskiem kształtowania się nowego porządku demograficznego [van de Kaa, 2008, za: Niewiadomska, Sobolewska-Poniedziałek, 2015]. Zmiana profilu demograficznego w Unii Europejskiej widoczna jest w podnoszącej się wartości mediany wieku tej populacji, która na przełomie lat 2000–2013 wzrosła o 3,9 lat, osiągając poziom 41,9 lat [Pauhofova, Dovalova, 2015, s. 191]. Przewiduje się, że proces ten będzie się sukcesywnie pogłębiał, prowadząc do wzrostu mediany w 2050 r. do 46 lat i w 2100 r. do 47 lat [Niewiadomska, Sobolewska-Poniedziałek, 2015, s. 69]. W Polsce oczekuje się również wzrostu intensywności tego zjawiska, w roku 2030 mediana wieku wynosić będzie 46,1 lat, zaś w 2050 już 51,8 lat. Tym samym plasując populację Polaków na 10 miejscu wśród najstarszych społeczeństw świata [ONZ, 2015].

Zaobserwowana dynamika zmian demograficznych wśród Europejczyków stanowi dylemat w obszarze rozwoju gospodarczego kontynentu.

[1] Uniwersytet Trzeciego Wieku (UTW) przy Uniwersytecie Gdańskim w Gdańsku. Filie Gdańskiego Uniwersytetu Trzeciego Wieku UG: Pruszcz Gdański i Pelplin; Kościerski Uniwersytet Trzeciego Wieku; UTW w Gdyni przy Centrum Aktywności Seniora; UTW w Sopocie, UTW w Malborku. Badanie w oddziałach UTW w województwie pomorskim trwało 12 miesięcy.

Jednym z zaproponowanych rozwiązań, wychodzącym naprzeciw tej negatywnej tendencji, jest rozwijana i propagowana przez Komisję Europejską koncepcja srebrnej gospodarki. Jej istotą jest wykorzystanie ekonomicznych możliwości pojawiających się w wyniku zaistnienia publicznych i prywatnych wydatków związanych ze starzeniem się społeczeństwa oraz istnienia swoistych potrzeb ludności w wieku 50 i więcej, i jako takich wchodzących w skład ogólnego rynku konsumenckiego, jednakże cechujących się istotnymi różnicami we wzorcach i priorytetach wydatkowania środków [Komisja Europejska, 2015].

Z wąskiego rozumienia koncepcji srebrnej gospodarki wyłania się idea srebrnego rynku oraz srebrnego przemysłu, w których główny nacisk położono na dostosowanie oferowanych na rynku towarów (i usług) do wymagań i potrzeb osób starszych [Niewiadomska, Sobolewska-Poniedziałek, 2015]. Te zaś z kolei generują zachowania konsumenckie rozumiane jako działania związane z poszukiwaniem, zakupem, użytkowaniem oraz oceną dóbr i usług, które mają zdolność do zaspokojenia potrzeb [Rosa, Perenc, 2011] wynikających z indywidualnego odczuwania potrzeb, i obejmujących całokształt obiektywnie i subiektywnie określonych, racjonalnych i emocjonalnych, świadomych i nieświadomych posunięć w trakcie podejmowania decyzji na rynku dóbr konsumpcyjnych oraz w czasie konsumpcji [Rudnicki, 2001, s. 15].

Wśród determinantów zachowań nabywców wyróżnia się czynniki ekonomiczne, marketingowe, kulturowe, społeczne, psychologiczne i osobowościowe [Rosa, Perenc, 2011]. W grupie tych ostatnich wielu badaczy przedmiotu jako istotną zmienną wskazuje wiek konsumentów. Wykorzystywanie demograficznych zmiennych dla poznania i profilowania konsumentów ma długą tradycję [Roscoe i inni, 1977] i są to parametry często stosowane przez marketingowców ze względu na łatwość ich użycia [Schewe, 1988]. Wiek chronologiczny jest najczęściej wykorzystywany jako zmienna do wydzielania i analizy segmentów rynkowych. Jakkolwiek zmienne demograficzne są tylko dobrym punktem startu do przeprowadzania analiz zachowań osób starszych [Ong i inni, 2008, s. 694]. Na przykład badania psychologiczne wskazują, że wiele osób w wieku starszym nie identyfikuje się z wiekiem metrykalnym, czując się o wiele młodziej [Szmigin, Carrigan, 2000]. Ponadto starsi konsumenci mogą wyrażać szerokie spektrum różnorodnych tożsamości, których zrozumienie jest pożądane celem tworzenia lepszych działań marketingowych [Szmigin, Carrigan, 2001]. Starsi konsumenci wydają się być różni od konsumentów młodszych wiekiem z powodu odmiennych motywów i zachowań osób starszych [Moschis i inni, 2004]. Przykładowo, starsze kobiety chętniej szukają informacji, próbują nowych marek i dokonują zakupów mniej ostrożnie niż

młodsze klientki [Birtwistle, Tsim, 2005]. Starsi konsumenci akceptują nowe zwyczaje lub nabywają nowy produkt, dopiero gdy odczuwają wrażenie uzyskiwania wymiernych korzyści z ich nabycia [Schiffman, Sherman, 1991]. Konsumenci starsi wiekiem nie kupują innowacji dla ich nowości; prawdopodobnie są oni wobec nich bardziej krytyczni, dokonując porównania konkretnej oferty z konkurencyjnymi propozycjami [Laventhal, 1997; Lunsford, Burnett, 1992; Szmigin, Carrigan, 2000]. Ponadto są istotnie wrażliwi na komunikaty typu „word-of-mouth" [East i inni, 2014].

Tak zwany srebrny rynek jest segmentem lukratywnym [de Asis, 2007]. W roku 2010 w ujęciu globalnym wartość wydatków konsumentów w wieku 60 i więcej wynosiła 8 bln USD, a szacunkowe dane sugerują, że w roku 2020 wyniosą 15 bln USD [AT Kearney, 2010]. Konsumenci tego segmentu dysponują wolnymi środkami i nie wahają się z nich korzystać [Reisenwitz, Iyer, 2007] i tak np. w krajach rozwiniętych ta grupa nie ponosi obciążeń z tytułu utrzymania dzieci i zobowiązań hipotecznych [Myers, Lumbers, 2008].

Wciąż jednak pokutuje powszechne przekonanie, że rynek konsumentów 60 i więcej jest nieatrakcyjny dla przedsiębiorstw dostarczających rozwiązań opartych o technologie informacyjne [Hough, Kobylanski, 2009]. Starsze osoby powinny być wspierane w nowych technologiach poprzez wyposażanie ich w niezbędne informacje, wsparcie edukacyjne, jak również kastomizację produktów komunikacyjnych, aby dopasować je do specyficznych potrzeb tej grupy nabywczej [Hassan i inni, 2017].

Zatem istotne jest zrozumienie wzorców zachowań starszych konsumentów, jak również kryteriów ewaluacyjnych, jakimi posługują się w procesach decyzyjnych na rynku osoby starsze, co pozwoli na generalne zrozumienie szerokich zmian zachowań konsumpcyjnych obserwowanych na srebrnym rynku [Ong i inni, 2008].

2. Samouczące się sieci neuronowe w analizie i identyfikacji wzorców zakupowych

Samouczenie się sieci neuronowych to bezwzorcowy proces odwzorowania wielowymiarowej przestrzeni wejściowej obiektów w niskowymiarową przestrzeń małej liczy jednostek funkcjonalnych, neuronów, z zachowaniem topograficznego podobieństwa obiektów. Sztuczna sieć sama z siebie doskonali sposób rozwiązywania określonego zadania, nie mając do dyspozycji wzorców zadań wraz z rozwiązaniami. Okazuje pewien poziom samoorganizacji. Dążenie do samoorganizacji jest cechą samouczącej się sieci neuronowej. Proces samouczenia i samoorganizacji zachodzi tylko wtedy, jeśli w wejściowym ciągu danych uczących istnieje pewna prawidłowość, na której może oprzeć się sieć [Tadeusiewicz, 1998; Duch i inni,

2000; Masters, 1995]. Jedną z bardziej znanych nienadzorowanych modeli sztucznych sieci neuronowych jest sieć typu GNG (*Growing Neural Gas*), zaproponowana w 1994 r. przez B. Frtizke [Fritzke, 1994]. Wykorzystywana jest przede wszystkim do grupowania wielowymiarowych zbiorów danych. Sieć GNG z założenia nie posiada stałej czy założonej *a priori* struktury. Struktura sieci tworzona jest w procesie samouczenia się sieci, począwszy od struktury najprostszej, złożonej jedynie z dwóch neuronów, do struktury bardziej złożonej. W procesie samouczenia struktura sieci zmienia się dynamicznie w taki sposób, że nowe neurony są wstawiane do sieci jedynie w tym miejscu sieci, w którym występuje największy błąd rozpoznawania wzorców (tj. błąd kwantyzacji). Dla sieci ustalana jest jedynie maksymalna złożoność sieci (tj. maksymalna liczba neuronów), a także maksymalny wiek neuronów, które nie biorą udziału w procesie samouczenia. W początkowej fazie procesu samouczenia, gdy sieć reprezentowana jest przez niewielką liczbę neuronów, zmiany położenia neuronów w sieci są szybkie i relatywnie duże. Zmiany te są coraz mniej gwałtowne, aż w końcowej fazie następuje uspokojenie sieci GNG. Podstawowe własności sieci GNG przedstawiono w tablicy 1. Sieć GNG znajduje dzisiaj zastosowanie w różnych dziedzinach i dyscyplinach nauki, np. w grupowaniu, analizie zwyczajów zakupowych [Decker, Monien, 2003; Decker, 2005; Migdał-Najman, 2011] w diagnostyce medycznej [Netto i inni, 2012], w rozpoznawaniu wzorców: kształtu liter [Datta i inni, 2001], konturów obrazu, detekcji ruchu, procesie kontroli urządzeń, naukach społecznych i innych [Memmert, Perl, 2009].

Tablica 1. Własności sieci GNG

Własności	Sieć GNG
Struktura sieci	zmienna
Liczba krytycznych parametrów sterujących	4
Jakość grupowania przy optymalnych parametrach	znakomita
Szybkość uczenia	szybka
Wymagania pojemności pamięci	mała
Dowolna konfiguracja skupień	tak
Martwe neurony	nie
Skręcenie się sieci	nie
Rozmycie skupień	dopuszczalne (możliwe połączenie sąsiednich skupień)
Skupienia nieseparowalne	niedopuszczalne (połączenie wszystkich nieseparowalnych skupień)
Wizualizacja danych wielowymiarowych	nie
Wizualizacja sieci	nie

Źródło: Opracowanie własne.

Do identyfikacji wzorców zakupowych uczestników Uniwersytetu Trzeciego Wieku wykorzystano sieć GNG. Przy budowie sieci GNG testowano następujące parametry: parametr lambda = 190 (jest to liczba iteracji, po której jest wstawiany nowy neuron do sieci), maksymalny wiek połączenia = 288 (informujący o tym, ile iteracji pozostaje w sieci neuron, który się nie uczy), krok uczenia neuronu zwycięzcy = 0,05, krok uczenia neuronów połączonych ze zwycięzcą = 0,0006. Analizowana sieć GNG składała się z 79 neuronów[2].

3. Zachowania i wzorce zakupowe konsumentów 60 i więcej nabywających telefony komórkowe typu smartfon

Celem badania była identyfikacja wzorców zachowań zakupowych osób w wieku 60 i więcej w obszarze produktu technologicznego na przykładzie telefonu komórkowego typu smartfon. Za przedmiot badania obrano następujące trzy obszary: generalne zachowanie konsumentów podczas zakupu nowego produktu; generalne zachowanie konsumentów podczas zakupu produktu innowacyjnego i generalne zachowanie konsumentów podczas zakupu produktu innowacyjnego „smartfon".

Do pomiaru powyższych konstruktów użyto 16 zmiennych opisujących: generalne zachowanie konsumentów podczas zakupu nowego produktu; postawy konsumentów wobec innowacyjnego produktu; używanie lub nieużywanie telefonu komórkowego; identyfikację użytkowania typu telefonu komórkowego; rodzaje zachowań konsumentów podczas zakupu telefonu komórkowego typu smartfon; postawy konsumentów wobec telefonu komórkowego typu smartfon; relatywny czas rozpoczęcia użytkowania przez konsumentów nowości w odniesieniu do zachowań osób z otoczenia; zainteresowanie zakupem nowej wersji telefonu komórkowego typu smartfon; relatywną częstotliwość zakupu nowości w odniesieniu do doświadczeń otoczenia; rozważenie zakupu nowej wersji produktu bez pozyskanej wiedzy o jej cechach; relatywną szybkość rozpoczęcia użytkowania nowości w odniesieniu do otoczenia; relatywną szybkość pozyskiwania wiedzy o nowej wersji produktu w odniesieniu do otoczenia; zainteresowanie nowinkami technologicznymi w telefonii komórkowej; częstotliwość odwiedzania działów w sklepach/supermarketach z telefonami komórkowymi; subiektywny poziom wiedzy o telefonach komórkowych i zainteresowanie nabyciem nowej wersji produktu.

W pozyskaniu danych pierwotnych zastosowano technikę sondażu, uzyskując wyniki respondentów przy zastosowaniu papierowego kwestionariusza ankiety. Pytania w badaniu zaadaptowano z kwestionariusza

[2] Więcej na temat sztucznej sieci GNG, parametrów sterujących siecią i własnościach sieci [Migdał-Najman, Najman, 2013].

Venkatesha i zespołu [2003] oraz Szmigin i Carrigan [2000]. Po translacji na język polski wprowadzono niezbędną adaptację do warunków lokalnych i dokonano ewaluacji poprawności ankiety w ramach pilotażu. Następnie przeprowadzono badanie właściwe.

Do identyfikacji występowania zjawiska w przypadku 10 stwierdzeń posłużono się 5-stopniową skalą Likerta (tab. 2), w której oznaczono 1 – zdecydowanie się nie zgadzam, zaś 5 – zdecydowanie się zgadzam. Do oceny stwierdzeń wykorzystano odpowiedzi jednokrotnego wyboru spośród dostępnej respondentom kafeterii.

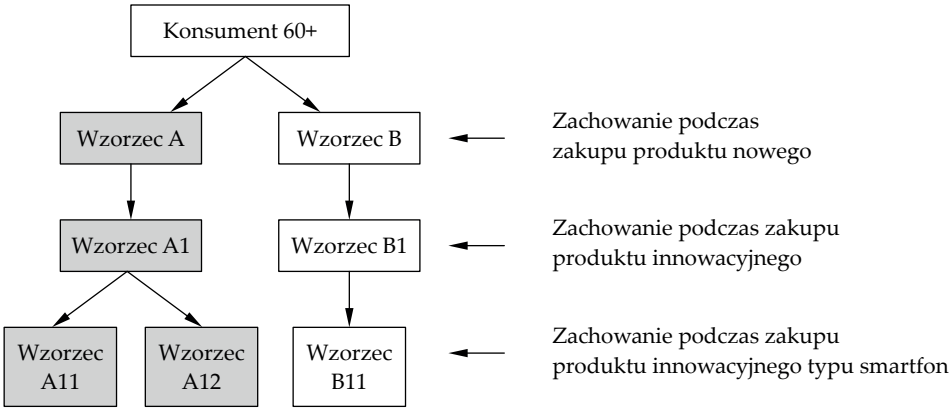
Badaniem objęto słuchaczy siedmiu oddziałów Uniwersytetów Trzeciego Wieku w województwie pomorskim. W okresie listopad 2014–październik 2015 rozdystrybuowano (tj. przekazano respondentom papierowe kwestionariusze w czasie wykładów) łącznie 720 kwestionariuszy. Zwrócono 523 wypełnionych kwestionariuszy. Wśród badanych 79% stanowiły kobiety; 90% respondentów zadeklarowała posiadanie telefonu komórkowego; 65% wskazała, że używa telefonu starego typu tzw. klawiszowy, 25% ankietowanych używa aparat nowego typu z ekranem dotykowym, nazywany smartfonem, 10% badanych nie udzieliła odpowiedzi na to pytanie. Do kolejnego etapu badania, tj. do identyfikacji wzorców zakupowych, wybrano uczestników, którzy zadeklarowali używanie telefonu komórkowego typu smartfon (tj. 25% badanych).

Użytkownikiem produktu innowacyjnego typu smartfon wśród uczestników Uniwersytetu Trzeciego Wieku jest przede wszystkim kobieta, w wieku 60–69 lat, mająca wykształcenie wyższe, będąca pracownikiem umysłowym. Na rysunku 1 zaprezentowano podstawowe wzorce zachowań zakupowych konsumentów 60 i więcej w stosunku do produktu nowego, produktu innowacyjnego i produktu innowacyjnego typu smartfon.

Zachowanie podczas zakupu produktu nowego

Wśród konsumentów 60 i więcej można wyróżnić dwa wzorce zachowań. Nazwijmy je: wzorzec A i wzorzec B. W grupie pierwszej (wzorzec A) decyzje o zakupie nowych, nieznanych dotąd produktów trwają długo, są poprzedzone rozważaniami, rozmowami z rodziną i/lub znajomymi a także poszukiwaniem informacji w różnorodnych źródłach: w prasie, radio, TV, Internecie. W grupy drugiej (wzorzec B) decyzje o zakupie nowych, nieznanych dotąd produktów są wynikiem zakupu zwyczajowego, np. pragnienie posiadania nowej wersji produktu (rys. 1).

Rysunek 1. Wzorce zakupowe konsumentów 60 i więcej wobec produktu nowego, innowacyjnego i typu smartfon



Źródło: Opracowanie własne.

Zachowanie podczas zakupu produktu innowacyjnego

Konsumenci z grupy pierwszej (wzorzec A) deklarują, że w stosunku do produktu innowacyjnego decyzja o zakupie jest podejmowana dopiero po okresie, kiedy przyjaciele/znajomi/rodzina już je nabędą i przetestują, a potem zarekomendują (wzorzec A1). Konsumenci z grupy drugiej (wzorzec B), w stosunku do produktu innowacyjnego nie wykazują jednej wyraźnej postawy, jednego typowego zachowania. Deklarowane są przez konsumentów z tej grupy różne zachowania (wzorzec B1).

Zachowanie podczas zakupu produktu innowacyjnego typu smartfon

Zachowanie konsumentów 60 i więcej w stosunku do produktu innowacyjnego typu smartfon w grupie pierwszej można podzielić na dwie postawy: wzorzec A11 i wzorzec A12. Konsumenci z grupy drugiej (wzorzec B), w stosunku do produktu innowacyjnego typu smartfon deklarują jedną postawę (wzorzec B11).

Wśród konsumentów z grupy wzorzec A11 decyzje o zakupie produktu innowacyjnego typu smartfon trwają długo, są poprzedzone rozważaniami, rozmowami z rodziną i/lub znajomymi, a także poszukiwaniem informacji w różnorodnych źródłach: w prasie, radiu, TV, Internecie. Wśród konsumentów z grupy wzorzec A12, znaleźli się użytkownicy, którzy produkt innowacyjny typu smartfon otrzymali w prezencie. Nie ponosili kosztów jego zakupu. Konsumenci z grupy wzorzec B11 deklarują, że decyzje o zakupie produktu innowacyjnego typu smartfon były wynikiem zakupu zwyczajowego, np. pragnieniem posiadania nowej wersji produktu. Grupa ta jest generalnie zainteresowana nowinkami w telefonii komórkowej.

Typowe zachowania konsumentów 60 i więcej podczas zakupu produktu innowacyjnego typu smartfon zaprezentowano w tablicy 2. W tablicy zakreślono ten wariant na skali 1–5, który badani respondenci wskazywali najczęściej. Wśród badanych dominowały stwierdzenia o negatywnym zabarwieniu, w szczególności te, które wskazywały brak zainteresowania nowymi wersjami i funkcjami telefonów typu smartfon, a także punktami, które zajmują się sprzedażą takich telefonów.

Tablica 2. Typowe zachowania podczas zakupu produktu innowacyjnego typu smartfon

Pytania	1	2	3	4	5
Generalnie w kręgu osób mi bliskich (rodzina/ przyjaciele/ znajomi) jestem ostatnią osobą, która kupiła/rozpoczęła użytkowanie telefonu komórkowego typu smartfon.	X				
Jeśli usłyszę o nowej wersji telefonu smartfon, a który byłby dostępny w sprzedaży w sklepach, był(a)bym zainteresowany(a) jego zakupem.	X				
W porównaniu z rodziną, przyjaciółmi, znajomymi rzadko kupuję/użytkuję najnowsze wersje telefonów smartfon.				X	
Rozważył(a)bym zakup nowej wersji telefonu smartfon, nawet jeśli nie znał(a)bym jego nowych funkcji, wyglądu i sposobu użytkowania.	X				
Generalnie jestem ostatnią osobą w moim środowisku (rodzina, przyjaciele, znajomi), która używa najnowszej wersji telefonu smartfon.			X		
Wiem o nowych wersjach telefonów smartfon o wiele wcześniej niż pozostałe osoby z mojego środowiska (rodziny, przyjaciół, znajomych).		X			
Generalnie jestem zainteresowany(a) nowinkami technologicznymi w telefonii komórkowej.		X			
Często odwiedzam działy w sklepach/ supermarketach z telefonami komórkowymi.	X				
Wiem więcej o telefonach komórkowych typu smartfon niż ludzie wokół mnie.		X			
Jeśli potrzebował(a)bym używać telefonu smartfon kupiał(a)bym najnowszy model, który jest dostępny w sprzedaży.		X			

Odpowiedzi podane na skali Likerta, gdzie: 1 – zdecydowanie się nie zgadzam, 5 – zdecydowanie się zgadzam.
Źródło: Opracowanie własne.

