

Warsaw University of Life Sciences – SGGW
Institute of Economics and Finance
Department of Econometrics and Statistics

**QUANTITATIVE METHODS
IN ECONOMICS**

**METODY ILOŚCIOWE W BADANIACH
EKONOMICZNYCH**

Volume XXI, No. 1

Warsaw 2020

EDITORIAL BOARD

Editor-in-Chief: Bolesław Borkowski

Vice-Editor-in-Chief: Hanna Dudek

Managing Editor: Grzegorz Koszela

Theme Editors:

Econometrics: Bolesław Borkowski

Multidimensional Data Analysis: Wiesław Szczesny

Mathematical Economy: Zbigniew Binderman

Analysis of Labour Market: Joanna Landmesser

Financial Engineering: Monika Krawiec

Data Science: Michał Gostkowski

Statistical Editor: Wojciech Zieliński

Technical Editors: Jolanta Kotlarska, Elżbieta Saganowska

Language Editor: Agata Cienkusz

Native Speaker: Yochanan Shachmurove

Editorial Assistant: Luiza Ochnio

SCIENTIFIC BOARD

Adnene Ajimi (University of Sousse, Tunisia)

Heni Boubaker (University of Sousse, Tunisia)

Vasily Dikussar (Doradnicyn Computing Centre of the Russian Academy of Sciences, Russia)

Peter Friedrich (University of Tartu, Estonia)

Paolo Gajo (University of Florence, Italy)

Agnieszka Gehringer (University of Göttingen, Germany)

Anna Maria Gil-Lafuente (University of Barcelona, Spain)

Jaime Gil-Lafuente (University of Barcelona, Spain)

Vasile Glavan (Moldova State University, Moldova)

Francesca Greselin (The University of Milano-Bicocca, Italy)

Ana Kapaj (Agriculture University of Tirana, Albania)

Jirawan Kitchaicharoen (Chiang Mai University, Thailand)

Yuriy Kondratenko (Black Sea State University, Ukraine)

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

Karol Kukula (University of Agriculture in Krakow, Poland)

Kesra Nermend (University of Szczecin, Poland)

Nikolas N. Olenev (Doradnicyn Computing Centre of the Russian Academy of Sciences, Russia)

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

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

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

Marina Z. Solesvik (Nord University, Norway)

Ewa Syczewska (Warsaw School of Economics, Poland)

Achille Vernizzi (University of Milan, Italy)

Andrzej Wiatrak (University of Warsaw, Poland)

Dorota Witkowska (University of Lodz, Poland)

ISSN 2082-792X

e-ISSN 2543-8565

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

Warsaw 2020, Volume XXI, No. 1

The original version is the paper version

Journal homepage: qme.sggw.pl

Published by Warsaw University of Life Sciences Press

CONTENTS

Aleksandra Dorosz, Hanna Dudek – Spożycie owoców w Polsce – Analiza na podstawie jednorównaniowych nieliniowych modeli ekonometrycznych	1
Bernard Kokczyński, Dorota Witkowska – Identyfikacja czynników wpływających na ocenę ryzyka kredytowego mikroprzedsiębiorstw działających w regionie łódzkim.....	11
Kacper Pacztkowski, Konrad Furmańczyk – Analiza debiutów na Rynku Głównym GPW w latach 2005-2019	20
Jacek Wolak – Analiza zmian poziomu życia w powiatach województwa małopolskiego	30
Marta Wójcik – Analiza porównawcza stanu służby zdrowia w województwach	39
Grzegorz M. Malinowski – Niepewność nauki a działanie, czyli – problemy z Evidence – Based Policy	49

SPOŻYCIE OWOCÓW W POLSCE – ANALIZA NA PODSTAWIE JEDNORÓWNIOWYCH NIELINIOWYCH MODELI EKONOMETRYCZNYCH

Aleksandra Dorosz

Hanna Dudek  <https://orcid.org/0000-0001-8261-2745>

Katedra Ekonometrii i Statystyki

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mail: aleksandra.dorosz@wp.pl; hanna_dudek@sggw.edu.pl

Streszczenie: W pracy podjęto się oceny dopasowania typowych nieliniowych modeli popytu na owoce w zależności od dochodów. W ekonometrycznej analizie wykorzystano funkcję potęgową, wykładniczą z odwrotnością i Törnquista dla dóbr pierwszej potrzeby. Modele oszacowano na podstawie danych z grup kwintylowych z lat 2013-2018 publikowanych przez Główny Urząd Statystyczny. Na podstawie oszacowanych parametrów modeli wyznaczono dochodowe elastyczności popytu. Stwierdzono, że modele opisane przez funkcję potęgową oraz funkcję Törnquista dla dóbr pierwszej potrzeby charakteryzowały się lepszym dopasowaniem do danych niż modele wykorzystujące funkcję wykładniczą z odwrotnością.

Słowa kluczowe: spożycie, owoce, gospodarstwa domowe, modele nieliniowe

JEL classification: C20, D12

WPROWADZENIE

We współczesnej literaturze dużą uwagę zwraca się na analizę stopnia zaspokojenia potrzeb materialnych gospodarstw domowych [Mikuła 2011; Dudek 2012; Sompolska-Rzechuła 2013; Kozera i in. 2014; Utzig 2016]. Zagadnienie to ściśle łączy się z problematyką konsumpcji. Ważnym bowiem elementem zachowań konsumenckich jest nabywanie środków zaspokajających różnorakie potrzeby.

<https://doi.org/10.22630/MIBE.2020.21.1.1>

W literaturze wiele uwagi poświęcono zarówno teoretycznym aspektom konsumpcji [Bywalec 2012; Zalega 2012], jak i empirycznym analizom dotyczącym modelowania zachowań konsumpcyjnych Polaków [Dudek 2008; Kurzawa, Wysocki 2009; Stanisławska, Wysocki 2011; Chudzian, Chrzanowska 2014; Gostkowski i in. 2014; Głowicka-Wołoszyn i in. 2016; Wołoszyn i in. 2016; Bąk 2017; Olewnicki i in. 2019]. W szczególności, wielu autorów podejmowało temat konsumpcji żywności w Polsce [Borowska i in. 2002; Gulbicka, Kwasek 2006; Dudek 2011; Szwacka-Mokrzycka 2019]. Żywność stanowi bowiem szczególnie ważną kategorię dóbr, zaspokajających elementarne potrzeby człowieka. Dlatego ta kategoria wymaga aktualnych badań i analiz. Temat ten podjęto także w niniejszym opracowaniu koncentrując się na spożyciu owoców. W pracy skupiono się na aspektach modelowania konsumpcji owoców w zależności od dochodów. W modelowaniu ekonometrycznym wykorzystano typowe podejście wykorzystujące funkcje: potęgową, Törnquista pierwszego rodzaju oraz wykładniczą z odwrotnością.

Celem pracy jest ocena dopasowania tych typowych nieliniowych modeli jednorodnaniowych w analizie popytu na owoce w Polsce. Modele te oszacowano na podstawie danych z lat 2013-2018 dotyczących grup kwintylowych. Ponadto, na podstawie analizy oszacowań dochodowych elastyczności popytu, w pracy podjęto się oceny intensywności reakcji gospodarstw domowych na zmiany dochodów.

CHARAKTERYSTYKA DANYCH NA TEMAT SPOŻYCIA OWOCÓW W POLSCE

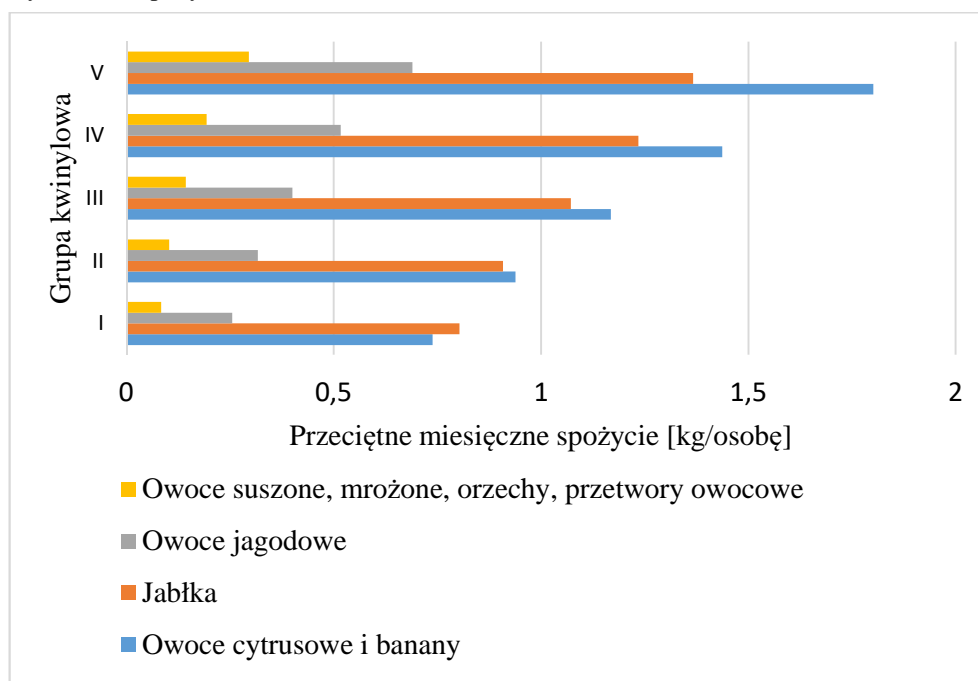
Jako materiał do analizy popytu konsumpcyjnego posłużono się danymi pochodzącymi z badań budżetów gospodarstw domowych, które są prowadzone przez Główny Urząd Statystyczny (GUS).