Zakończenie

Artykuł przedstawia propozycję analizy wzorców zachowań nabywczych konsumentów przy wykorzystaniu sztucznej sieci neuronowej typu GNG. Sieć GNG posiada duży potencjał w rozpoznawaniu takich ukrytych wzorców zachowań nabywczych i umożliwiła zidentyfikowanie ich wśród osób w wieku 60 i więcej w obszarze zakupu produktu technologicznego na przykładzie telefonu komórkowego typu smartfon. W wyniku zastosowania samouczącej się sieci neuronowej typu GNG zidentyfikowano trzy specyficzne wzorce zachowań nabywczych, tj.:

- konsumenci „rozważający", którzy decyzję o zakupie smartfona podejmują po długich rozważaniach, którym towarzyszą konsultacje z osobami z bliskiego otoczenia oraz poszukiwania informacji w różnorodnych źródłach (np. w prasie, radio, TV, Internecie);
- konsumenci „zwyczajowi", którzy decyzję o zakupie podejmują w krótkim okresie od momentu pojawienia się produktu na rynku, gdyż zakup nowości jest dla nich zakupem zwyczajowym; jest to grupa nabywcza generalnie zainteresowana nowinkami technologicznymi na rynku telefonii komórkowej;
- konsumenci „obdarowani", którzy są użytkownikami smartfonów, jednakże weszli w ich posiadanie w wyniku otrzymania produktu w postaci prezentu lub podarunku od osób bliskich z ich otoczenia.

Pozyskane wyniki mają implikacje zarówno teoretyczne, jak i praktyczne. W warstwie teoretycznej pogłębiona została wiedza o strukturze rynku konsumentów w wieku 60 i więcej, szczególnie w aspekcie zrozumienia tego segmentu rynkowego z uwzględnieniem zachowań nabywczych produktów technologicznych. Wyniki sugerują, iż segment konsumentów 60 i więcej nie jest jednorodny i można w jego obrębie wydzielić mikrosegmenty o specyficznych cechach. W obszarze praktycznym warto zwrócić uwagę na konstrukcje przekazów marketingowych dedykowanych konsumentom w wieku 60 i więcej w zależności od mikrosegmentu. Na przykład dla konsumentów „rozważających" należy wykorzystywać do budowania komunikatu np. osoby referencyjne będące w bliskim otoczeniu nabywców 60 i więcej oraz umieszczać szczegółowe informacje o produkcie w mediach dedykowanych tej grupie konsumenckiej. Dla konsumentów „zwyczajowych" warte podkreślania w przekazie jest, iż produkt jest im niezbędny w tzw. życiu codziennym, a jego zakup ma charakter rutynowy. Wyniki sugerują, że również w działalności marketingowej nie należy pomijać osób „obdarowanych" niedokonujących samodzielnie zakupu smartfonów, ale otrzymujących je w postaci prezentu. W przypadku tej grupy np. komunikat marketingowy powinien być dwukierunkowy, tj. dedykowany do

osób zakupujących prezent oraz bezpośrednio użytkowników w wieku 60 i więcej.

Literatura

Badowska S., Rogala A.(2016), *Identifying the portrayal of 50+ consumers in Polish print advertising*, „Economics and Business Review", Vol. 2(16), No. 4.

Birtwistle G., Tsim C. (2005), *Consumer purchasing behaviour: an investigation of the UK mature women's clothing marke*t, „Journal of Consumer Behaviour", Vol. 4, No. 6.

Datta A., Parui S.K., Chaudhuri B.B. (2001), *Skeletonization by a topology-adaptive self-organizing neural network*, „Pattern Recognition", Vol. 34, No. 3.

de Asis K. (2007), *Tapping the gray market*, Brand Papers, www.brandchannel.com.

Decker R. (2005), *Market basket analysis by means of a growing neural network*, „International Review of Retail, Distribution and Consumer Research", Vol. 15, No. 2.

Decker R., Monien K. (2003), *Market basket analysis with neural gas networks and self-organizing maps*, „Journal of Targeting, Measurement and Analysis for Marketing", Vol. 11, No. 4.

Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R. (red.) (2000), *Sieci neuronowe*, PAN, Akademicka Oficyna Wydawnicza Exit, Warszawa.

East R., Unclesc M.D., Lomax W. (2014), *Hear nothing, do nothing: The role of word of mouth in the decision-making of older consumers*, „Journal of Marketing Management", Vol. 30, No. 7–8.

Fritzke B. (1994), *Growing cell structures – a self-organizing network for unsupervised and supervised learning*, „Neural Networks", Vol. 7, No. 9.

Hassan Z., Rosnah M.Y., Sapuan M.S., Sharifah N., Rizal H.M. (2017), *Assessing the usability and ergonomic considerations on communication technology for older Malaysians*, „Universal Access in the Information Society", Vol. 16, No. 2.

Hough M., Kobylanski A. (2009), *Increasing elder consumer interactions with information technology*, „Journal of Consumer Marketing", Vol. 26, No. 1.

Kearney A.T. (2010), *What do mature consumers want?*, Global Business Policy Council, AT Kearney, Virginia.

Komisja Europejska (2015), *Growing the Silver Economy*, Background paper.

Leventhal R.C. (1997), *Aging consumers and their effects on the market place*, „Journal of Consumer Marketing", Vol. 14, No. 4.

Lunsford D.A., Bumet M.S. (1992), *Marketing Product Innovations to the Elderiy: Understanding the Barriers to Adoption*, „Journal of Consumer Marketing", Vol. 9, No. 4.

Masters T. (1995), *Advanced algorithms for neural networks*, Wiley, New York.

Memmert D., Perl J. (2009), *Analysis and simulation of creativity learning by means of artificial neural networks*, „Human Movement Science", Vol. 28, No. 2.

Migdał-Najman K. (2011), *Analiza porównawcza samouczących się sieci neuronowych typu SOM i GNG w poszukiwaniu reguł asocjacyjnych*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, „Taksonomia", nr 18, „Klasyfikacja i analiza danych – teoria i zastosowania", nr 176.

Migdał-Najman K., Najman K. (2013), *Samouczące się sztuczne sieci neuronowe w grupowaniu i klasyfikacji danych. Teoria i zastosowania w ekonomii*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.

Moschis G.P., Curasi C., Bellenger D. (2004), *Patronage motives of mature consumers in the selection of food and grocery stores*, „The Journal of Consumer Marketing", Vol. 21, No. 2–3.

Myers H., Lumbers M. (2008), *Understanding older shoppers: a phenomenological investigation*, „Journal of Consumer Marketing", Vol. 25, No. 5.

Netto S.M.B., Silva A.C., Nunes R.A., Gattass M. (2012), *Automatic segmentation of lung nodules with growing neural gas and support vector machine*, „Computers in Biology and Medicine", Vol. 41, No. 11.

Niewiadomska A., Sobolewska-Poniedziałek E. (2015), *Srebrna gospodarka – nowy paradygmat rozwoju starzejącej się Europy*, „Ekonomia XXI w.", nr 3.

Ong F.S., Kitchen J.Ph., Jama A.T. (2008), *Consumption patterns and silver marketing: an analysis of older consumers in Malaysia*, „Marketing Intelligence & Planning", Vol. 26, No. 7.

ONZ (2015), *World Population Prospects: The 2015 Revision*, https://esa.un.org/unpd/wpp/; dostęp: 17.02.2017.

Pauhofova I., Dovalova G. (2015), *Potential of silver economy in the European Union (Selected Views)*, „European Scientific Journal", August, Special Edition.

Reisenwitz T., Iyer R. (2007), *A comparison of younger and older baby boomers: investigating the viability of cohort segmentation*, „Journal of Consumer Marketing", Vol. 24, No. 4.

Rosa G., Perenc J. (2011), *Zachowania nabywców*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Roscoe A.M. Jr, LeClaire A. Jr, Schiffman L.G. (1977), *Theory and management applications of demographics in buyer behavior in consumer and industrial buying behavior*, w: A.D. Woodside, J.M. Sheth, P.D. Bennett (eds.), *Consumer and Industrial Buying Behavior*, Elsevier, New York.

Rudnicki L. (2001), *Zachowania konsumentów na rynku*, PWE, Warszawa 2001.

Schewe C. (1988), *Marketing to our aging populations: responding to psychological changes*, „Journal of Consumer Marketing", Vol. 5, No. 3.

Schiffman L.G., Sherman E. (1991), *Value Orientations of New-Age Elderly: The Coming of an Ageless Market*, „Journal of Business Research", Vol. 22, No. 2.

Szmigin I., Carrigan M. (2001), *Time, Consumption and the Older Consumer: An Interpretive Study of the Cognitively Young*, „Psychology and Marketing", Vol. 18, No. 10.

Szmigin, I., Carrigan, M. (2000), *The Older consumer as Innovator: Does Cognitive Age hold the Key?*, „Journal of Marketing Management", Vol. 16, No. 5.

Tadeusiewicz R. (1998), *Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami*, Akademicka Oficyna Wydawnicza PLJ, Warszawa.

Van de Kaa D.J. (2008), *Demographic transitions*, w: Zeng Yi (ed.), *Demography, Encyclopedia of Life Support Systems (EOLSS)*, Developed under the Auspices of the UNESCO (Oxford, UK: Eolss Publishers), http://www.eolss.net, dostęp: 25.04.2015.

Venkatesh V., Morris M.G., Davis G.B., Davis F.D. (2003), *User acceptance of information technology: towards a unified view*, „MIS Quarterly", Vol. 27, No. 3.

Streszczenie

Artykuł przedstawia propozycję zastosowania sztucznej sieci neuronowej typu GNG do identyfikacji wzorców zachowań nabywczych konsumentów w wieku 60 i więcej w obszarze zakupu produktu technologicznego na przykładzie telefonu komórkowego typu smartfon. Zastosowana metoda jest efektywnym narzędziem analizy wielowymiarowego zbioru danych i identyfikacji wzorców zakupowych. Otrzymane wyniki sugerują występowanie wśród respondentów trzech specyficznych wzorców zachowań nabywczych. Wyniki rozszerzają współczesną wiedzę o zachowania konsumentów w wieku 60 i więcej. Równolegle mogą one mieć zastosowanie w kształtowaniu strategii marketingowych dla tej szczególnej grupy odbiorców.

Słowa kluczowe

konsument 60 i więcej, wzorce zachowań konsumenckich, wzorce nabywcze, samouczące się sieci neuronowe typu GNG, smartfon

The use of self-learning neural networks for analyzing purchasing behaviour and identifying their patterns among consumers at the age of 60 and over (Summary)

The paper presents a proposal to analyze consumer behaviour patterns using self-learning GNG neural networks. The GNG netowork has a great potential for identifing specific patterns of behaviour of people aged 60 and over in the field of purchasing a technological product on the example of a smartphone. The employed method is a highly effective tool for analyzing multidimensional data sets and identifying purchase patterns. The results suggest that three specific consumer patterns of purchase behaviour exist among the respondents. The findings expand the current knowledge concerning purchasing behaviour of consumers at the age of 60 and more. Parallely they may be used in shaping marketing strategies for this specific consumer group.

Keywords

consumer 60 and over, consumer behaviour patterns, purchasing patterns, self-learning neural network GNG, smartphone

Radosław Murkowski*

Płeć a zaawansowanie procesu starzenia się ludności państw europejskich

Wstęp

„Starość" w aspekcie jednostkowym jest to zjawisko o charakterze biologicznym będące ostatnim etapem cyklu życia jednostki, w którym następuje pogorszenie się jej zdrowia, sprawności i atrakcyjności fizycznej. Starzenie się to naturalny i nieuchronny proces, który zależy od wielu czynników, na które człowiek nie zawsze ma wpływ. Natomiast „starość" w aspekcie zbiorowym oznacza taką strukturę ludności według wieku, w której istnieje relatywnie wysoki odsetek osób starszych, podczas gdy z procesem starzenia się populacji mamy do czynienia wtedy, gdy ten odsetek się zwiększa. Proces starzenia się ludności państw Europy powodowany jest z jednej strony przez wzrost długości życia ludzkiego, a z drugiej – jest pogłębiany przez utrzymywanie się dzietności na bardzo niskim poziomie. Ponadto w niektórych krajach europejskich zmiany te są jednocześnie zintensyfikowane poprzez zwiększoną emigrację młodych osób, a w innych łagodzone przez zjawisko odwrotne. Na proces starzenia się ludności mają również wpływ zmiany w zakresie struktury ludności według wieku spowodowane pojawianiem się wyżów lub niżów demograficznych.

Zasadniczym celem artykułu jest ocena zaawansowania procesu starzenia się populacji kobiet i mężczyzn w krajach europejskich w zależności od przyjętej definicji i metody pomiaru tego zjawiska. Płeć jest jednym z najważniejszych czynników charakteryzujących zróżnicowanie procesów ludnościowych w badaniach demograficznych, w tym również procesu starzenia demograficznego. Zróżnicowanie poziomu zaawansowania starzenia się populacji kobiet i mężczyzn wynika z dwóch prawidłowości demograficznych. Po pierwsze, statystycznie rzecz biorąc, mężczyźni rodzą się częściej niż kobiety, co jest uwarunkowane przede wszystkim czynnikami biologicznymi. Po drugie, współcześnie kobiety żyją przeciętnie o kilka lat dłużej niż mężczyźni praktycznie we wszystkich rozwiniętych populacjach, dlatego dominują one liczebnie w starszych grupach wiekowych. Czynniki społeczne różnicujące długość życia kobiet i mężczyzn to m.in. większa skłonność kobiet do dbania o zdrowie czy częstsze zaangażowanie

* Dr, Katedra Statystyki i Demografii, Wydział Ekonomii, Uniwersytet Ekonomiczny w Poznaniu, al. Niepodległości 10, 61-875 Poznań, r.murkowski@ue.poznan.pl

mężczyzn w zachowania niebezpieczne dla zdrowia i życia człowieka. Z drugiej strony niektórzy biolodzy zwracają uwagę m.in. na genetyczne różnice pomiędzy kobietami a mężczyznami, które powodują, że różnica w średniej długości życia kobiet i mężczyzn nigdy nie zostanie całkowicie zniwelowana [Fihel, 2011, s. 22]. Dlatego nawet jeśli mężczyźni zaczęliby unikać zachowań ryzykownych związanych ze swoją płcią, to statystycznie kobiety i tak żyłyby dłużej, różnica bowiem jednego lub dwóch lat jest uwarunkowana genetycznie [Candore i inni, 2006]. Ponadto mężczyźni i kobiety różnią się pod względem zapadalności na poszczególne choroby, bowiem wśród kobiet częściej przyczyną zgonu są choroby układu krążenia, natomiast wśród mężczyzn nowotwory złośliwe [tamże, 2006]. Biorąc pod uwagę te uwarunkowania w zakresie starzenia się obu płci, można postawić pytanie: czy wiek określający początek starości kobiet oraz mężczyzn nie powinien być różny? Ponadto, czy te uwarunkowania starzenia się kobiet i mężczyzn nie powinny znaleźć się w konsekwencji w sposobie pomiaru tego zjawiska z uwzględnieniem płci? Prawidłowa diagnoza procesu starzenia demograficznego kobiet i mężczyzn wymaga zastosowania właściwych narzędzi poznawczych służących do jego pomiaru. Dlatego można postawić hipotezę, że w zależności od przyjętej metody pomiaru starzenia się ludności kobiet i mężczyzn otrzymamy różne wyniki w zakresie poziomu zaawansowania tego zjawiska.

Aktualnie za początek okresu starości uznaje się najczęściej wiek 60 lat lub 65 lat. Dlatego w tradycyjnym ujęciu pomiar zaawansowania procesu starzenia się danej populacji polega najczęściej na obliczeniu proporcji osób w wieku 60 bądź 65 lat i więcej wśród ogółu ludności. Jednakże należy podkreślić, że założenie, że starość rozpoczyna się w momencie osiągnięcia przez daną osobę określonego wieku jest daleko arbitralne i abstrahuje od przemian w zakresie życia w zdrowiu czy umieralności badanej populacji. Dlatego do pomiaru zaawansowania procesu starzenia się populacji kobiet i mężczyzn wykorzystano również koncepcje badawcze, które w sposób płynny traktują granicę pomiędzy wiekiem średnim a wiekiem starczym zarówno w czasie, jak i w przestrzeni. Alternatywny pomiar zaawansowania starzenia się danej populacji polega m.in. na mierzeniu proporcji osób, którym pozostało nie więcej niż np. 15 lat życia. Dlatego wraz ze wzrostem długości życia ludzkiego jak i z poprawą stanu zdrowia zmienia się wiek, który uznajemy za początek starości kobiet oraz mężczyzn.

Badanie zawansowania starzenia się ludności objęło większość państw europejskich wraz z Turcją (leży tylko częściowo w Europie) i Cyprem (pomimo że nie leży w Europie, to jest członkiem Unii Europejskiej). Dane porównano dla wszystkich państw dla 2014 r. Ponadto proces starzenia

się populacji dla wybranych państw zaprezentowano w okresie od 1960 do 2050 r. (prognoza na podstawie szacunków Eurostatu).

1. Miary procesu starzenia się ludności

Pomiar procesu starzenia się populacji dokonywany jest najczęściej za pomocą prostej kalkulacji odsetka osób uznawanych za „stare" wśród ogółu ludności. Za osobę starą uznajemy arbitralnie najczęściej tę, dla której od momentu jej narodzin upłynęła określona i przyjęta z góry liczba lat – zazwyczaj 60 lub 65 lat. Początek okresu starości ustalany jest w tym tradycyjnym ujęciu w sposób chronologiczny (kalendarzowy). Wydaje się jednak, że taki sposób pomiaru tego zjawiska odzwierciedla przede wszystkim zmiany w strukturze ludności według wieku, a niekoniecznie sam proces starzenia demograficznego populacji. Jednakże poza tradycyjnym ujęciem procesu starzenia się ludności, istnieją również alternatywne sposoby jego pomiaru [Abramowska-Kmon, 2011] wynikające z przyjęcia innego początku okresu starości niż tego związanego z wiekiem kalendarzowym.

Badacze W.C. Sanderson, S. Scherbov i W. Lutz [Sanderson, Scherbov 2005, 2007, 2008, 2010; Lutz i inni, 2008] zaproponowali, by obok wieku chronologicznego związanego z liczbą lat, które jednostka już przeżyła, stosować również wiek prospektywny (potencjalny) oparty na liczbie lat, które jednostka może jeszcze przeżyć. Jako początek okresu starości w ujęciu prospektywnym przyjęli wiek, dla którego oczekiwane dalsze trwanie życia wynosi 15 lat zgodnie z aktualnie obserwowanym porządkiem „wymierania" ludności. Próg starości jest w konsekwencji płynny, najczęściej różny dla kobiet i mężczyzn, oraz zmienia się wraz z poprawą jakości życia mieszkańców. Poziom zaawansowania procesu starzenia się poszczególnych populacji można określać w tym ujęciu za pomocą proporcji osób, którym pozostało nie więcej niż 15 lat życia (Prop. 15-). Jednakże pomimo swoich zalet również i w tym podejściu w sposób arbitralny klasyfikuje się poszczególne jednostki za starsze, uznając za nie akurat te, którym pozostało do przeżycia jeszcze średnio dokładnie nie więcej niż 15 lat.

Początek okresu starości można skorelować także z wiekiem biologicznym człowieka, który określa stopień ogólnej sprawności i żywotności człowieka, czyli odzwierciedla aktualny stan zdrowia organizmu, tj. obecność lub brak procesów patologicznych. Przykładowo, siedemdziesięciolatek będący w dobrej formie fizycznej byłby zapewne młodszy biologicznie niż sześćdziesięciolatek cierpiący na poważną chorobę [Balcombe, Sinclair, 2001]. Pomimo że proces starzenia się organizmu jest nieodwracalny, to pojawianie się znamion starości można opóźniać poprzez np. zdrowy i aktywny styl życia. Dlatego prawidłowa klasyfikacja danej osoby jako

„starej" powinna uwzględniać nie tylko fakt dożycia przez nią określonego wieku, ale również obiektywny stan zaawansowania procesu biologicznego jej starzenia się, np. obecność lub nieistnienie pewnych zjawisk związanych ze stanem zdrowia. Dlatego autor proponuje, aby do pomiaru zaawansowania procesu starzenia się poszczególnych populacji wykorzystać obiektywne sumaryczne miary stanu zdrowia populacji i za początek okresu starości uznać wartość oczekiwanego dalszego trwania życia w zdrowiu noworodka. Wskaźnik przeciętnego trwania życia w zdrowiu określa dla osoby o danej płci i wieku średnią liczbę lat życia bez niepełnosprawności (w zdrowiu), a zatem dodaje wymiar jakościowy do oczekiwanej liczby lat życia. Nie zawsze musi być tak, że wzrostowi długości ludzkiego życia towarzyszyć będzie wprost proporcjonalny wzrost długości jego życia w dobrym zdrowiu. Na potrzeby niniejszego badania autor uznał za początek okresu starości oczekiwaną długość życia w zdrowiu noworodka (*Healthy Adjusted Life Expectancy* – HALE) szacowaną przez Światową Organizację Zdrowia (WHO) na podstawie globalnego badania zapadalności na choroby. Wielkość ta określa przeciętną liczbę lat, jaką osoba z danej populacji może przeżyć w zdrowiu na podstawie danych o umieralności oraz danych epidemiologicznych o zapadalności i chorobowości [Wróblewska, 2008; Ojrzyńska, 2013]. Ocena zaawansowania procesu starzenia się poszczególnych państw polegałaby w tym przypadku na mierzeniu proporcji osób, które osiągnęły wiek HALE wśród ogółu ludności (Prop. HALE+). Ponadto Eurostat publikuje wartości przeciętnego dalszego trwania życia w zdrowiu oparte na reprezentatywnych badaniach ankietowych, które bazują na subiektywnym postrzeganiu siebie jako osoby chorej (z pewnym uproszczeniem można takie osoby utożsamiać również jako „stare"). Jednakże jasne jest, że biologiczny aspekt starości warunkuje również sferę psychiczną człowieka i w konsekwencji postrzeganie siebie jako osoby starej, dlatego wykorzystywanie danych Eurostatu i WHO w zakresie długości życia w zdrowiu powinno dawać zbieżne rezultaty poziomu zaawansowania starzenia się ludności.

2. Zaawansowanie procesu starzenia się kobiet i mężczyzn w ujęciu tradycyjnym

W okresie ostatnich kilkudziesięciu lat miał miejsce w krajach wysoko rozwiniętych proces podnoszenia się udziału liczby ludności w najstarszych grupach wiekowych, który był przede wszystkim konsekwencją spadku dzietności i jednocześnie wzrostu szans dożycia okresu starości. Wraz z postępem medycyny w zakresie leczenia schorzeń przewlekłych czy leczenia chorób zakaźnych systematycznie zwiększa się – już praktycznie w całej Europie – wartość przeciętnego dalszego życia, co powoduje, że

wielopokoleniowość staje się kwestią coraz bardziej powszechną. Z drugiej strony w wyniku obniżenia się dzietności proces starzenia się ludności Europy jest często intensyfikowany przez spadek udziału najmłodszych grup wiekowych w strukturze ludności według wieku. Dlatego współczesna Europa stała się jedną z tych części świata, gdzie stopień zaawansowania procesu demograficznego starzenia jest na relatywnie wysokim poziomie.

W tradycyjnym ujęciu za kraje o najniższym stopniu zaawansowania procesu demograficznego starzenia uznamy w 2014 r. państwa Europy Wschodniej (Rosja, Mołdawia, Białoruś), niektóre kraje leżące na Bałkanach (Albania i Macedonia) oraz niektóre państwa Europy Północnej (Irlandia i Islandia) (rys. 1). Jednocześnie osoby będące w wieku 65 lat i więcej stanowiły w 2014 r. w Turcji mniej niż 8% ogółu jej ludności – najmniej wśród wszystkich badanych państw. Z drugiej strony w takich państwach Europy Południowej lub Środkowej jak Włochy, Niemcy czy Grecja ponad 20% ogółu ludności stanowiły osoby w wieku 65 lat i więcej. Ponadto w świetle aktualnych projekcji demograficznych należy oczekiwać, że w przyszłości w niektórych państwach Europy frakcja osób w wieku 65 lat i więcej może w przyszłości przekroczyć poziom 30% ogółu ludności. Jednocześnie intensyfikuje się w Europie proces podwójnego starzenia się ludności, dzięki systematycznemu wzrostowi liczby najstarszych osób pośród wszystkich osób uznawanych za „stare".

W tradycyjnym ujęciu większą szansę na osiągnięcie starości mają kobiety niż mężczyźni, które dominują wśród osób będących w wieku 65 lat więcej we wszystkich badanych państwach Europy. Jednakże sam stopień feminizacji osób starszych jest w Europie wysoce zróżnicowany terytorialnie (tab. 1). Skala dominacji kobiet wśród osób starszych zależy głównie od stanu zdrowia samych mężczyzn, a ten jest konsekwencją m.in. ich skłonności do prowadzenia niezdrowego trybu życia czy podejmowania ryzykownych zachowań. Wyrazem tego zróżnicowania jest m.in. różnica pomiędzy oczekiwanym dalszym trwaniem życia noworodka płci męskiej i żeńskiej. Mężczyźni mieszkający w państwach znajdujących się na najwyższym poziomie rozwoju społeczno-ekonomicznego żyją statystycznie najczęściej tylko o około 4 lata krócej niż kobiety. Z drugiej strony w przypadku państw Europy Wschodniej różnica ta przekracza aktualnie wartość nawet 10 lat, dlatego dla takich państw jak Rosja, Białoruś czy państwa bałtyckie nawet blisko 70% wszystkich osób, które osiągnęły wiek 65 lat stanowiły kobiety (rys. 2). W przypadku państw Europy Północnej kobiety w wieku 65 lat i więcej stanowiły mniej niż 55% ogółu ludności w tym wieku.

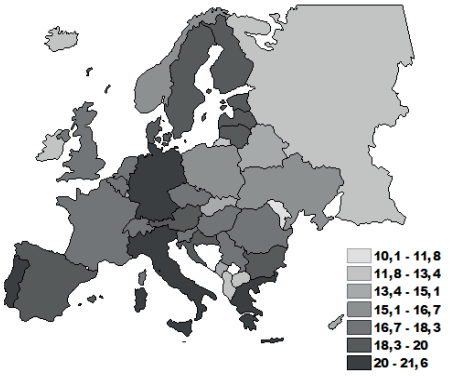
Tablica 1. Odsetek osób w wieku 65 lat i więcej w 2014 r.

Wyszczególnienie	Prop. K 65+ [%]	Prop. M 65+ [%]	Prop. K 65+ w 65+ [%]	Wyszczególnienie	Prop. K 65+ [%]	Prop. M 65+ [%]	Prop. K 65+ w 65+ [%]
Albania	12,8	11,7	51,9	Macedonia	14,0	11,1	55,6
Austria	20,7	16,0	57,5	Malta	20,2	16,2	55,6
Belgia	20,1	15,7	56,9	Mołdawia	12,3	8,0	62,4
Białoruś	17,9	9,7	68,0	Czarnogóra	15,4	11,6	57,6
Bułgaria	22,9	16,5	59,5	Niemcy	23,4	18,4	56,9
Chorwacja	21,8	15,2	60,5	Norwegia	17,7	14,4	54,8
Cypr	15,0	13,4	54,1	Portugalia	22,4	17,5	58,5
Czechy	20,3	14,8	58,7	Rosja	16,7	9,0	68,2
Dania	19,9	16,9	54,5	Rumunia	19,6	13,8	59,8
Estonia	23,2	13,3	66,4	Serbia	20,5	15,9	57,7
Finlandia	22,1	17,2	57,1	Słowacja	16,6	10,8	61,7
Francja	20,3	15,9	57,6	Słowenia	20,8	14,6	59,2
Grecja	22,4	18,8	55,8	Szwajcaria	19,7	15,7	56,2
Hiszpania	20,5	16,0	57,0	Szwecja	21,2	17,9	54,3
Holandia	19,1	16,0	54,8	Turcja	8,9	6,8	56,5
Irlandia	13,6	12,0	53,8	Ukraina	19,2	11,1	66,8
Islandia	14,3	12,5	53,1	Węgry	21,2	13,9	62,7
Litwa	22,9	13,6	66,4	W. Brytania	19,1	16,1	54,9
Luksemburg	15,8	12,4	55,9	Włochy	23,9	19,1	57,1
Łotwa	24,0	13,6	67,5	Polska	18,0	12,1	61,4

Objaśnienia skrótów użytych w tabeli: Prop. K 65 + oznacza odsetek kobiet w wieku 65 lat i więcej wśród ogółu kobiet; Prop. M 65 + oznacza odsetek mężczyzn w wieku 65 lat i więcej wśród ogółu mężczyzn; Prop. K 65 + w 65+ oznacza odsetek kobiet w wieku 65 lat i więcej wśród ogółu osób w wieku 65 lat i więcej.

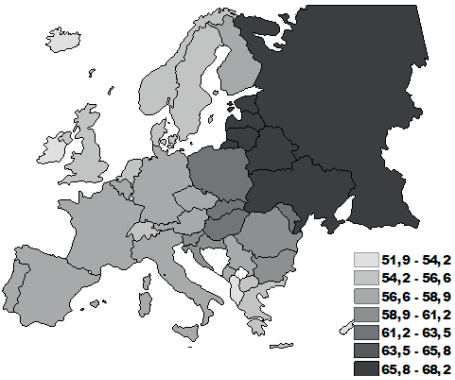
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 1. Odsetek osób w wieku 65 lat i więcej wśród ogółu ludności w 2014 r.



Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 2. Odsetek kobiet w wieku 65 lat i więcej wśród ogółu osób w wieku 65 lat i więcej w 2014 r.



Źródło: Opracowanie własne na podstawie danych Eurostat.

3. Zaawansowanie procesu starzenia się kobiet i mężczyzn w ujęciu alternatywnym

W przypadku pomiaru zaawansowania procesu starzenia się ludności w alternatywnym ujęciu za początek okresu starości przyjmuje się wiek, który jest różny zarówno dla kobiet, jak i mężczyzn czy dla poszczególnych populacji. Próg starości określony może być przez wiek, w którym danej osobie zgodnie z aktualnym porządkiem wymierania pozostało jeszcze nie więcej niż 15 lat życia lub według propozycji autora przez wiek równy oczekiwanemu dalszemu trwaniu życia w zdrowiu noworodka.

Tablica 2. Odsetek osób, dla których oczekiwane dalsze trwanie życia nie przekracza 15 lat w populacji w 2014 r.

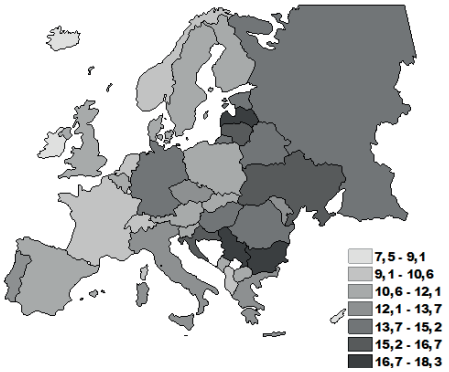
Wyszczególnienie	Prop. 15-kobiet [%]	Prop. 15-mężczyzn [%]	Prop. 15-kobiet w 15- [%]	Wyszczególnienie	Prop. 15-kobiet [%]	Prop. 15-mężczyzn [%]	Prop. 15-kobiet w 15- [%]
Albania	9,4	9,6	49,1	Macedonia	12,2	11,8	50,6
Austria	12,0	11,2	52,9	Malta	10,3	9,8	51,2
Belgia	11,8	11,7	51,0	Mołdawia	11,3	13,2	48,0
Białoruś	14,0	13,9	53,8	Czarnogóra	12,0	11,4	51,8
Bułgaria	17,9	18,5	50,6	Niemcy	15,3	14,1	52,9
Chorwacja	16,2	14,4	54,7	Norwegia	9,7	9,2	50,9
Cypr	8,4	8,7	50,5	Polska	11,2	10,4	53,4
Czechy	12,3	12,8	50,0	Portugalia	13,3	13,0	53,0

Wyszczególnienie	Prop. 15-kobiet [%]	Prop. 15-mężczyzn [%]	Prop. 15-kobiet w 15- [%]	Wyszczególnienie	Prop. 15-kobiet [%]	Prop. 15-mężczyzn [%]	Prop. 15-kobiet w 15- [%]
Dania	11,0	11,7	48,9	Rosja	14,4	15,7	51,5
Estonia	15,6	13,0	57,8	Rumunia	15,1	14,4	52,3
Finlandia	12,2	11,2	52,9	Serbia	17,0	17,1	51,1
Francja	10,3	9,5	53,6	Słowacja	11,3	10,7	52,7
Grecja	14,2	13,1	53,4	Słowenia	12,7	10,7	54,7
Hiszpania	11,1	10,4	52,4	Szwajcaria	10,4	9,6	52,6
Holandia	10,5	10,4	50,8	Szwecja	11,6	11,6	50,1
Irlandia	7,6	8,1	49,0	Turcja	5,6	5,8	48,8
Islandia	7,5	7,7	49,1	Ukraina	16,1	15,0	55,5
Litwa	16,6	14,8	56,7	Węgry	15,4	14,7	53,6
Luksemburg	8,5	8,6	49,5	W. Brytania	10,9	10,7	51,3
Łotwa	17,9	15,8	57,2	Włochy	13,5	12,9	52,6

Objaśnienia skrótów użytych w tabeli: Prop. 15- oznacza odsetek kobiet lub mężczyzn, którym zostało nie więcej niż 15 lat życia wśród ogółu kobiet lub mężczyzn; Prop. 15- kobiet w 15- oznacza odsetek kobiet, którym zostało nie więcej niż 15 lat życia wśród ogółu osób, którym zostało nie więcej niż 15 lat życia
Źródło: Opracowanie własne na podstawie danych Eurostat.

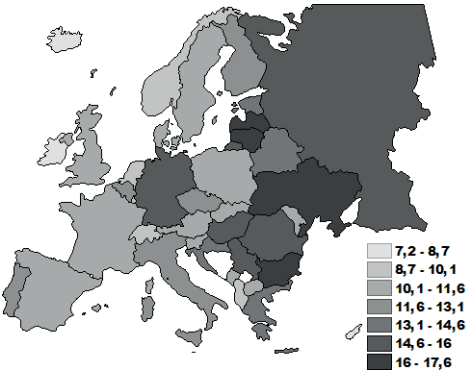
Stopień zaawansowania procesu starzenia się ludności w przypadku alternatywnego ujęcia jest najwyższy dla państw Europy Wschodniej a najniższy dla krajów z Europy Zachodniej – zupełnie odwrotnie jak to było w tradycyjnym ujęciu tego procesu (zob. rys. 3 i 4 oraz por. rys. 1). Bułgaria była państwem, które w 2014 r. charakteryzowało się najwyższym odsetkiem osób starych – aż 18,2% populacji Bułgarii stanowiły osoby, którym pozostało nie więcej niż 15 lat życia i jednocześnie około 17,2% ludności stanowiły osoby, które były w wieku wyższym niż wartość oczekiwanego dalszego trwania życia w zdrowiu noworodka. Wśród państw, które w Europie charakteryzowały się wysokim odsetkiem osób starych, znajdowały się jeszcze takie państwa, jak Serbia, Łotwa, Litwa, Ukraina, Rosja czy Węgry. Natomiast najmniejszy odsetek osób, które uznalibyśmy za „stare", znajdował się w Turcji – około 5,7% populacji miało przed sobą średnio jeszcze nie więcej niż 15 lat życia oraz 7,1% ludności było w wieku wyższym niż wartość HALE. Za państwa z najmniejszym odsetkiem osób starszych uznamy także takie kraje, jak: Islandia, Irlandia, Cypr, Luksemburg czy Norwegia – dla tych państw mniej niż 10% ogółu ludności to osoby starsze.

Rysunek 3. Odsetek osób, dla których oczekiwane dalsze trwanie życia nie przekracza 15 lat w populacji w 2014 r.



Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 4. Odsetek osób w wieku wyższym niż HALE w populacji w 2014 r.



Źródło: Opracowanie własne na podstawie danych Eurostat.

Wyniki pomiaru stopnia zaawansowania procesu starzenia się poszczególnych populacji w Europie w obu ujęciach alternatywnych są do siebie bardzo zbliżone (różnica pomiędzy początkiem okresu starości w obu tych podejściach w zdecydowanej większości badanych populacji nie przekraczała jednego roku) i jednocześnie diametralnie różnią się w stosunku do tradycyjnego ujęcia tego zjawiska. Z drugiej jednak strony można oczekiwać, że różnica pomiędzy oboma ujęciami alternatywnymi może w przyszłości się jednak pogłębiać, gdy wzrostowi średniej długości życia nie będzie towarzyszyć proporcjonalny wzrost jego długości w pełnym zdrowiu.

Tablica 3. Odsetek osób w wieku wyższym niż HALE w populacji w 2014 r.

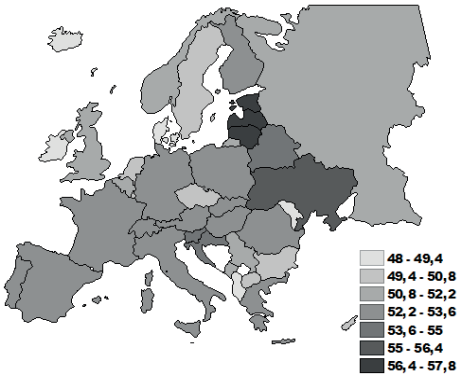
Wyszczególnienie	Prop. HALE+ kobiet [%]	Prop. HALE+ mężczyzn [%]	Prop. HALE+ kobiet w HALE+ [%]
Albania	8,3	10,0	44,9
Austria	11,9	10,8	53,4
Belgia	12,4	11,1	53,6
Białoruś	13,6	14,2	52,4
Bułgaria	16,9	18,2	49,5
Chorwacja	14,4	13,0	54,3
Cypr	8,2	8,7	49,9
Czechy	11,7	12,1	50,0
Dania	11,1	10,7	51,3

Wyszczególnienie	Prop. HALE+ kobiet [%]	Prop. HALE+ mężczyzn [%]	Prop. HALE+ kobiet w HALE+ [%]
Estonia	15,6	12,8	58,2
Finlandia	12,5	11,6	52,7
Francja	11,4	10,1	54,7
Grecja	13,4	13,3	51,6
Hiszpania	11,8	10,6	53,4
Holandia	10,3	9,0	54,0
Irlandia	7,5	7,4	50,8
Islandia	7,5	6,9	52,2
Litwa	17,3	16,8	54,7
Luksemburg	9,4	8,0	54,0
Łotwa	17,4	15,6	57,0
Macedonia	10,0	10,2	49,3
Malta	10,5	8,9	54,4
Mołdawia	10,0	11,1	49,4
Czarnogóra	11,2	10,2	52,9
Niemcy	15,3	13,9	53,3
Norwegia	9,5	8,6	52,3
Polska	11,5	11,2	52,2
Portugalia	13,3	12,7	53,7
Rosja	13,7	15,9	50,1
Rumunia	14,4	15,0	50,1
Serbia	14,8	14,9	51,0
Słowacja	10,6	10,5	51,5
Słowenia	12,4	10,7	54,0
Szwajcaria	10,4	9,0	54,1
Szwecja	11,8	10,5	52,8
Turcja	7,1	7,1	49,8
Ukraina	16,0	16,4	53,2
Węgry	15,3	14,1	54,4
W. Brytania	11,2	10,4	52,6
Włochy	14,3	11,7	56,4

Objaśnienia skrótów użytych w tabeli: Prop. HALE+ oznacza odsetek kobiet lub mężczyzn w wieku wyższym niż HALE wśród ogółu kobiet lub mężczyzn; Prop. HALE+ kobiet w HALE+ oznacza odsetek kobiet w wieku wyższym niż HALE wśród ogółu osób w wieku wyższym niż HALE

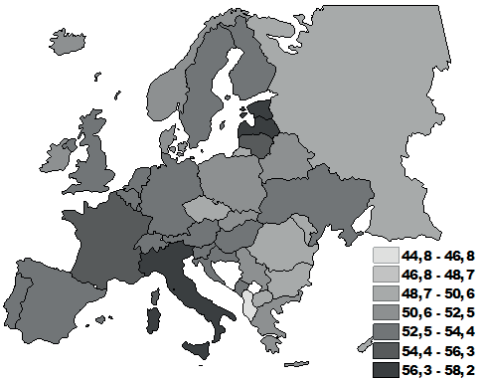
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 5. Odsetek kobiet, dla których oczekiwane dalsze trwanie życia nie przekracza 15 lat wśród ogółu osób, dla których oczekiwane dalsze trwanie życia nie przekracza 15 lat w 2014 r.



Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 6. Odsetek kobiet w wieku wyższym niż HALE wśród ogółu osób w wieku wyższym niż HALE w 2014 r.



Źródło: Opracowanie własne na podstawie danych Eurostat.

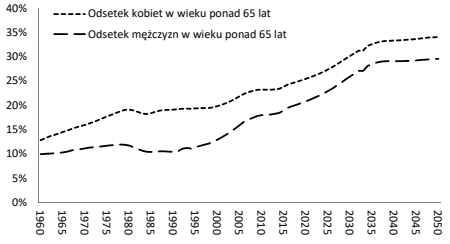
W alternatywnym pomiarze zaawansowania starzenia się ludności stopień feminizacji osób starszych jest znacznie mniejszy niż w przypadku tradycyjnego ujęcia. Frakcja kobiet, którym zostało nie więcej niż 15 lat życia wśród wszystkich kobiet, jest dla większości państw Europy bardzo zbliżona do poziomu notowanego w populacji mężczyzn – dla około 75% wszystkich badanych państw kształtuje się na poziomie nieróżniącym się więcej niż o 1 p.p. (tab. 2). Podobne rezultaty otrzymamy dla drugiego ujęcia alternatywnego, gdzie różnica pomiędzy frakcją kobiet i mężczyzn będących w wieku wyższym niż wartość przeciętnego dalszego trwania życia w zdrowiu tylko dla trzech państw z 40 badanych przekraczała 2 p.p. (tab. 3). Dla porównania w ujęciu tradycyjnym odsetek kobiet w wieku 65 lat i więcej wśród wszystkich kobiet był zawsze wyższy niż dla mężczyzn i tylko dla trzech państw różnica ta wynosiła mniej niż 2 p.p., podczas gdy przykładowo dla Łotwy przekraczała nawet 10 p.p. (por. tab. 1 i por. z tab. 2–3). Ponadto w ujęciu tradycyjnym w populacji osób „starych", tj. np. w wieku 65 lat i więcej, we wszystkich państwach rozwiniętych dominują kobiety, podczas gdy w ujęciu alternatywnym dominacja ta nie jest już tak wyraźna. Kobiety, którym zostało jeszcze nie więcej niż 15 lat życia, stanowiły w połowie badanych państw około 50% (+/–2%) wszystkich osób uznawanych za „stare". Podobnie było dla drugiego ujęcia alternatywnego, gdzie kobiety będące w wieku wyższym niż HALE stanowiły dla połowy badanych państw około 50% (+/–3%) wszystkich osób uznawanych za „stare". Jednocześnie dla niektórych państw w Europie

liczba mężczyzn uznanych za osoby starsze była wyższa niż kobiet (m.in. w Albanii, Mołdawii i Turcji). Z drugiej strony, w szczególności dla państw bałtyckich, dominacja kobiet, które uznamy za starsze, była nadal wysoka – na 100 osób uznanych za starsze 58 osób to były kobiety (zob. rys. 3–4). Reasumując, w alternatywnym pomiarze zaawansowania starzenia się ludności współczynnik feminizacji osób starszych jest znacznie mniejszy niż w tradycyjnym ujęciu.

4. Dynamika procesu starzenia się ludności Europy w ujęciu tradycyjnym i alternatywnym

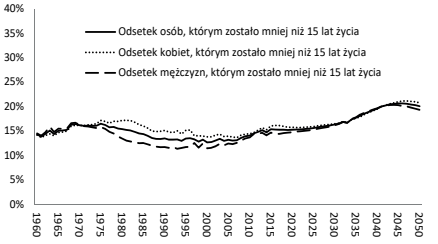
Dotychczasowe rozważania w zakresie pomiaru zaawansowania starzenia się ludności Europy dotyczyły aż 40 państw, ale tylko w ujęciu statycznym, tj. dla 2014 r. Dlatego, by ocenić zaobserwowane prawidłowości, autor porównał również wyniki starzenia się populacji wybranych państw Europy w długim okresie, tj. od 1960 r. aż do 2050 r. Zaprezentowano zmiany poziomu zaawansowania starzenia się populacji kobiet i mężczyzn w ujęciu tradycyjnym oraz alternatywnym dla populacji Niemiec, Słowacji, Portugalii, Norwegii oraz Estonii (kraje te z pewnym uproszczeniem mogą reprezentować Europę Północną, Wschodnią, Południową, Zachodnią i Środkową). Jeden rysunek prezentuje ujęcie tradycyjne przedstawiające odsetek kobiet i mężczyzn będących w wieku 65 lat i więcej wśród ogółu kobiet bądź mężczyzn (rys. 7–15). Natomiast drugi wykres przedstawia ujęcie alternatywne reprezentowane przez odsetek kobiet lub mężczyzn, którym zostało mniej niż 15 lat życia wśród ogółu kobiet lub mężczyzn. W związku z tym, że wyniki przeciętnego dalszego trwania życia w zdrowiu nie są dostępne dla długich okresów czasu, to nie jest możliwe zaprezentowanie wyników drugiego ujęcia alternatywnego w ujęciu dynamicznym.

Rysunek 7. Zmiana odsetka osób w wieku 65 lat i więcej w Niemczech w latach 1960–2050



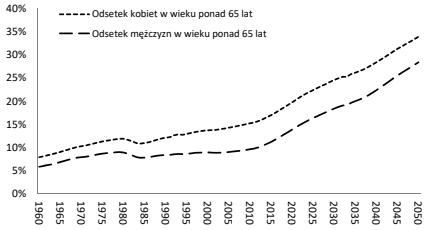
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 8. Zmiana odsetka osób, którym zostało mniej niż 15 lat życia w Niemczech w latach 1960–2050



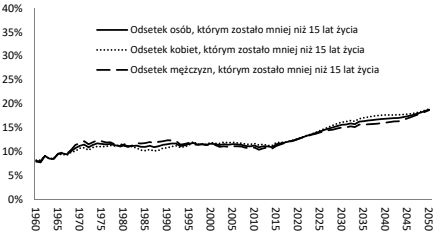
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 9. Zmiana odsetka osób w wieku 65 lat i więcej na Słowacji w latach 1960–2050



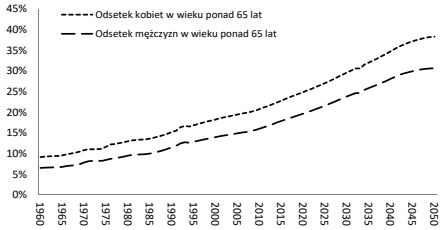
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 10. Zmiana odsetka osób, którym zostało mniej niż 15 lat życia na Słowacji w latach 1960–2050



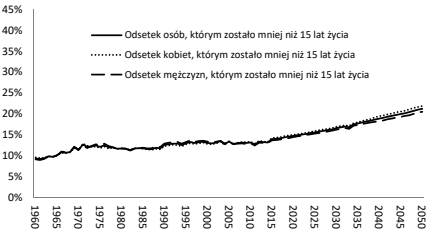
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 11. Zmiana odsetka osób w wieku 65 lat i więcej w Portugalii w latach 1960–2050



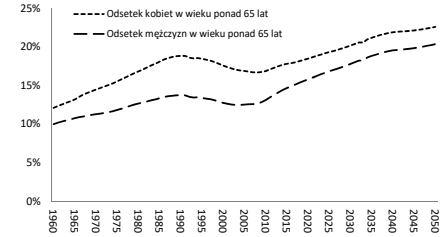
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 12. Zmiana odsetka osób, którym zostało mniej niż 15 lat życia w Portugalii w latach 1960–2050



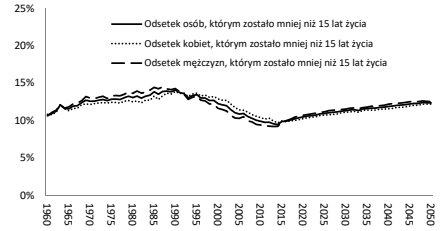
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 13. Zmiana odsetka osób w wieku 65 lat i więcej w Norwegii w latach 1960–2050



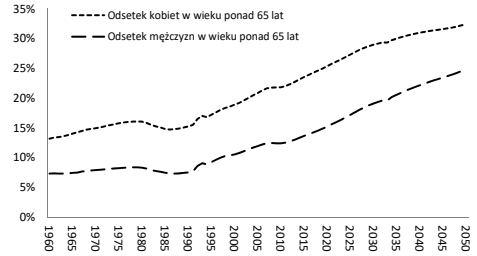
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 14. Zmiana odsetka osób, którym zostało mniej niż 15 lat życia w Norwegii w latach 1960–2050



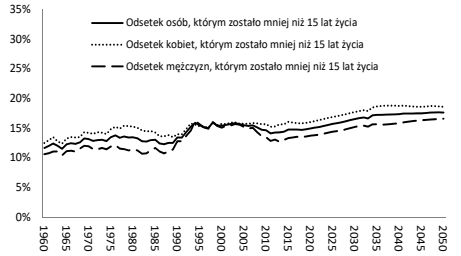
Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 15. Zmiana odsetka osób w wieku 65 lat i więcej w Estonii w latach 1960–2050



Źródło: Opracowanie własne na podstawie danych Eurostat.

Rysunek 16. Zmiana odsetka osób, którym zostało mniej niż 15 lat życia w Estonii w latach 1960–2050



Źródło: Opracowanie własne na podstawie danych Eurostat.

W wyniku porównania ujęcia tradycyjnego i alternatywnego można zwrócić szczególną uwagę na kilka faktów. Po pierwsze, frakcja kobiet będących w wieku 65 lat i więcej wśród ogółu kobiet była zawsze wyraźnie większa niż dla populacji mężczyzn we wszystkich państwach Europy (rys. 7, 9, 11, 13, 15). Prawidłowość ta jest powszechnie znana i wynika z faktu, że kobiety w krajach rozwiniętych żyją statystycznie dłużej niż mężczyźni, w wyniku czego dominują liczebnie w starszych grupach wiekowych. Po drugie, frakcja osób w wieku 65 lat i więcej wśród ogółu ludności zwiększa się w ostatnich latach praktycznie we wszystkich państwach Europy (choć jednak z różnym nasileniem). Po trzecie, wzrost odsetka osób, którym zostało nie więcej niż 15 lat życia wśród ogółu ludności dla badanych państw w Europie nie jest już tak oczywisty i tym bardziej powszechny (rys. 8, 10, 12, 14, 16). W wielu państwach Europy wzrost wieku określającego początek starości zniwelowało wpływ starzenia się ich struktury ludności według wieku. Po czwarte, jak już wspomniano wcześniej, w ujęciu alternatywnym istnieje mniejsze zróżnicowanie w poziomie starzenia się populacji kobiet i mężczyzn – dla zaprezentowanych państw (poza Estonią) frakcja kobiet, którym zostało nie więcej niż 15 lat życia, wśród ogółu kobiet kształtowała się w okresie ostatnich kilkudziesięciu lat na praktycznie takim samym poziomie jak dla mężczyzn (podobne prawidłowości można było zaobserwować również dla większości pozostałych badanych państw Europy).

Zakończenie

Płeć jest jednym z najważniejszych czynników charakteryzujących zróżnicowanie procesów ludnościowych w badaniach demograficznych, w tym również w poziomie starzenia się różnych populacji. Z jednej strony w tradycyjnym ujęciu poziom feminizacji starszych grup wieku

jest w Europie duży, choć stopień tej dominacji jest różny – od bardzo wysokiego w krajach Europy Wschodniej do relatywnie niskiego w państwach Europy Zachodniej. Przyczyną takiego stanu jest inny porządek wymierania mężczyzn związany z ich nadumieralnością w wieku średnim powiązaną z czynnikami o charakterze społecznym i środowiskowym. Dlatego różne uwarunkowania procesu starzenia się kobiet i mężczyzn potwierdzają zasadność, aby początek starości kobiet i mężczyzn był inny – najczęściej wyższy dla kobiet. W związku z tym stopień feminizacji osób starych w alternatywnym ujęciu procesu starzenia populacji jest znacznie mniejszy niż w jego tradycyjnym ujęciu.

Literatura

Abramowska-Kmon A. (2011), *O nowych miarach zaawansowania procesu starzenia się ludności*, „Studia Demograficzne", nr 1 (159).

Balcombe N.R., Sinclair A. (2001), *Ageing: definitions, mechanisms and the magnitude of the Problem*, „Best Practice and Resarch Clinical Gastroenterology", No. 15.

Candore G., Balistreri C.R., Listi F., Grimaldi M.P., Vasto S., Colonna-Romano G., Franceschi C., Lio D., Caselli G., Caruso C. (2006), *Immunogenetics, gender, and longevity*, „Annals of the New York Academy of Sciences", Vol. 1089, No. 1.

Fihel A. (2011), *Płeć a trwanie życia. Analiza demograficzna*, Wydawnictwo Uniwersytetu Warszawskiego, Warszawa.

Lutz W., Sanderson W., Scherbov S. (2008), *The coming acceleration of global population ageing*, „Nature", Vol. 451, No. 7179.

Ojrzyńska A. (2013), *Ocena trwania życia w zdrowiu populacji Polski z wykorzystaniem sumarycznych miar stanu zdrowia*, „Śląski Przegląd Statystyczny", nr 11.

Sanderson W., Scherbov S. (2005), *A new perspective on population aging*, „Demographic Research", Vol. 16.

Sanderson W., Scherbov S. (2005), *Average remaining lifetimes can increase as human populations age*, „Nature", Vol. 435, No. 7043.

Sanderson W., Scherbov S. (2008), *Rethinking Age and Aging*, „Population Bulletin", Vol. 62, No. 4.

Sanderson W., Scherbov S. (2010), *Remeasuring aging*, „Science", Vol. 329, No. 5997.

Wróblewska W. (2008), *Sumaryczne miary stanu zdrowia populacji*, „Studia Demograficzne", nr 153–154.

Streszczenie

Płeć jest jednym z ważniejszych czynników charakteryzujących zróżnicowanie procesów ludnościowych w badaniach demograficznych, w tym również w poziomie starzenia się różnych populacji. Starzenie się ludności oznacza wzrost odsetka osób starych w populacji. Jednakże niezmiernie istotne jest to, jaki wiek uznamy za początek starości. W ramach artykułu zaprezentowane zostało zróżnicowanie zaawansowania procesu starzenia się populacji kobiet i mężczyzn za pomocą tradycyjnych i alternatywnych metod jego pomiaru polegających m.in. na mierzeniu proporcji osób, którym pozostało takie same przeciętne dalsze trwanie

życia. W alternatywnym pomiarze zaawansowania starzenia się ludności stopień feminizacji osób starszych jest znacznie mniejszy niż w przypadku tradycyjnego ujęcia.

Słowa kluczowe
starzenie się ludności, starość, początek okresu starości, alternatywne miary starzenia się populacji

Gender and the advancement of demographic aging of European populations (Summary)

Gender is one of the most important factors characterizing the diversity of population processes in demographic research including the level of aging of various populations. The population aging means that there is an increase in the proportion of old people in the population. However it is very important that we consider age to be the beginning of old age. The article presents the differentiation of the aging process of the population of men and women using traditional and alternative methods of measuring it, including: to measure the proportion of people who have remained the same life expectancy. In the alternative measurement of the advancement of aging, the degree of feminisation of older people is much smaller than in the traditional approach.

Keywords
population aging, senility, beginning of the period of old age, alternative measures of population aging

Beata Bieszk-Stolorz*

Modele regresji w ocenie wpływu płci na formę wyjścia z bezrobocia rejestrowanego

Wstęp

Liczne badania potwierdzają zróżnicowaną sytuację kobiet i mężczyzn na rynku pracy. Według danych BAEL w Polsce w IV kwartale 2016 r. kobiety charakteryzowały się niższym współczynnikiem aktywności zawodowej osób w wieku 15 lat i więcej, niższym wskaźnikiem zatrudnienia i wyższą stopą bezrobocia. Kobiety przeważały również wśród biernych zawodowo (61,8%), a odsetek kobiet wśród ogółu zarejestrowanych w urzędach pracy wyniósł 53,3%. Przemiany dokonujące się na współczesnym rynku pracy powodują, że sytuacja kobiet w ostatnich latach ulega stopniowej zmianie. Jest to wynik upowszechniania elastycznych form zatrudnienia, wzrostu wykształcenia oraz przemian społecznych. Przechodzenie od tradycyjnego do partnerskiego modelu rodziny może wpływać na polepszenie sytuacji kobiet na rynku pracy, gdyż trudniejsza sytuacja kobiet wynika głównie z podwójnej roli, jaką pełnią, czyli aktywnych zawodowo matek lub opiekunek [Kotowska i inni, 2007].

Głównym zadaniem urzędów pracy jest pomoc w znalezieniu zatrudnienia oraz podejmowanie działań związanych z aktywizacją osób bezrobotnych. Podjęcie pracy nie jest jednak jedynym powodem wyrejestrowania. Bezrobotni przechodzą na rentę lub emeryturę, otrzymują zasiłki przedemerytalne, decydują się na kontynuację nauki w systemie dziennym. Jednym z częstszych powodów wyrejestrowań z urzędu jest odmówienie bez uzasadnionej przyczyny przyjęcia propozycji odpowiedniego zatrudnienia. Wszystkie te przyczyny stanowią różnego rodzaju zdarzenia konkurujące. Do ich zbadania można zastosować modele regresji z obszaru analizy trwania. Celem artykułu jest wykorzystanie modelu Lunna-McNeila i wielomianowego modelu logitowego do oceny wpływu płci na szansę i intensywność wychodzenia z bezrobocia rejestrowanego z powodu podjęcia pracy, wykreślenia lub pozostałych przyczyn. Przyczyny te wyodrębniono, korzystając z funkcji skumulowanej częstości *CIF*. W badaniu wykorzystano dane indywidualne 22 078 osób bezrobotnych

* Dr hab., Instytut Ekonometrii i Statystyki, Wydział Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński, ul. Mickiewicza 64, 71–101 Szczecin, beatus@wneiz.pl

zarejestrowanych w 2013 r. w Powiatowym Urzędzie Pracy w Szczecinie (PUP) i obserwowanych do końca 2014 r.

1. Metodyka badania

Wraz z rozwojem demografii wyodrębniła się grupa metod pozwalająca na analizę czasu trwania życia ludzkiego, nazywana analizą trwania (przeżycia). Metody te od wielu lat stosuje się również do badania czasu trwania zjawisk społeczno-ekonomicznych. Przyjmuje się, że czas trwania jednostki w danym stanie, aż do momentu wystąpienia określonego zdarzenia kończącego obserwację, jest zmienną losową T. Ponieważ okres obserwacji jest ograniczony, część jednostek może nie zaznać zdarzenia przed jego końcem. W takich przypadkach czas trwania dla badanej jednostki jest znany jedynie częściowo. Są to obserwacje cenzurowane prawostronnie. Zdarzeniami kończącymi obserwację mogą być: zakończenie działalności firmy, podjęcie zatrudnienia, zaprzestanie spłaty kredytu. Metodami analizy trwania można badać żywotność firm [Markowicz, 2012], aktywność ekonomiczną ludności [Landmesser, 2013], dynamikę ubóstwa w miejskich i wiejskich gospodarstwach domowych [Sączewska-Piotrowska, 2016], ryzyko kredytowe [Wycinka, 2015], czas trwania bezrobocia [Bieszk-Stolorz, Markowicz, 2012a; Bieszk-Stolorz, 2013].

Podstawowym pojęciem analizy trwania jest funkcja trwania, nazywana również funkcją przeżycia, zdefiniowana następująco:

$$S(t) = P(t > T) = 1 - F(t) \tag{1}$$

gdzie: T – czas trwania zjawiska, $F(T)$ – dystrybuanta rozkładu zmiennej losowej T.

Funkcja ta pozwala na wyznaczenie prawdopodobieństwa niezajścia zdarzenia co najmniej do czasu t. Dystrybuanta $F(T)$ wyraża prawdopodobieństwo, że zdarzenie zajdzie najpóźniej do czasu t. W przypadku badania czasu trwania w bezrobociu, jeżeli zdarzeniem jest podjęcie pracy przez osobę zarejestrowaną w urzędzie pracy, to estymator funkcji trwania informuje o tym, jakie jest prawdopodobieństwo pozostania w rejestrze, a estymator dystrybuanty pozwala określić, jakie jest prawdopodobieństwo podjęcia pracy.

W przypadku gdy nie jest znany rozkład czasu trwania analizowanego zjawiska, do estymacji funkcji trwania powszechnie stosuje się estymator Kaplana-Meiera [Kaplan, Meier, 1958]:

$$\hat{S}(t) = \prod_{j:t_j \le t} \left(1 - \frac{d_j}{n_j}\right) \tag{2}$$

gdzie: $t_1 < t_2 < \ldots < t_i < \ldots < t_n$ – momenty zachodzenia zdarzeń, d_j – liczba zdarzeń w momencie t_j, n_j – liczba jednostek zagrożonych do momentu t_j.

Drugą funkcją stosowaną w analizie trwania jest funkcja hazardu, opisująca chwilowy potencjał pojawiającego się zdarzenia w momencie t pod warunkiem przetrwania do czasu t. Jest ona zdefiniowana następująco [Kleinbaum, Klein, 2005]:

$$h(t) = \lim_{\Delta t \to 0} \frac{P\left(t \le T < t + \Delta t \middle| T \ge t\right)}{\Delta t} \tag{3}$$

Funkcja hazardu wyraża intensywność zajścia zdarzenia. Często stosowany jest model proporcjonalnych hazardów Coxa, który można zapisać następująco:

$$h(t, X) = h_0(t)\exp\left(\sum_{j=1}^{p} \beta_j X_j\right) \tag{4}$$

gdzie: $h_0(t)$ – hazard bazowy, X_1, X_2, … ,X_p – zmienne objaśniające.

Interpretacji podlegają wyrażenia $HR = \exp(\beta i)$, nazywane ilorazami hazardu, które pozwalają ocenić intensywność względną wystąpienia danego zdarzenia.

W badaniach za obserwacje cenzurowane przyjmuje się sytuacje, w których badana jednostka znika z pola widzenia lub występuje zdarzenie kończące obserwację, które wyklucza zajście zdarzenia właściwego [Pepe, 1991], czyli zdarzenie konkurujące. W takich sytuacjach ciekawych wyników dostarcza wykorzystanie modeli ryzyk konkurujących [Klein, Bajorunaite, 2004]. W tym przypadku występują dwie zmienne losowe T i C, opisujące odpowiednio czas do zajścia pewnego zdarzenia i czas do cenzurowania. W przypadku K konkurujących rodzajów ryzyka, obserwacji podlegają pary (X, δ), gdzie $X = \min(T, C)$ i $\delta = 0, 1, \ldots, K$. Przyjmuje się, że $\delta = 0$ dla obserwacji cenzurowanej oraz $\delta = 1, \ldots, k$ dla obserwacji kończących się zdarzeniem (jednym z k konkurujących).

Do oceny skumulowanego prawdopodobieństwa wystąpienia zdarzenia z powodu k przed czasem t, przy założeniu, że jednostka jest narażona na wystąpienie któregokolwiek z ryzyk konkurujących k [Bryant, Dignam, 2004] wykorzystuje się funkcję skumulowanej częstości występowania (*Cumulative Incidence Function — CIF*), nazywaną również subdystrybuantą czasu trwania T. Funkcja ta została po raz pierwszy zaproponowana przez Kalbfleischa i Prentice'a (2002) i jest zdefiniowana następująco [Klein, Moeschberger, 2003, s. 52]:

$$CIF_k(t) = P(t \le T, \delta = k) = \int_0^t S(u)h_k(u)du = \int_0^t S(u)dH_k \quad \text{dla } k = 1, 2, 3, \ldots, K \quad (5)$$

gdzie: $H_k(t)$ – funkcja hazardu skumulowanego dla określonego k, $S(t)$ – funkcja przeżycia.

Jeżeli $t_1 < t_2 < \ldots < t_i < \ldots < t_n$ są momentami zachodzenia zdarzeń, to skumulowana funkcja hazardu $H_k(t)$ dla przyczyny k może być określona przez estymator Nelsona-Aalena:

$$\hat{H}_k(t) = \sum_{j:t_{j \le t}} \frac{d_{kj}}{n_j} \quad (6)$$

gdzie: d_{kj} – liczba zdarzeń z powodu zaistnienia przyczyny k, n_j – liczba osób zagrożonych w czasie t_j.

Po połączeniu estymatorów (2) i (6) powstaje estymator funkcji skumulowanej częstości występowania z powodu zaistnienia przyczyny k [Marubini, Valsecchi, 1995]:

$$\hat{CIF}_k(t) = \sum_{j:t_j \le t} \hat{S}(t_{j-1}) \frac{d_{kj}}{n_j} \quad (7)$$

Równość funkcji skumulowanych częstości dla n podgrup weryfikuje test Graya [1988]. Porównuje on średnie ważone hazardów skumulowanej funkcji częstości. Hipoteza zerowa zakłada brak różnic między funkcjami skumulowanych częstości wyznaczonych dla podgrup. Statystyka testowa ma rozkład chi-kwadrat o $n - 1$ stopniach swobody. W przypadku braku występowania zdarzeń konkurujących test Graya sprowadza się do zwykłego testu log-rank.

Do oceny intensywności względnej wystąpienia danego zdarzenia do czasu t można wykorzystać model Lunna-McNeila, którego podstawą jest warstwowy model Coxa z interakcjami. Jeżeli przez $g = 1, 2, \ldots, K$ oznaczy się warstwy będące rodzajami ryzyka, to wersja alternatywna tego modelu ma postać [Kleinbaum, Klein, 2005, s. 4]:

$$h_g(t, X) = h_{0g}(t)\exp\left(\sum_{k=1}^{K}\sum_{j=1}^{p}\delta_{kj}D_k X_j\right) \quad (8)$$

gdzie: $X_1, X_2, \ldots, X_p$ – zmienne objaśniające, $D_1, D_2, \ldots, D_K$ – K zmiennych sztucznych.

Dane do modelu Lunna-McNeila muszą zostać pogrupowane w specjalny sposób. W przypadku występowania K rodzajów ryzyka wyjściowe dane należy zduplikować K razy. Obiekt i o czasie trwania t_i występuje K razy. Każdej obserwacji (również tej zduplikowanej) nadaje się status e_i, który informuje o przynależności obiektu i do określonej grupy ryzyka.

Przyjmuje się, że e_i jest równe 1, jeżeli obserwacja obiektu i zakończyła się zajściem zdarzenia typu k ($k = 1, 2, \ldots, K$) oraz e_i ma wartość 0 w przeciwnym przypadku. Dla obserwacji cenzurowanej e_i jest równe 0 dla każdego $k = 1, 2, \ldots, K$. Sztuczne zmienne dychotomiczne $D_1, D_2, \ldots, D_K$ reprezentują K rodzaje ryzyka: D_k jest równe 1 dla ryzyka typu k i 0 dla pozostałych rodzajów ryzyka. Warstwy $g = 1, 2, \ldots, K$ odpowiadają K zdarzeniom konkurującym. W modelu Lunna-McNeila nie interpretuje się bezpośrednio parametrów δ_{kj}, lecz postać $\exp(\delta_{kj})$. Jeżeli X_j jest zmienną objaśniającą dychotomiczną, to wyrażenie

$$HR_{g=k}\left(X_j = 1 / X_j = 0\right) = \exp\left(\delta_{kj}\right) \tag{9}$$

interpretuje się jako hazard względny (intensywność względną) zajścia zdarzenia typu k.

Innym sposobem oceny ryzyk konkurujących jest zastosowanie w badaniu wielomianowego modelu logitowego, które w analizie trwania traktuje się jako modele hazardu z czasem dyskretnym. Parametry tych modeli pozwalają wyznaczyć szansę względną zajścia zdarzeń konkurujących [Bieszk-Stolorz, Markowicz, 2012b]. Wielomianowy model logitowy ma następującą postać [Kleinbaum, Klein, 2002]:

$$P(Y = 0 \mid X_1, X_2, \ldots, X_m) = \frac{1}{1 + \sum_{j=1}^{k} \exp\left(\alpha_{j0} + \sum_{i=1}^{m} \alpha_{ji} X_i \right)} \tag{10}$$

$$P(Y = j \mid X_1, X_2, \ldots, X_m) = \frac{\exp\left(\alpha_{j0} + \sum_{i=1}^{m} \alpha_{ji} X_i \right)}{1 + \sum_{j=1}^{k} \exp\left(\alpha_{j0} + \sum_{i=1}^{m} \alpha_{ji} X_i \right)} \quad \text{dla } j = 1, 2, \ldots, k \tag{11}$$

gdzie: j – numer wariantu zmiennej objaśnianej; $j = 0, 1, \ldots, k$, i – numer zmiennej objaśniającej; $i = 1, 2, \ldots, m$.

Jeżeli X_j jest zmienną objaśniającą dychotomiczną, to wyrażenie:

$$OR_{g=k}\left(X_j = 1 / X_j = 0\right) = \exp\left(\alpha_{kj}\right) \tag{12}$$

interpretuje się jako szansę względną zajścia zdarzenia typu k.

2. Dane wykorzystane w badaniu

Stosowanie modeli z obszaru analizy trwania wymaga informacji o każdej jednostce wchodzącej w skład badanej kohorty. W prezentowanej w artykule analizie wykorzystano anonimowe dane indywidualne 22 078 (w tym 9770 kobiet) osób bezrobotnych pozyskane z Powiatowego Urzędu

Pracy w Szczecinie. Kohortę stanowiły osoby zarejestrowane w 2013 r. i obserwowane do końca 2014 r. Zdarzeniem kończącym obserwację każdej jednostki był moment wyrejestrowania z urzędu. Badaniu podlegał czas od momentu zarejestrowania do momentu wyrejestrowania z określonego powodu. Dzięki systemowi informatycznemu Syriusz w urzędach pracy gromadzone są obszerne dane na temat zarejestrowanych bezrobotnych. Wśród kilkudziesięciu powodów wyrejestrowania znajdują się związane z podjęciem zatrudnienia, przejściem na rentę lub emeryturę, kontynuacją nauki, wyjazdem za granicę. Powody te pogrupowano. Natomiast 1856 (w tym 913 kobiet) obserwacji nie zakończyła się zdarzeniem, czyli wyrejestrowaniem do końca 2014 r. Stanowią one obserwacje cenzurowane prawostronnie. Liczebność wszystkich grup podana jest w tablicy 1.

Tablica 1. Rodzaje zdarzeń kończących obserwację

Powody wyrejestrowania			
Powody główne	Liczba (kobiety)	Powody cząstkowe	Liczba (kobiety)
Praca	9633 (4809)	Praca niesubsydiowana	7807 (4003)
		Praca subsydiowana	929 (424)
		Działalność gospodarcza	897 (382)
Wykreślenie	8965 (3264)	Wykreślenie	8965 (3264)
Pozostałe	1624 (784)	Renta/zasiłek/ emerytura	685 (380)
		Wyjazd za granicę	445 (201)
		Pozostałe	494 (203)

Źródło: Opracowanie własne.

Na każdy z trzech głównych powodów wyrejestrowania składały się powody cząstkowe, które zawarto w siedmiu grupach (tab. 1). Podjęcie pracy (Praca) składa się z trzech głównych podgrup: podjęcie pracy lub innego zatrudnienia niesubsydiowanego, podjęcie pracy subsydiowanej przez urząd, działalność gospodarcza. Wykreślenie jest grupą powodów, w których zarejestrowana osoba bezrobotna nie wykazała się chęcią współpracy z urzędem i z powodu niestawiennictwa w urzędzie w określonym terminie lub z powodu odmowy przyjęcia propozycji pracy została wykreślona z rejestru. Pozostałe powody wyrejestrowania (Pozostałe) są mniej liczne i z nich z osobna miał marginalny wpływ na prawdopodobieństwo wyrejestrowania z urzędu. Dlatego postanowiono rozpatrywać je razem.

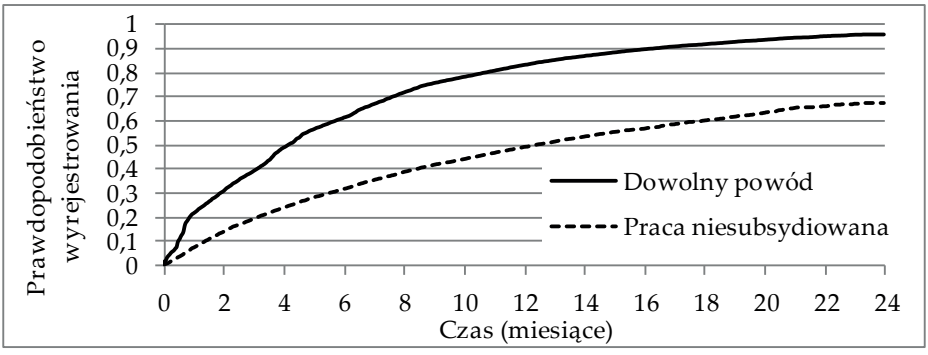
3. Analiza wpływu płci na formę wyjścia z bezrobocia

Badania przeprowadzono w trzech etapach. Pierwszy z nich – badanie wstępne – polegał na wyznaczeniu estymatorów skumulowanej częstości

występowania w celu wyodrębnienia grup form wyjścia z bezrobocia rejestrowanego.

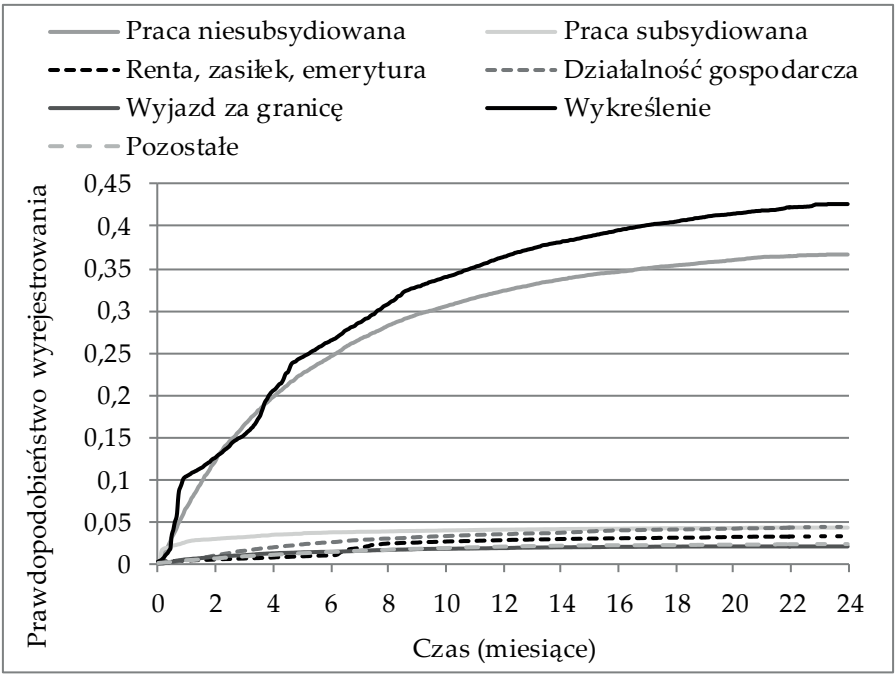
Na rysunku 1 przedstawiono estymatory skumulowanej częstości dla dowolnego powodu wyjścia z bezrobocia i dla wyjścia do pracy niesubsydiowanej. Informują one o prawdopodobieństwie wyrejestrowania z urzędu przed upływem czasu t. W przypadku dowolnego powodu wyrejestrowania obserwacją cenzurowaną jest ta, która nie zakończyła się przed upływem 2014 r. Dla pracy niesubsydiowanej oprócz poprzedniego rodzaju cenzurowania występuje również drugi rodzaj: wyrejestrowanie z powodu innego niż podjęcie pracy. Przykładowo, prawdopodobieństwo wyrejestrowania z dowolnego powodu przed upływem 6 miesięcy od zarejestrowania było równe 0,6, a prawdopodobieństwo podjęcia pracy było równe 0,3. Wynika z tego, że na prawdopodobieństwo wyrejestrowania z urzędu duży wpływ mogły mieć inne przyczyny niż podjęcie pracy. Aby zbadać ten wpływ (ryzyko konkurujące), wyznaczono funkcje CIF_k ($k = 1, 2, \ldots, 7$) dla siedmiu przyczyn wyrejestrowania z urzędu pracy (rys. 2). Podjęcie pracy niesubsydiowanej i wykreślenie osiągają wartości zdecydowanie większe niż pozostałe przyczyny. Pozostałe: praca niesubsydiowana, podjęcie działalności gospodarczej, przejście na rentę, zasiłek i emeryturę, wyjazd za granicę oraz pozostałe nie przekroczyły wartości 0,05 i miały wpływ marginalny na ryzyko całkowite. Analiza wartości tych funkcji wskazuje, że po 4 miesiącu od zarejestrowania prawdopodobieństwo wyrejestrowania bezrobotnego przez urząd z powodu wykreślenia było większe od prawdopodobieństwa podjęcia pracy niesubsydiowanej.

Rysunek 1. Estymatory *CIF* dla dowolnego powodu wyjścia z bezrobocia i podjęcia pracy niesubsydiowanej



Źródło: Opracowanie własne.

Rysunek 2. Estymatory skumulowanej funkcji częstości wyrejestrowań z różnych powodów



Źródło: Opracowanie własne.

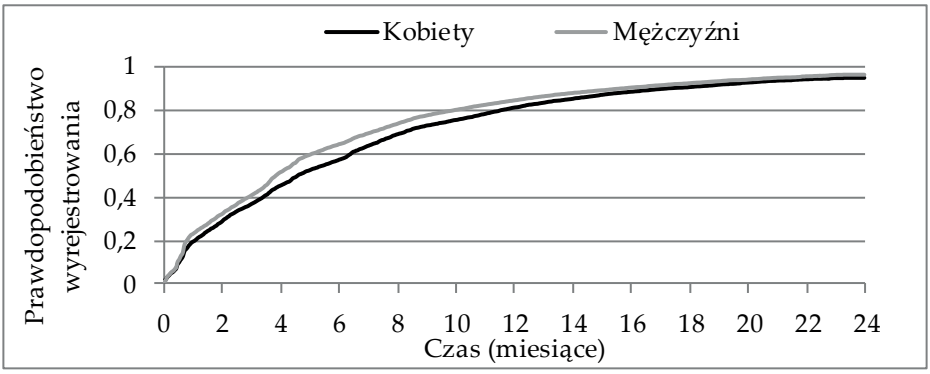
Dla osoby bezrobotnej ważne jest podjęcie ogólnie pojętej pracy. Dlatego drugi etap badania polegał na pogrupowaniu siedmiu przyczyn wyrejestrowania w trzy grupy form wyjścia z bezrobocia i wyznaczeniu funkcji CIF_k dla $k = 1, 2, 3$ dla bezrobotnych kobiet i mężczyzn. Zdarzenia polegające na podjęciu pracy subsydiowanej i niesubsydiowanej oraz podjęcie działalności gospodarczej zawarto w jednej grupie nazwanej Praca. Niezmieniona została grupa Wykreślenie. Wyjazd za granicę, pobieranie renty, zasiłku, emerytury dołączono do grupy Pozostałe. Różnice w przebiegu krzywych dla obu płci zbadano korzystając z testu Graya (tab. 2). Najpierw wyznaczono estymatory skumulowanej częstości dla dowolnego powodu wyrejestrowania, które ze względu na brak zdarzeń konkurujących są równe dystrybuancie rozkładu czasu trwania w bezrobociu (rys. 3).

Tablica 2. Wynik testu Graya dla grup bezrobotnych według płci

Przyczyna wyrejestrowania	Test Graya (Chi-kwadrat)	*p*
Dowolny powód	77,644	0,000
Praca	190,786	0,000
Wykreślenie	413,120	0,000
Pozostałe	10,932	0,001

Źródło: Opracowanie własne.

Rysunek 3. Estymatory skumulowanej funkcji częstości wyrejestrowań z dowolnego powodu dla kobiet i mężczyzn



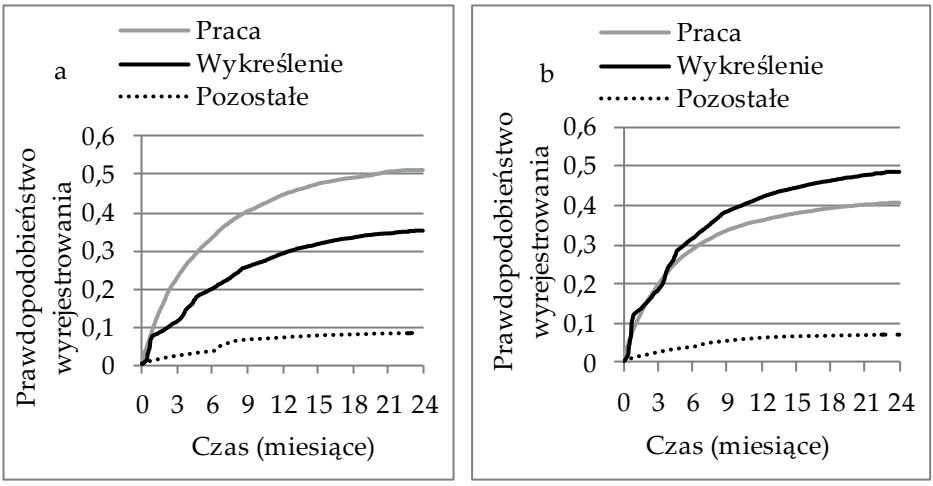
Źródło: Opracowanie własne.

Analizując wykresy obu estymatorów, widać, że prawdopodobieństwo wyrejestrowania bezrobotnych mężczyzn z dowolnego powodu jest nieco większe niż w przypadku kobiet. Istotność różnic między krzywymi została potwierdzona testem Graya ($\chi^2 = 77{,}644$; $p = 0{,}000$). Prawdopodobieństwo pozostania w rejestrze urzędu po 12 miesiącach od momentu zarejestrowania było równe 0,16 w przypadku mężczyzn i 0,19 dla kobiet. Jest to jednocześnie prawdopodobieństwo przejścia w stan długotrwałego bezrobocia.

Do bardziej szczegółowych wniosków na temat przyczyn wyrejestrowania prowadzi analiza ryzyk konkurujących uwzględniająca trzy rodzaje powodów wyrejestrowania (zdarzeń konkurujących): podjęcie pracy (w sensie ogólnym), wykreślenie i pozostałe. Test Graya wskazał na różnice w przebiegu krzywych skumulowanej częstości występowania dla grup wyodrębnionych ze względu na płeć (tab. 2). Podjęcie pracy było najczęstszym powodem wyrejestrowania dla kobiet w całym okresie trwania w bezrobociu (rys. 4a). Na drugim miejscu było wykreślenie. W przypadku mężczyzn najbardziej prawdopodobnym powodem wyrejestrowania było wykreślenie (począwszy od 4 miesiąca od zarejestrowania), a następnie podjęcie pracy (rys. 4b). Pozostałe przyczyny były na trzecim miejscu

zarówno dla kobiet, jak i mężczyzn i miały znaczenie marginalne. Prawdopodobieństwo wyrejestrowania kobiet do 24 miesiąca od zarejestrowania z powodu podjęcia pracy było równe 0,51 (mężczyźni 0,41), z powodu wykreślenia 0,35 (mężczyźni 0,49), z powodu pozostałych przyczyn 0,084 (mężczyźni 0,071).

Rysunek 4. Estymatory skumulowanej funkcji częstości (*CIFk*) dla różnych powodów wyrejestrowania kobiet (a) i mężczyzn (b)



Źródło: Opracowanie własne.

Analizując przebieg krzywych, zauważalna jest regularna krzywizna estymatorów powodu Praca. Estymatory prawdopodobieństwa wykreślenia i pozostałych powodów nie posiadają takiej własności. Dla grupy Pozostałe, zarówno w przypadku kobiet, jak i mężczyzn, w siódmym miesiącu od zarejestrowania wystąpił niewielki skok wartości. Analiza danych wykazała, że był on spowodowany zwiększoną liczbą wyrejestrowań w tym okresie z powodu przyznania zarejestrowanej osobie bezrobotnej prawa do pobierania świadczenia/zasiłku przedemerytalnego. Dla zdarzenia Wykreślenie znaczny skok w pierwszym miesiącu był związany ze zwiększoną liczbą wyrejestrowań z powodu niestawiennictwa osoby bezrobotnej (kobiet i mężczyzn) w PUP w wyznaczonym terminie. Gdyby wszystkie obserwowane osoby zostały wyrejestrowane do końca 2014 r., to suma estymatorów CIF_k dla wszystkich rodzajów ryzyka w 24 miesiącu byłaby równa 1. Ponieważ w badaniu wystąpiły obserwacje cenzurowane, to suma ta jest mniejsza od 1. Niezerowa różnica, która powstała, pozwala określić prawdopodobieństwo pozostania w rejestrze bezrobotnych po 24 miesiącach od momentu zarejestrowania. W przypadku kobiet było ono równe 0,05, a dla mężczyzn 0,04.

W trzeciej części badania wykorzystano modele regresji z obszaru analizy trwania. Model Lunna-McNeila (wersja alternatywna) posłużył do zbadania wpływu płci na intensywność różnych form wychodzenia z bezrobocia. Dychotomiczna zmienna objaśniająca X wyraża płeć osoby bezrobotnej, przy czym przyjmuje wartość 1 dla kobiet i 0 dla mężczyzn (grupa referencyjna). Sztuczne zmienne dychotomiczne D_k ze wzoru (8) określają w tym przypadku formę ryzyka konkurującego i przyjmują wartość 1 dla ryzyka numer g i 0 w przeciwnym przypadku. W badaniu g przyjmuje następujące wartości: 1 dla podjęcia pracy, 2 w przypadku wykreślenia i 3 dla pozostałych powodów wyrejestrowania. Poziom istotności parametrów przyjęto na poziomie 0,01. Wyniki estymacji parametrów modelu Lunna-McNeila przedstawiono w tablicy 3.

Tablica 3. Oceny parametrów modelu Lunna-McNeila

Parametr	Ocena	Błąd standardowy	Statystyka Walda	p	HR
d_1	0,1270	0,0204	38,8026	0,0000	1,14
d_2	–0,4389	0,0220	399,2858	0,0000	0,64
d_3	0,0390	0,0500	0,6168	0,4322	1,04
$\chi^2 = 451{,}836$, $p = 0{,}0000$					

Źródło: Opracowanie własne.

Do zbadania szansy względnej wyjścia z bezrobocia kobiet w stosunku dla mężczyzn wykorzystano wielomianowy model logitowy, w którym uwzględniono cztery stany kończące obserwację. Zmienną objaśniającą X jest, tak jak w poprzednim modelu, płeć, a zmienna objaśniana Y przyjęła następujące wartości: 0 dla pozostania w rejestrze, 1 dla podjęcia pracy, 2 w przypadku wykreślenia i 3 dla pozostałych powodów wyrejestrowania. Wyniki estymacji parametrów wielomianowego modelu logitowego przedstawiono w tablicy 4.

Do interpretacji wyników estymacji parametrów modeli: hazardu i logitowych wykorzystano odpowiednio ilorazy hazardu i ilorazy szans (tab. 3–4). Z przeprowadzonej analizy wynika, że kobiety o prawie 14% intensywniej niż mężczyźni podejmowały pracę i o nieco ponad 35% mniej intensywniej były wykreślane z rejestru. Intensywności wyrejestrowania z pozostałych powodów dla kobiet i mężczyzn były zbliżone (brak istotności parametru d_3). Szansa (ryzyko) wykreślenia z rejestru dla kobiet była o 61% niższa niż dla mężczyzn, wyrejestrowania z powodu podjęcia pracy i z pozostałych powodów dla kobiet i mężczyzn były zbliżone (brak istotności parametrów α_1 *i* α_3).

Tablica 4. Wyniki estymacji parametrów wielomianowego modelu logitowego

Parametr	Ocena	Błąd standardowy	Statystyka Walda	*p*	*OR*
α_{10}	1,6323	0,0356	2101,672	0,0000	–
α_1	0,0292	0,0507	0,332	0,5645	1,03
α_{20}	1,7993	0,0352	2619,721	0,0000	–
α_2	–0,5254	0,0514	104,644	0,0000	0,59
α_{30}	–0,1157	0,0474	5,943	0,0148	–
α_3	–0,0367	0,0680	0,291	0,5897	0,96

Źródło: Opracowanie własne.

Zakończenie

Zastosowane w badaniu metody analizy trwania: funkcja skumulowanej częstości, model Lunna-McNeila oraz wielomianowy model logitowy pozwoliły na ocenę wpływu płci na prawdopodobieństwo, intensywność względną i szansę względną wyrejestrowania z urzędu pracy z określonego powodu. Estymator skumulowanej funkcji częstości wskazał, że podjęcie pracy było najbardziej prawdopodobnym powodem wyrejestrowania w przypadku kobiet, a w grupie mężczyzn – wykreślenie. Wynika z tego, że chęć podjęcia zatrudnienia nie była jedyną przyczyną zarejestrowana się w urzędzie. Pozostałe powody miały znaczenie marginalne. Parametry modelu Lunna-McNeila wskazały, że kobiety intensywniej niż mężczyźni podejmowały pracę, mężczyźni zaś intensywniej byli wykreślani z urzędu. Płeć osoby bezrobotnej nie wpływała na intensywność wyrejestrowania z pozostałych powodów. Na podstawie parametrów wielomianowego modelu logitowego wykazano, że szansa (ryzyko) wykreślenia z rejestru była niższa dla kobiet. Płeć osoby bezrobotnej nie wpływała na szanse wyrejestrowania z powodu podjęcia pracy i z pozostałych powodów.

Literatura

Bieszk-Stolorz B. (2013), *Analiza historii zdarzeń w badaniu bezrobocia*, Volumina.pl Daniel Krzanowski, Szczecin.

Bieszk-Stolorz B., Markowicz I. (2012a), *Modele regresji Coxa w analizie bezrobocia*, CeDeWu, Warszawa.

Bieszk-Stolorz B., Markowicz I. (2012b), *Wykorzystanie wielomianowego modelu logitowego do oceny szansy podjęcia pracy przez bezrobotnych*, w: K. Jajuga, M. Walesiak (red.), „Taksonomia", nr 19: *Klasyfikacja i analiza danych – teoria i zastosowania*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", nr 242.

Bryant J. , Dignam J.J. (2004), *Semiparametric models for cumulative incidence functions*, „Biometrics", Vol. 60, No. 1.

Gray R.J. (1988), *A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk*, „The Annals of Statistics", Vol. 16, No. 3.

Kalbfleisch J.D., Prentice R.L. (2002), *The Statistical Analysis of Failure Time Data, Second Edition*, John Wiley & Sons, Inc., Hoboken, New Jersey.

Kaplan E.L., Meier P. (1958), *Nonparametric Estimation from Incomplete Observations*, „Journal of American Statistical Association", Vol. 53, No. 282.

Klein J.P., Bajorunaite R. (2004), *Inference for Competing Risks*, w: N. Balakrishnan, C.R. Rao (eds.), *Handbook of Statistics: Advances in Survival Analysis,* Elsevier, Amsterdam.

Klein J.P., Moeschberger M.L. (2003), *Survival Analysis: Techniques for Censored and Truncated Data. Second Edition*, Springer-Verlag, New York.

Kleinbaum D., Klein M. (2002), *Logistic Regression. A Self-Learning Text*, Springer-Verlag, New York.

Kleinbaum D., Klein M. (2005), *Survival Analysis. A Self-Learning Text*, Springer-Verlag, New York.

Kotowska I.E., Sztanderska U., Wóycicka I. (red.) (2007), *Aktywność zawodowa i edukacyjna a obowiązki rodzinne w Polsce w świetle badań empirycznych*, Wydawnictwo Naukowe SCHOLAR, Warszawa.

Landmesser J. (2013), *Wykorzystanie metod analizy czasu trwania do badania aktywności ekonomicznej ludności w Polsce*, Wydawnictwo SGGW, Warszawa.

Markowicz I. (2012), *Statystyczna analiza żywotności firm*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Marubini E., Valsecchi M. (1995), *Analysing Survival Data from Clinical Trials and Observational Studies*, JohnWiley & Sons, New York.

Pepe M.S. (1991), *Inference for Events With Dependent Risks in Multiple Endpoint Studies*, „Journal of the American Statistical Association", Vol. 86, No. 415.

Sączewska-Piotrowska A. (2016), *Dynamika ubóstwa w miejskich i wiejskich gospodarstwach domowych*, „Wiadomości Statystyczne", nr 7.

Wycinka E. (2015), *Modelowanie czasu do zaprzestania spłat rat kredytu lub wcześniejszej spłaty kredytu jako zdarzeń konkurujących*, „Problemy Zarządzania", nr 3(55), t. 2.

Streszczenie

Podjęcie pracy nie jest jedynym powodem wyrejestrowania z urzędu pracy. Bezrobotni przechodzą na rentę lub emeryturę, otrzymują zasiłki przedemerytalne, decydują się na kontynuację nauki w systemie dziennym. Jednym z częstszych powodów jest odmówienie bez uzasadnionej przyczyny przyjęcia propozycji odpowiedniego zatrudnienia. Wszystkie te przyczyny stanowią różnego rodzaju zdarzenia konkurujące. Do ich zbadania można zastosować modele regresji z obszaru analizy trwania. Celem artykułu jest wykorzystanie modelu Lunna-McNeila i wielomianowego modelu logitowego do oceny wpływu płci na szansę i intensywność wychodzenia z bezrobocia rejestrowanego z powodu podjęcia pracy, wykreślenia lub pozostałych przyczyn. Przyczyny te wyodrębniono, korzystając z funkcji skumulowanej częstości *CIF*. W celu zbadania różnic we wpływie płci na prawdopodobieństwo wyrejestrowania z powodu określonej przyczyny zastosowano test Graya. W badaniu skorzystano z danych indywidualnych 22 078 osób bezrobotnych zarejestrowanych w 2013 r. w Powiatowym Urzędzie Pracy w Szczecinie i obserwowanych do końca 2014 r. W analizowanym okresie kobiety

intensywniej niż mężczyźni podejmowały pracę oraz mniej intensywniej były wykreślane z rejestru. Intensywności wyrejestrowania z pozostałych powodów dla kobiet i mężczyzn nie różniły się od siebie w sposób istotny statystycznie. Szanse mężczyzn na wykreślenie były większe niż dla kobiet. W przypadku pozostałych przyczyn różnice nie były istotne statystycznie.

Słowa kluczowe
ryzyka konkurujące, skumulowana funkcja częstości, wielomianowy model logitowy, model Lunna-McNeila

Regression models in assessment of the influence of gender on the form of registered unemployment exit (Summary)

Accepting a job is not the only reason of deregistration from the labour office. The unemployed persons become retired, can apply for invalidity pension, receive early retirement benefits or start full time studies. One of the most common causes of de-registering is the unemployed person's unjustified refusal to accept a job offer. The above causes are regarded as competing events of various kinds. They can be analysed by means of the regression models from the domain of survival analysis. The goal of the article is application of the Lunn-McNeal model and multinomial logit model for the assessment of the influence of gender on the odds and intensity of registered unemployment exit because of accepting a job, refusal or other causes. These causes were extracted by means of the cumulative incidence function. The Gray test was used in order to analyse the differences in the influence of gender on the probability of deregistration because of specific reason. The analysis was performed on the basis of individual data of 22,078 unemployed persons, registered in the Poviat Labour Office in Szczecin in 2013 and observed until the end of the year 2014. In the analysed period women accepted jobs more intensively than men and were deregistered less intensively. The intensities of deregistration because of other causes for women and men were not significantly different. The odds of refusal for men were higher than for women. In case of the other causes the differences were not statistically significant.

Keywords
competing risks, cumulative incidence function, multinomial logit model, Lunn-McNeil model

Rumiana Górska*
Piotr Staszkiewicz**

Zastosowanie algorytmu lasów losowych do prognozowania modyfikacji opinii biegłego rewidenta

Wstęp

Biegły rewident, badając sprawozdanie finansowe, może wydać opinię pozytywną (bez zastrzeżeń) lub modyfikowaną. Określa się, iż raport biegłego rewidenta jest modyfikowany, jeżeli opinia biegłego rewidenta zawiera zastrzeżenia lub ocena sprawozdania jest negatywna albo zakres badania jest ograniczony.

Celem artykułu jest sprawdzenie, czy decyzję biegłego rewidenta wydania opinii pozytywnej lub zmodyfikowanej można prognozować, a także dodatkowo określenie czynników wpływających na tę decyzję przy zastosowaniu algorytmu lasów losowych. Algorytm lasów losowych jest uważany za jedną z najdokładniejszych metod klasyfikacji, polegającą na łączeniu wielu drzew klasyfikacyjnych. Badanie przeprowadzono na próbie 774 opinii rocznych sprawozdań finansowych polskich przedsiębiorstw z lat 2010–2013, spośród których 286 były modyfikowane. Uzyskane wyniki klasyfikacji i prognozy opinii biegłego rewidenta porównano z wynikami modelu logitowego.

Artykuł składa się z trzech części. W pierwszej zawarty jest opis problemu badawczego i przegląd literatury. Druga zawiera opis danych oraz metodologię. W trzeciej części przedstawiono wyniki empiryczne i wnioski.

1. Problem badawczy i przegląd literatury

Jedno z pierwszych badań czynników wpływających na opinię biegłego rewidenta było to przeprowadzone w 1974 r. przez E.I. Altmana i T. McGougha [1974], którzy zestawili model bankructwa przedsiębiorstwa z typem opinii wydanych dla podmiotów ogłaszających następnie bankructwo. Następne badania dotyczyły wykorzystania różnych modeli predykcyjnych do przewidywania typu opinii [Dopuch i inni, 1987; Mutchler, 1985; Koh, Killough, 1990]. W innym badaniu [Kennedy, Shaw, 1991] autorzy badali możliwości predykcji wyłącznie na podstawie opinii z poprzedniego

* Dr, Instytut Ekonometrii, Szkoła Główna Handlowa w Warszawie, al. Niepodległości 162, 02-554 Warszawa, e-mail: rgorska@sgh.waw.pl
** Dr, Instytut Finansów Korporacji i Inwestycji, Szkoła Główna Handlowa w Warszawie, al. Niepodległości 162, 02-554 Warszawa, piotr.staszkiewicz@sgh.waw.pl

roku. Zaś K. Keasey i R. Watson byli prekursorami wykorzystania opinii biegłego rewidenta jako zmiennej niezależnej w modelu predykcji bankructwa [Keasey, Watson, 1987]. Podobnie E. Altman i inni [2016].

W Polsce prognozowaniem opinii biegłego rewidenta przy użyciu modeli logitowych zajmowali się M. Gruszczyński i B. Pajdo [2003], którzy w swojej pracy odnieśli się do dwustopniowego modelu Ch. Pryor i J. Terza, a także B. Gostomczyk [2013]. M. Iwanicz-Drozdowska przeprowadzała zaś badania w zespole wspomnianego wyżej E. Altmana [Altman i inni, 2016]. R. Górska i P. Staszkiewicz [2018] przedstawili dyskusję relacji między wynagrodzeniem rewizyjnym i nierewizyjnym firmy autorskiej a jakością badania przybliżaną typem opinii biegłego rewidenta.

We wspomnianych badaniach głównym przedmiotem zainteresowania były możliwości wykorzystania opinii do prognozowania zagrożenia upadłości lub bankructwa lub ustalenie zmiennych wpływających na opinię. Wszak we wskazanych podejściach badacze nie określali, na ile wybór metody prognozowania wpływa na jej dokładność. Mając na uwadze zidentyfikowaną lukę, w prezentowanym badaniu przeprowadzono także porównanie dokładności dwóch wspomnianych metod prognozowania.

W badaniu uwzględniono trzy grupy zmiennych jako determinanty opinii biegłego rewidenta: zmienne finansowe, charakterystyki podmiotu badanego i charakterystyki firmy audytorskiej.

Podstawowym zakresem wypowiedzi biegłego rewidenta jest rzetelność obrazu finansowego przedsiębiorstwa przedstawionego w sprawozdaniu finansowym, dlatego głównymi determinantami opinii powinny być zmienne finansowe. Dla uzyskania porównania między różnymi podmiotami zastosowano podstawowe wskaźniki finansowe.

Grupy kapitałowe i podmioty notowane na rynku publicznym (podmioty zainteresowania publicznego) poddawane są szczególnym dodatkowym procesom rewizyjnym. Stąd drugą grupą zmiennych uwzględnionych w modelu są charakterystyki podmiotu badanego.

W modelu uwzględniono także charakterystyki firmy audytorskiej. Duże międzynarodowe sieci audytorskie mają rozbudowany system wewnętrznej kontroli z uwagi na transgraniczne świadczenie usług. Wielkość wynagrodzenia firmy audytorskiej związana jest z jakością i pracochłonnością badania. Dodatkowe usługi nierewizyjne, świadczone równolegle z audytem, mogą prowadzić zarówno do efektu transferu wiedzy i kompetencji, jak i w szczególnych sytuacjach mogą wpływać na percepcję niezależności firmy audytorskiej, stąd ich obecność w testowanej specyfikacji modelu predykcji. W związku z powyższym w tym badaniu stawiamy hipotezę, iż każda grupa zmiennych może mieć wpływ na opinię biegłego rewidenta.

Za model odniesienia przyjęliśmy model logitowy szeroko stosowany w poprzednich badaniach.

2. Dane i metodologia

Dane finansowe wykorzystane w badaniu pochodzą z baz EMIS i Amadeus, zaś dane o wynagrodzeniu biegłych rewidentów zostały zebrane ręcznie w oparciu o sprawozdania finansowe jednostek dostępne elektronicznie w Internecie lub w Monitorze Polskim B. Dane obejmują 774 sprawozdania 219 firm z okresu 2010–2013. Biegli rewidenci wydali 488 opinii bez zastrzeżeń oraz 286 opinii modyfikowanych.

Wyboru zmiennych dokonano na postawie wcześniejszych badań [Hay i inni, 2006; Górska, Staszkiewicz, 2018]. Szerszy przegląd badań nad rewizją przedstawił P. Staszkiewicz [2017].

Zmienne, które są brane pod uwagę w badaniu potencjalnych czynników wpływających na opinie biegłych rewidentów, podzielono umownie na trzy grupy. Opis zmiennych podano poniżej:

Wskaźniki finansowe:

- ROA – Stopa zwrotu z aktywów,
- CACL – Wskaźnik bieżącej płynności (Aktywa obrotowe/Zobowiązania bieżące),
- TDTA – Wskaźnik zadłużenia (Zobowiązania ogółem/Aktywa ogółem),
- INVREC – Zapasy i należności/Aktywa ogółem.

Inne charakterystyki firmy:

- TA – Aktywa ogółem,
- SUB – Liczba podmiotów zależnych,
- LISTED – Czy spółka jest notowana na GPW? (1 – tak).

Zmienne związane z audytem:

- BIG4 – Czy firma audytorska jest jedną z czterech największych firm audytorskich (Deloitte, Ernst&Young, KPMG, PwC) (1 – tak),
- AF – Wynagrodzenie za audyt,
- NAF – Wynagrodzenie za usługi konsultingowe, niezwiązane z audytem.

Metodologia

Punkt odniesienia stanowi model logitowy o następującej specyfikacji:

$$Y = \beta_0 + \beta_1 \text{ROA} + \beta_2 \text{CACL} + \beta_3 \text{TDTA} + \beta_4 \text{INVREC} + \beta_5 \text{TA} + \\ + \beta_6 \text{SUB} + \beta_7 \text{LISTED} + \beta_8 \text{BIG4} + \beta_9 \text{AF} + \beta_{10} \text{NAF} + \varepsilon \quad (1)$$

gdzie:

Y – zmienna zero-jedynkowa (1 – raport modyfikowany, 0 – raport bez zastrzeżeń).

W przeprowadzonym badaniu zbiór danych podzielono na dwie części: zbiór uczący, obejmujący sprawozdania z okresu 2010–2012, i zbiór testowy, obejmujący sprawozdania z roku 2013. Model oszacowano na podstawie danych w okresu 2010–2012, a następnie wykorzystano do predykcji opinii z okresu 2013.

Badanie powtórzono przy zastosowaniu algorytmu lasów losowych, przy analogicznym podziale danych. Obliczenia dokonano w środowisku R.

Najogólniej mówiąc, algorytm lasów losowych polega na stworzeniu i łączeniu wielu różnych drzew klasyfikacyjnych. Każde drzewo jest tworzone na losowej próbie n obserwacji pobieranych ze zwracaniem ze zbioru uczącego (próba bootstrapowa). Ponadto atrybuty (zmienne) wybierane przy konstrukcji każdego drzewa również są wybierane losowo. W trakcie budowania drzewa w każdym węźle podział odbywa się poprzez wylosowanie bez zwracania m spośród p atrybutów ($m \leq p$). Parametr m ustalany jest najczęściej w następujący sposób: $m = \sqrt{p}$, sugerowany w literaturze przedmiotu [Hastie i inni, 2009, s. 611]. Proces ten jest powtarzany w następnym węźle, a budowanie drzewa bez przycinania trwa, jeżeli to możliwe, do momentu uzyskania w liściach elementów z jednej tylko klasy. Dany wektor obserwacji jest klasyfikowany przez wszystkie drzewa, a ostatecznie zaklasyfikowany do klasy, w której wystąpił najczęściej.

Algorytm lasów losowych pozwala na uniknięcie problemów związanych z pojedynczym drzewem klasyfikacyjnym, takich jak niestabilność drzewa, czyli duża wrażliwość na małe zmiany w próbie, a także problem „przeuczenia" i potrzeba „przycinania" drzewa.

W przedstawionym badaniu skonstruowano model lasu losowego na podstawie zbioru uczącego. Jako miarę niejednorodności klas w węzłach zastosowano współczynnik Giniego. Następnie dokonano prognozy opinii biegłego rewidenta dla zbioru testowego.

Prognozowanie na podstawie lasu losowego polega na:

- określeniu prognoz z każdego drzewa wchodzącego w skład lasu,
- wyznaczeniu prognozy całego modelu na podstawie głosowania większościowego.

Prognozowania dokonuje się w następujący sposób:

$$\hat{F}^*(x_i) = \arg\max_j \{\sum_{m=1}^{M} I(\hat{F}_m(x_i = C_j))\} \tag{2}$$

gdzie:

$\hat{F}^*(x_i)$ – klasa, do której została przypisana obserwacja x_i przez las losowy,

$\hat{F}_m(x_i = C_j)$ – klasa, do której została przypisana obserwacja x_i przez m-te drzewo,

C_j – klasa j.

Algorytm lasów losowych pozwala także wskazać i uporządkować zmienne dyskryminujące opinię, co umożliwia zrealizowanie dodatkowego celu badania – określenie czynników wpływających na decyzje biegłego rewidenta.

3. Wyniki empiryczne

Na podstawie zbioru uczącego, składającego się z 574 obserwacji (sprawozdania z lat 2010–2012), oszacowano model logitowy oraz zbudowano model lasu losowego. Następnie na podstawie obu modeli sporządzono prognozy opinii biegłego rewidenta dla 200 sprawozdań z roku 2013 i porównano je z rzeczywistymi opiniami. W tablicy 1 przedstawiono tzw. tabele trafności dla algorytmu lasów losowych, modelu logitowego zarówno dla zbioru uczącego, jak i testowego.

Tablica 1. Tabele trafności

	Zbiór	Uczący		Testowy	
Model	Przewidywane / Empiryczne	0	1	0	1
Las losowy	0	**311**	55	**99**	36
	1	109	**99**	23	**42**
Model logitowy	0	**334**	22	**114**	8
	1	163	**45**	61	**17**

Oznaczenia:
0 – raport bez zastrzeżeń,
1 – raport modyfikowany.
Źródło: Opracowanie własne.

Dla zbioru uczącego liczba przypadków poprawnej klasyfikacji lasu losowego wynosi 410 spośród wszystkich 574 przypadków, czyli 71,2% wszystkich obserwacji. Dla modelu logitowego liczba ta wynosi 389, czyli 67,8%. Dla zbioru testowego analogicznie liczba przypadków poprawnej predykcji lasu losowego wynosi 141 spośród 200, czyli 70,5% przypadków, a dla modelu logitowego – 131, czyli 65,5%.

Na podstawie danych z tabel trafności obliczono także:

– TPR (*True Positive Rate*), czyli odsetek poprawnie sklasyfikowanych jedynek (raportów modyfikowanych), według wzoru:

$$TPR = \frac{TP}{T} \tag{3}$$

gdzie: TP – liczba poprawnie sklasyfikowanych 1, T – liczba empirycznych 1;

oraz

– TNR (*True Negative Rate*), czyli odsetek poprawnie sklasyfikowanych zer (raportów bez zastrzeżeń), według wzoru:

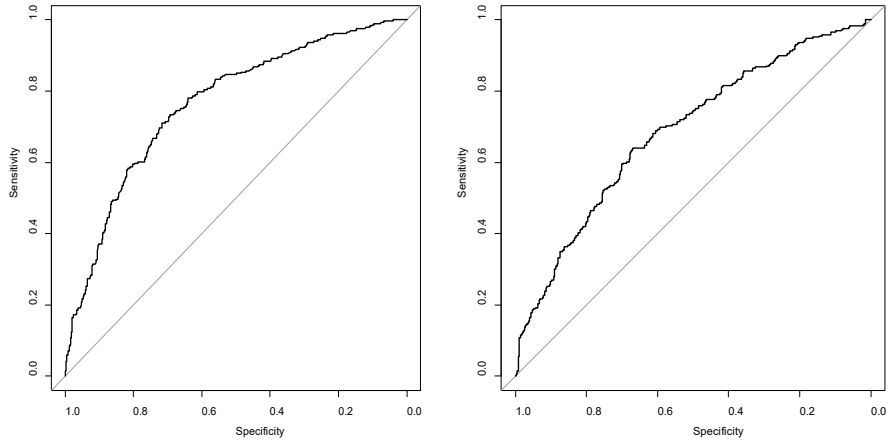
$$TNR = \frac{TN}{N} \tag{4}$$

gdzie: TN – liczba poprawnie sklasyfikowanych 0, N – liczba empirycznych 0.

TPR, znany również jako czułość testu (*Sensitivity*), czyli odsetek poprawnie sklasyfikowanych jedynek, dla lasu losowego wynosi 64,3%, a dla modelu logitowego 67,2%. TNR, znany jako swoistość testu (*Specifity*), dla lasu losowego wynosi 74,1%, a dla modelu logitowego 67,8%,

Na rysunku 1 pokazano krzywe ROC (*Reciever Operating Characteristic*), prezentujące graficznie zależność pomiędzy swoistością (TNR) a czułością testu (TPR), obliczone dla całego zbioru danych. (Dla ułatwienia interpretacji na osi odciętych przedstawiono 1-TNR).

Rysunek 1. Krzywa ROC dla lasu losowego i modelu logitowego

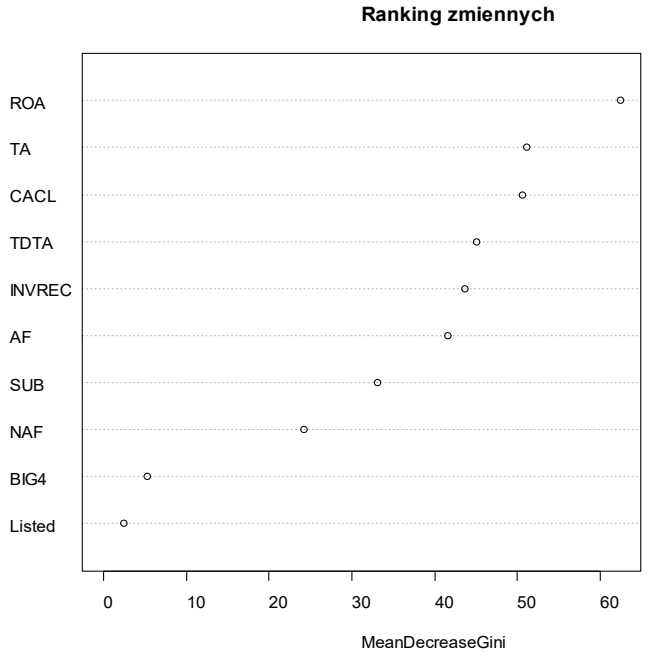


Źródło: Opracowanie własne.

Pole pod krzywą (*Area Under the Curve* – AUC) jest miarą dokładności testu. Dla lasu losowego wynosi ona 76,3%, a dla modelu logitowego 68,5%. Obliczono także 95% przedziały ufności dla tej miary – dla lasu losowego przedział ten jest następujący: (72,9; 79,7), a dla modelu logitowego (64,6; 72,4). Zatem algorytm lasów losowych charakteryzuje się większą dokładnością predykcji w stosunku do modelu logitowego i może być używany do prognozowania.

Ponadto algorytm ten pozwala na zidentyfikowanie i uporządkowanie zmiennych dyskryminujących opinię biegłego rewidenta. Na rysunku 2 pokazano wykres uporządkowanych zmiennych wpływających na opinię biegłego rewidenta. Na osi Y są uporządkowane zmienne wpływające na opinię biegłego rewidenta według stopnia ich ważności. W największym stopniu na opinię tę wpływają zmienne finansowe, dotyczące sytuacji finansowej firmy (ROA, TA, CACL, TDTA, INVREC). Następną zmienną dyskryminującą opinię jest wynagrodzenie za usługi audytorskie (AF), lecz można to wytłumaczyć faktem, iż często zmienna ta jest skorelowana z wielkością przedsiębiorstwa, czyli zmienną Aktywa ogółem (TA). Najmniejszy wpływ ma fakt notowania spółki na giełdzie (LISTED) oraz wielkość audytora (BIG4).

Rysunek 2. Ranking zmiennych wpływających na opinię biegłego rewidenta



Źródło: Opracowanie własne.

Wyniki estymacji modelu logitowego podano w tablicy 2. Zmienne istotnie wpływające na opinię biegłego rewidenta to stopa zwrotu z aktywów ROA, zmienna BIG4 oraz Aktywa ogółem TA. W celu weryfikacji modelu podano wynik testu ilorazu wiarygodności wskazujący na łączną istotność zmiennych objaśniających, a także liczbę poprawnych predykcji oraz tzw. zliczeniowy R^2=67,8%. Dodatkowo w celu dokonania oceny współliniowości obliczono także czynniki inflacji wariancji (VIF).

Tablica 2. Estymacja Logit z wykorzystaniem 574 obserwacji

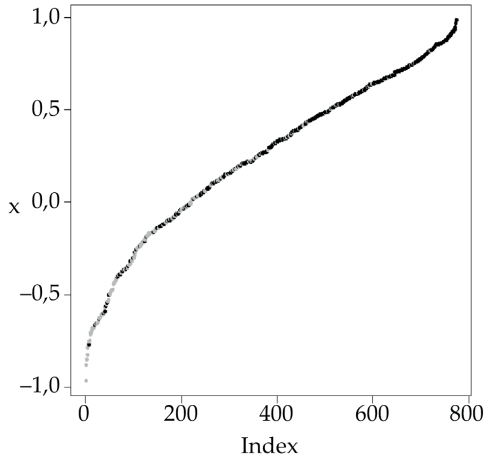
Zmienna zależna: Opinion

Zmienna	Współczynnik	Błąd stand.	Z	Wartość -p		VIF
const	0,082607	0,448445	0,1842	0,8539		
ROA	−3,82504	0,754818	−5,067	4,03E-07	***	1,071
CACL	0,003282	0,006282	0,5224	0,6014		1,026
TDTA	−0,135601	0,734048	−0,1847	0,8534		2,848
INVREC	0,189107	0,596918	0,3168	0,7514		2,855
TA	8,70E-08	4,55E-08	1,912	0,0559	*	3,164
SUB	0,004401	0,006668	0,6599	0,5093		1,272
LISTED	−0,557401	0,43287	−1,288	0,1979		1,045
BIG4	−0,754963	0,294749	−2,561	0,0104	**	1,293
AF	−7,24185e-07	6,25E-07	−1,158	0,2467		4,191
NAF	2,77E-07	7,61E-07	0,3639	0,716		3,469

Test ilorazu wiarygodności: Chi-kwadrat(4) = 52,402 [0,0000]
Liczba przypadków „poprawnej predykcji" = 389 (67,8%)
Źródło: Opracowanie własne.

Na rysunku 2 pokazano tzw. granicę lasu losowego. Wykres ten obrazuje podział danych ze względu na interesującą nas cechę (raport niemodyfikowany vs. raport modyfikowany). Jeżeli na wykresie byłyby widoczne dwa rozłączne podzbiory, wówczas metoda klasyfikacji oparta na algorytmie lasów losowych charakteryzowałaby się 100% skutecznością. Na rysunku punkty ciemniejsze odpowiadają niemodyfikowanym raportom, a jaśniejsze – modyfikowanym.

Rysunek 2. Granica lasu losowego



Źródło: Opracowanie własne.

Zbiory te nie są rozłączne, co wynika z faktu, iż istnieją także inne czynniki determinujące decyzje biegłego rewidenta poza zmiennymi, które są brane pod uwagę w badaniu. Stąd klasyfikacja, a także predykcja na podstawie tego zbioru danych zawsze będzie obarczona błędem.

Oceniając powyższe wyniki, należy wszakże zwrócić uwagę na fakt, iż algorytm drzew losowych wskazuje na inne uporządkowanie zmiennych dyskryminujących niż model logitowy.

Zakończenie

Przedstawione badanie wykazało, iż algorytm lasów losowych charakteryzuje się większą dokładnością przy prognozowaniu opinii biegłego rewidenta w porównaniu do modelu logitowego. Zmienne wpływające na opinię biegłego rewidenta to przede wszystkim wskaźniki finansowe badanej spółki. Najmniejsze znaczenie ma fakt notowania spółki na giełdzie oraz wielkość audytora.

Literatura

Altman E.I. (1982), *Accounting implications of failure prediction models*, „Journal of Accounting, Auditing and Finance", Vol. 6, No. 1.

Altman E.I., McGough T. (1974), *Evaluation of a company as a going concern*, „Journal of Accountancy", Vol. 138, No. 6.

Altman E., Iwanicz-Drozdowska M., Laitinen E., Suvas A. (2016), *Financial and non-financial variables as long-horizon predictors of bankruptcy*, „The Journal of Credit Risk", Vol. 12, No. 4.

Dopuch N., Holthausen R.W., Leftwich R.W. (1987), *Predicting Audit Qualification with Financial and Market Variables*, „The Accounting Review", Vol. 72, No. 3.

Gostomczyk B. (2013), *Ryzyko utraty zdolności jednostki do kontynuowania działalności w teorii i praktyce gospodarczej*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse. Rynki finansowe. Ubezpieczenia", nr 61, t. 2.

Górska R., Staszkiewicz P. (2018) *Non audit fee and the Auditiee's Failure Risk*, „Zeszyty Teoretyczne Rachunkowości" (w druku).

Gruszczyński M., Pajdo B. (2003), *Wskaźniki finansowe a opinia biegłego rewidenta*, „Bank i Kredyt", nr 34(5).

Hastie T., Tibshirani R., Friedman J., (2009), *The Elements of Statistical Learning. Data mining, inference and prediction*, Springer, Second Edition.

Hay D., Knechel R., Li V., (2006), *Non-Audit Services and Auditor Independence: New Zealand Evidence*, „Journal of Business Finance and Accounting", Vol. 33, No. 5–6.

Keasey K., Watson R. (1987), *Non-Financial Symptoms and the Prediction of Small Company Failure: A Test of Argenti's Hypotheses*, „Journal of Business Finance & Accounting", Vol. 14, No. 3.

Kennedy D.B., Shaw W.H. (1991), *Evaluating financial distress resolution using prior audit opinions*, „Contemporary Accounting Research", Vol. 8, No. 1.

Koh H.C., Killough L. (1990), *The use of multiple discriminant analysis in the assessment of the the going concern status of an audit client*, „Journal of Business Finance & Accounting", Vol. 17, No. 2.

Mutchler J.F. (1985), *A Multivariate Analysis of the Auditor's Going-Concern Opinion Decision*, „Journal of Accounting Research", Vol. 23, No. 2.

Staszkiewicz P., (2017), *Quo Vadis, Auditor? Synteza Badań Nad Rewizją Finansową w Polskiej Domenie Publicznej*, „Zeszyty Teoretyczne Rachunkowości", nr 95(151).

Streszczenie

Pytanie badawcze przedstawione w niniejszej analizie odnosi się do problemu, na ile prognozowanie typu opinii biegłego rewidenta jest właściwsze za pomocą algorytmu lasów losowych w porównaniu do regresji logistycznej. Na próbie 774 opinii biegłych rewidentów wydanych w latach 2010–2013 oszacowano modele prognozowania typu opinii. Wyniki wskazują na lepszą trafność modelu w przypadku zastosowania algorytmu lasów losowych. Wyniki wskazują także na niską zdolność predykcyjną zmiennych określających wielkość firmy audytorskiej oraz fakt notowania podmiotu badanego na rynku publicznym. Badanie dostarcza wstępne wyniki dla gospodarki polskiej, które dają podstawy do dalszych badań nad zastosowaniem metod uczenia maszynowego w rewizji finansowej.

Słowa kluczowe

lasy losowe, regresja logistyczna, audyt, rewizja finansowa, klasyfikacja

Application of the random forest algorithm for prediction of the modification of auditor opinion (Summary)

This paper asks wheatear the random forest algorithm outperforms logit modeling of the audit opinion. We have used a sample of the 774 audit opinions issued in years 2010-2013 in Poland . We applied the unformal set of control variables and compared the results of the logit regression and the random forest classification. The results indicate that random forest algorithm outperforms the logit regression. In addition, our results support the conclusion that the size of the auditor and the listing of the company's shares are weak predictors of the type of the opinion. The study provides robust results about Polish economy, which can be used for the further application of the machine learning methods in audit research.

Keywords

random forest, logit regression, audit, classification

Roman Machuga*

Modelowanie systemu informacyjnego rachunkowości zarządczej na przykładzie przedsiębiorstwa branży spożywczej

Wstęp

Szybki wzrost objętości informacji, podwyższenie jej jakości i udoskonalenie składu strukturowego i form przedstawienia stworzyły warunki wstępne dla opracowania nowych zasobów wsparcia informacyjnego oraz przepływów informacji. Dlatego można mówić nie o rachunkowości lub informacji księgowej w systemie zarządzania procesami produkcyjnymi, a o istnieniu oddzielnego niezależnego systemu informacyjnego rachunkowości zarządczej. W nowoczesnych warunkach organizacji pracy w przedsiębiorstwach system informacyjny rachunkowości jest fundamentalny (centralny), ponieważ łączy praktycznie cały niezbędny dla zarządzania zbiór informacji. Taki system można przedstawić w postaci zwyczajnego modelu informacyjnego – modelowanie jest jedną z głównych metod jego badania. System informacyjny rachunkowości zarządczej jest uogólnionym modelem działalności produkcyjno-gospodarczej przedsiębiorstw przemysłowych oraz innych podmiotów gospodarczych.

System informacyjny rachunkowości zarządczej, jak i inne systemy informacji, ma swoją konkretną funkcję, cel i zadania. Głównym celem funkcjonowania takiego systemu informacyjnego w przedsiębiorstwie jest wsparcie informacyjne procesu decyzyjnego.

Rachunkowość na wszystkich etapach rozwoju społeczeństwa ludzkiego była zasobem tworzenia informacji, która dalej wykorzystywana była do oceny stanu realnego przedsiębiorstwa, działań gospodarczych, faktów i podjęcia w końcu uzasadnionych decyzji zarządczych. Innymi słowy, zabezpieczenie informacyjne procesu podjęcia decyzji zarządczych to zadanie podstawowe, główne przeznaczenie funkcjonalne rachunkowości zarządczej w gospodarce rynkowej.

Branża spożywcza charakteryzuje się częstymi zmianami w środowisku zewnętrznym: zmiana popytu konsumentów w zależności od sezonu, zmniejszenie lub zwiększenie cen zakupu surowca, nowi dostawcy itp.

* Asystent, Katedra Metod Ilościowych, Wydział Nauk Ekonomicznych, Uniwersytet Warmińsko-Mazurski w Olsztynie, ul. Michała Oczapowskiego 2, 10-719 Olsztyn, roman.machuga@uwm.edu.pl

Dlatego próba modelowania informacyjnego systemu rachunkowości zarządczej będzie podjęta dla hipotetycznego przedsiębiorstwa.

Celem artykułu jest zaproponowanie uproszczonego modelu funkcjonowania rachunkowości zarządczej na przykładzie przedsiębiorstwa branży spożywczej. Głównymi metodami badawczymi są analiza systemowa i modelowanie. Analiza systemowa została zastosowana do poznania układów organizacyjnych i sposobów ich działania wewnątrz przedsiębiorstwa, do opisania i wyjaśniania relacji zachodzących w systemie, zwłaszcza do analizy procesu podejmowania decyzji zarządczych. Do opracowania schematów procesów decyzyjnych w przedsiębiorstwie oraz funkcjonowania informacyjnego systemu rachunkowości zarządczej zastosowano modelowanie graficzne, w tym z wykorzystaniem algorytmizacji za pośrednictwem zunifikowanego języka modelowania UML.

Przeprowadzone badania nie mają charakteru kompleksowego i stanowią w perspektywie fundament pogłębionych analiz.

1. Podejmowanie decyzji zarządczej

Problematyka podejmowania decyzji zarządczych oraz modelowania w rachunkowości zarządczej dość często jest przedmiotem badania polskich naukowców. Na przykład J. Czarnecki w swej pracy rozpatruje rachunkowość zarządczą w kontekście koncepcji podejmowania decyzji [Czarnecki, 2013], a B. Nita bada możliwości modelowania rachunkowości na potrzeby zarządzania [Nita, 2009]. Natomiast M. Wójcik pokazuje proces podejmowania decyzji zarządczych w przedsiębiorstwie oraz rolę rachunkowości zarządczej w tym procesie [Wójcik, 2009], zaś E. Nowak w swojej książce rozpatruje rachunkowość zarządczą jako system informacyjny oraz źródło informacji zarządczej [Nowak, 2018]. W dotychczasowych badaniach nie zajmowano się możliwościami zastosowania algorytmizacji do podejmowania decyzji zarządczych oraz funkcjonowania informacyjnego systemu rachunkowości zarządczej.

Pod pojęciem decyzji zarządczej rozumiemy realizację jednego z poprzednio opracowanych wariantów (projektów). V.M. Tomashevskyi wydziela kilka głównych kierunków podejmowania decyzji na podstawie wyników modelowania:

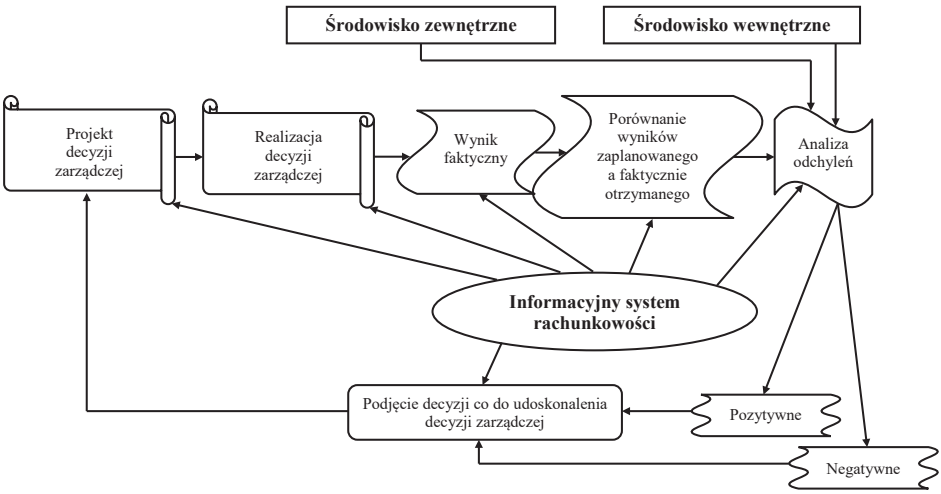
- wyszukiwanie najlepszych pod względem konkretnych kryteriów efektywności znaczeń parametrów złożonych systemów zarządzania,
- wyszukiwanie znaczenia optymalnego kryterium efektywności systemu,
- porównanie wariantów alternatywnych struktury systemu oraz wyznaczenie najlepszego wśród nich,
- modelowanie sytuacji przypadkowych według scenariusza typu „co będzie, jeśli…" [Tomashevskyi, 2005, s. 300].

Podobnego zdania jest W. Rebizant, który pisze, że „Niezależnie od przyjętego modelu decyzyjnego w procesie decyzyjnym powinna zostać podjęta optymalna decyzja, będąca jedną z decyzji dopuszczalnych, spełniająca najlepiej sformułowane kryteria oceny. Decyzją dopuszczalną jest decyzja (wybór) spełniająca wszystkie warunki ograniczające. Pod pojęciem kryterium decyzyjnego kryje się natomiast przyporządkowanie ilościowej lub jakościowej miary korzyści, użyteczności, kosztów i zysków danej decyzji" [Rebizant, 2012, s. 31].

Dlatego fakt podjęcia konkretnej decyzji zarządczej w żadnym przypadku nie można uważać za zakończenie procesu zarządzania. Ponieważ ten proces koniecznie przewiduje otrzymanie wyniku oczekiwanego, to po praktycznej realizacji decyzji należy zaplanować wykonanie, na przykład, na górnym poziomie hierarchii zarządzania, kompleksu nowych zadań powiązanych z porównaniem wyniku zaplanowanego z faktycznie otrzymanym, fiksacją odchyleń, analizą ich podstaw. Zebraną i opracowaną w taki sposób informację przez blok tzw. sprzężenia zwrotnego przekazuje się dla podjęcia już nowej decyzji zarządczej dotyczącej udoskonalenia modelu imitacyjnego (opracowania nowego projektu decyzji). Schemat tego procesu przedstawiono na rysunku 1.

Na podstawie zaproponowanego schematu procesu analizy efektywności podjętych decyzji zarządczych można opracować uproszczony model funkcjonowania systemu informacyjnego rachunkowości zarządczej na przykładzie przedsiębiorstwa branży spożywczej.

Rysunek 1. Schemat procesu analizy efektywności podjętych decyzji zarządczych



Źródło: Opracowanie własne.

2. Model funkcjonowania systemu informacyjnego rachunkowości zarządczej

Opracowanie uproszczonego modelu funkcjonowania systemu informacyjnego rachunkowości zarządczej będzie realizowane dla hipotetycznego przedsiębiorstwa, którego głównym kierunkiem działalności gospodarczej jest produkcja i sprzedaż lodów.

W analizowanym przedsiębiorstwie skutecznie funkcjonuje system informacyjny rachunkowości zarządczej. W systemie informacyjnym, jak i w jakimkolwiek systemie, istnieje środowisko wewnętrzne i zewnętrzne.

Ze środowiska wewnętrznego na podstawie danych rachunkowości dla użytkowników dostępne są informacje o:

- kosztach produkcji, magazynowania i dostawy wyrobów,
- cenach surowca, wyrobów pomocniczych i energii,
- liczbie wyrobów produkowanych miesięcznie w ciągu roku,
- rezerwach na zwiększenie zdolności produkcyjnych itd.

Ze środowiska zewnętrznego do systemu wpływa szereg różnorodnych informacji. Na przykład:

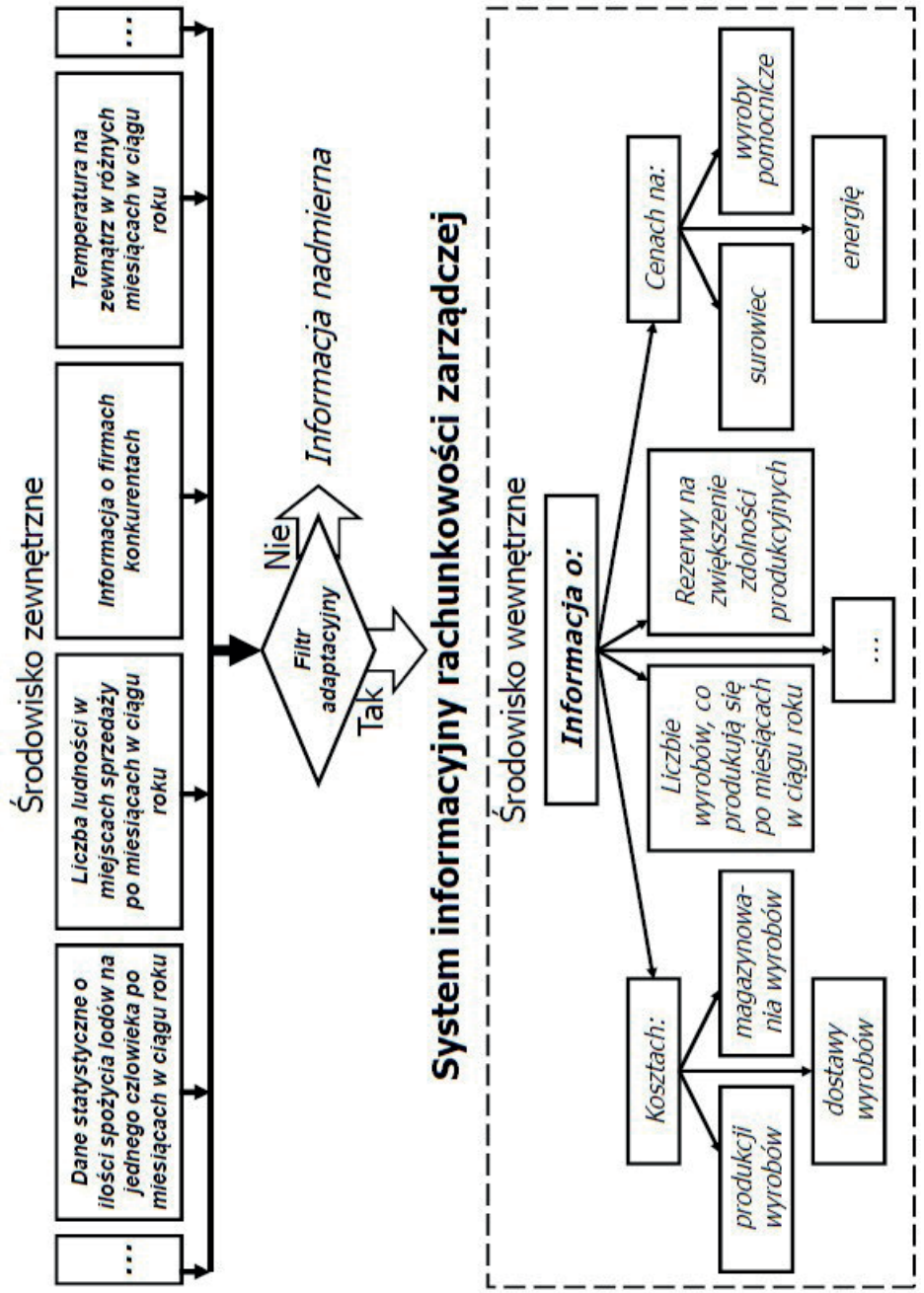
- dane statystyczne o ilości spożycia lodów na jednego człowieka w miesiącu w ciągu roku,
- liczba ludności w miejscach sprzedaży w miesiącach w ciągu roku,
- informacja o firmach konkurentach,
- temperatura na zewnątrz w różnych miesiącach w ciągu roku.

Informacja nadchodząca ze środowiska zewnętrznego bardzo często nie jest informacją księgową czy nawet ekonomiczną. Informacja może być zarówno korzystna, jak i niekorzystna dla działalności przedsiębiorstwa. Dla wyboru wśród dużych zbiorów informacji korzystnej, niezbędnej dla efektywnego funkcjonowania w przedsiębiorstwie systemu informacyjnego rachunkowości zarządczej, potrzeba zastosować oryginalny filtr adaptacyjny, w którym warunki będą zmieniać się w zależności od typu informacji, potrzeb produkcji, wymagań do podjęcia decyzji zarządczych itd.

Uproszczony schemat opisanych etapów funkcjonowania systemu informacyjnego rachunkowości zarządczej pokazano na rysunku 2.

Jeżeli informacja wpływająca odpowiada wymaganiom filtru adaptacyjnego, to wpływa ona dalej w system, jeśli nie – to jest ona nadmierną i zostaje poza granicami systemu informacyjnego. Na konieczność zastosowania filtru adaptacyjnego zwracał uwagę jeszcze w 1983 r. K.N. Narybajev w swej monografii poświęconej automatyzacji podsystemu rachunkowości [Narybajev, 1983, s. 83].

Rysunek 2. Uproszczony schemat funkcjonowania systemu informacyjnego rachunkowości zarządczej (część 1)



Źródło: Opracowanie własne.

Na podstawie informacji nadchodzącej ze środowiska zewnętrznego (przez filtr adaptacyjny) i wewnętrznego w systemie informacyjnym rachunkowości zarządczej wytwarzają się kompleksowe dane systematyzowane dla przygotowania wariantów projektów decyzji zarządczych. Na podstawie wytworzonych danych personel zarządczy bezpośrednio przygotowuje projekty decyzji, np. o zwiększeniu lub zmniejszeniu zakupu surowcu (ilość, koszty dostaw itp.), zwiększeniu lub zmniejszeniu zdolności produkcyjnych, zmianie technologii czy cyklu technologicznego, konieczności przemyślenia polityki cenowej itd. Dla opracowanych projektów decyzji koniecznie prowadzi się przybliżone obliczenie kosztów, dochodu, podatków.

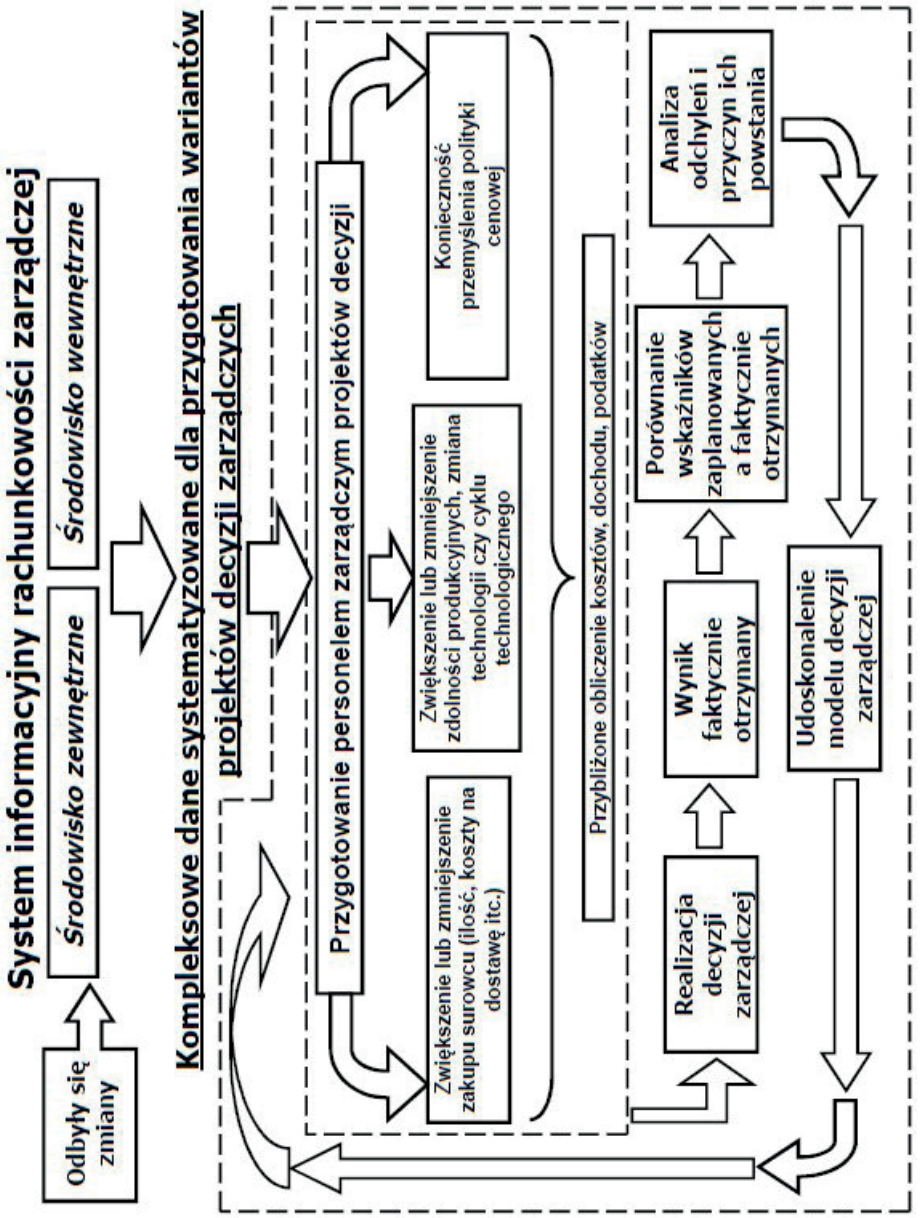
Po przygotowaniu wariantów projektów decyzji zarządczych i wyborze optymalnego następuje etap jego realizacji. Kolejny etap to otrzymanie wyniku faktycznego. Faktyczny wynik otrzymany musi być koniecznie porównany ze wskaźnikami zaplanowanymi. Na podstawie porównania wykonuje się analizę odchyleń oraz przyczyn ich powstania.

Kolejny etap funkcjonowania systemu to udoskonalenie modelu imitacyjnego (projektu decyzji zarządczej) z uwzględnieniem zmian, które odbyły się (lub mogą odbyć się) w środowisku zewnętrznym, oraz opracowanie nowych projektów decyzji zarządczych. Uproszczony schemat opisanych powyżej etapów przedstawiono na rysunku 3.

Etapy przygotowania projektów decyzji zarządczych, realizacji optymalnej decyzji zarządczej, otrzymanie wyniku faktycznego, porównanie wskaźników zaplanowanych a faktycznie otrzymanych, analiza odchyleń i przyczyn ich powstania oraz udoskonalenie modelu imitacyjnego systemu informacyjnego rachunkowości zarządczej to proces cykliczny, który powtarza się tak długo, dopóki przedsiębiorstwo działa [Machuga, 2009].

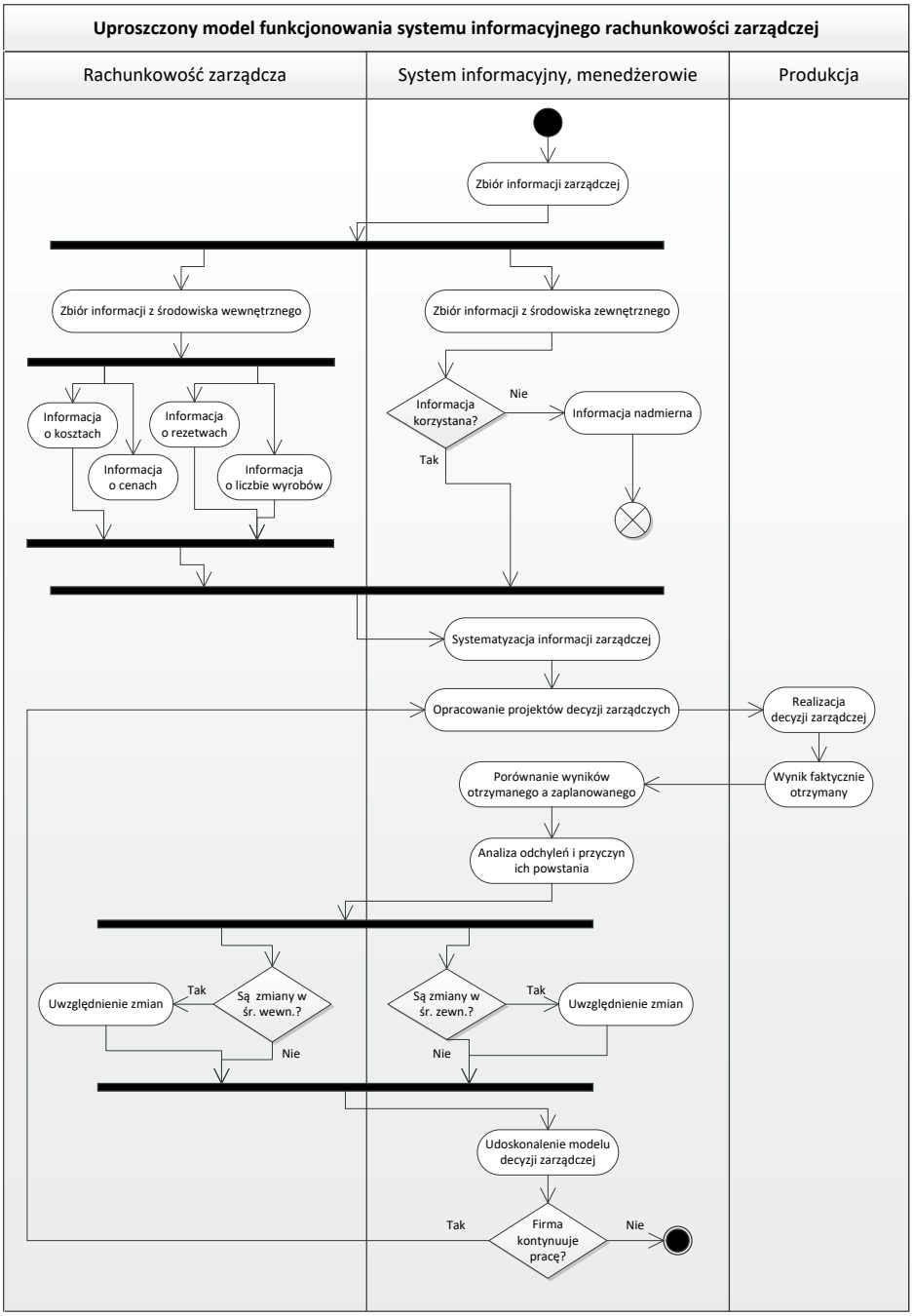
Przedstawione na rysunkach 2 i 3 uproszczone schematy funkcjonowania systemu informacyjnego rachunkowości zarządczej dodatkowo zostały opracowane z wykorzystaniem zunifikowanego języka modelowania UML. Otrzymany model pokazano na rysunku 4.

Rysunek 3. Uproszczony schemat funkcjonowania systemu informacyjnego rachunkowości zarządczej (część 2)



Źródło: Opracowanie własne.

Rysunek 4. Diagram aktywności – model systemu w języku UML



Źródło: Opracowanie własne.

Zakończenie

Na podstawie przeprowadzonego badania można wyciągnąć następujące wnioski:

1. W procesie podejmowania decyzji zarządczych głównym dostawcą niezbędnych informacji jest rachunkowość zarządcza.
2. W informacyjnym systemie rachunkowości zarządczej w procesie decyzyjnym musi być wykorzystywany filtr adaptacyjny dla filtrowania informacji ze środowiska zewnętrznego.
3. Proces decyzyjny z wykorzystaniem informacyjnego systemu rachunkowości zarządczej to proces cykliczny.
4. Wymagane etapy procesu decyzyjnego: opracowanie projektów i podjęcie decyzji, realizacja decyzji, otrzymanie wyników, porównanie wyników z zaplanowanymi, analiza odchyleń, udoskonalenie modelu decyzji zarządczej.
5. Opracowany uproszczony model funkcjonowania informacyjnego systemu rachunkowości zarządczej może być wykorzystany dla szczegółowego modelowania podprocesów procesu decyzyjnego.
6. Przeprowadzone badania stanowią w perspektywie fundament pogłębionych analiz.

Literatura

Czarnecki J. (2013), *Rachunkowość zarządcza zgodna z koncepcją podejmowania decyzji opartych na dowodach*, „Zeszyty Teoretyczne Rachunkowości", nr 128.

Machuga R.I. (2009), *System informacyjny rachunkowości w zarządzaniu przedsiębiorstwami przemysłu spożywczego*, Tarnopol.

Narybajev K.N. (1983), *Organizacja i metodologia rachunkowości w warunkach systemów automatyzowanych*, „Финансы и статистика", Moskwa.

Nita B. (2009), *Modelowanie rachunkowości na potrzeby zarządzania dokonaniami przedsiębiorstwa*, „Zeszyty Teoretyczne Rachunkowości", nr 105.

Nowak E. (2018), *Rachunkowość zarządcza w przedsiębiorstwie*, CeDeWu, Warszawa.

Rebizant W. (2012), *Metody Podejmowania Decyzji*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, http://www.kierunkizamawiane.pwr.wroc.pl/materialy/SkryptMPDpopr2WRebizant.pdf, dostęp: 30.03.2017.

Tomashevskyi V.M. (2005), *Modelowanie systemów*, Видавнича група BHV, Kijów.

Wójcik M. (2009), *Proces podejmowania decyzji w przedsiębiorstwie*, Wydawnictwo E-bookowo.

Streszczenie

Celem artykułu jest zaproponowanie uproszczonego modelu funkcjonowania rachunkowości zarządczej na przykładzie przedsiębiorstwa branży spożywczej. Modelowany system przewiduje wykorzystanie informacji zarówno księgowej ze środowiska wewnętrznego, jak i różnorodnej informacji ze środowiska zewnętrznego. Zaproponowano wykorzystanie filtra adaptacyjnego dla określenia

korzystności informacji wchodzącej. Sam proces decyzyjny w systemie informacyjnym jest procesem cyklicznym, składającym się z etapów przygotowania projektów decyzji zarządczych, realizacji decyzji, otrzymania wyniku faktycznego, porównania wyniku otrzymanego z zaplanowanym, analizy odchyleń i przyczyn ich powstania, udoskonalenia modelu decyzji zarządczej. Dodatkowo został opracowany i przedstawiony uproszczony model funkcjonowania systemu informacyjnego rachunkowości z wykorzystaniem zunifikowanego języka modelowania UML.

Słowa kluczowe
system informacyjny, rachunkowość zarządcza, decyzja, modelowanie

Modeling of management accounting information system (Summary)

The purpose of the paper is to try to create an information system in management accounting. The modelled system provides for the use of both company-generated accounting information (internal data) and the environment external (external data). It was proposed to use an adaptive filter to determine the usefulness of incoming information. The decision-making process itself in the information system is a cyclical process. It consists of the following stages: preparation of draft management decisions, implementation of the decision, receipt of the actual result, comparison of the result obtained with the scheduled, analysis of the deviations and causes thereof, improvement of the management decision model. In addition, a model for the operation of the accounting information system was developed and presented using a unified UML modelling language.

Keywords
information system, management accounting, decision, modeling