Dane wykorzystane do badania to wtórne dane GUS opublikowane w corocznych opracowaniach [Badania Budżety Gospodarstw Domowych 2014-2019]. W pracy skupiono się na informacjach na temat:

- 1) przeciętnych miesięcznych dochodów rozporządzalnych na 1 osobę w gospodarstwach domowych według grup kwintylowych oraz
- 2) przeciętnego miesięcznego spożycia wybranych grup owoców na 1 osobę w gospodarstwach domowych według grup kwintylowych, przy czym rozpatrzono:
 - owoce cytrusowe i banany,
 - jabłka,
 - owoce jagodowe,
 - owoce suszone, mrożone, orzechy i przetwory owocowe.

Na rysunku 1 przedstawiono informację na temat przeciętnego miesięcznego spożycia owoców w Polsce w latach 2013-2018 w grupach kwintylowych.

Rysunek 1. Spożycie owoców w Polsce w latach 2013-2018



Źródło: opracowanie własne na podstawie danych z [Badania Budżetów Gospodarstw Domowych 2014-2019]

Z rysunku 1 można w szczególności odczytać, że w gospodarstwach domowych, których miesięczne dochody rozporządzone plasowały się w najwyższej grupie kwintylowej spożywano w przeliczeniu na 1 osobę miesięcznie przeciętnie ok. 1,8 kg owoców cytrusowych i bananów, prawie 1,4 kg jabłek, ok. 0,7 kg owoców jagodowych oraz ok. 0,3 kg owoców suszonych, mrożonych, orzechów i przetworów owocowych. Z kolei gospodarstwa z najbiedniejszej grupy charakteryzowały się znacznie niższą konsumpcją wszystkich rozpatrywanych typów owoców. Z informacji przedstawionych na rys. 1 wynika zatem, że konsumpcja owoców była silnie uzależniona od dochodów. Dlatego też, w pracy rozpatrzono modele opisujące tę zależność.

JEDNORÓWNIOWE NIELINIOWE MODELE POPYTU NA ŻYWNOSĆ

W badaniach ekonometrycznych popytu na żywność najczęściej stosuje się następujące funkcje [Kukuła 2009; Kubicová, Lušňáková 2010]:

- potęgowa: $f(x) = ax^\beta$, (1)
- Törnquista dla dóbr pierwszej potrzeby: $f(x) = \frac{ax}{\beta+x}$, (2)

- wykładniczą z odwrotnością: $f(x) = \exp\left(\alpha + \beta \frac{1}{x}\right)$ (3)

W pracy rozpatrzono możliwość wykorzystania każdej z wymienionych funkcji. Warto nadmienić, że ważną własnością funkcji potęgowej jest stała elastyczność funkcji względem zmiennej niezależnej, tzn.

$$E(x) = f'(x) \cdot \frac{x}{f(x)} = \beta \quad (4)$$

Parametr β informuje zatem o procentowej zmianie $f(x)$ spowodowanej wzrostem zmiennej x o 1%. Elastyczności funkcji Törnquista dla dóbr pierwszej potrzeby oraz funkcji wykładniczej z odwrotnością zależą natomiast od poziomu zmiennej niezależnej, wynosząc odpowiednio:

$$E(x) = \frac{\beta}{\beta+x} \quad (5)$$

$$E(x) = -\frac{\beta}{x} \quad (6)$$

Ponadto, inaczej niż w wypadku funkcji potęgowej, parametr β nie ma tu bezpośredniej interpretacji, natomiast parametr α w funkcji Törnquista oraz $\exp(\alpha)$ w funkcji wykładniczej z odwrotnością oznaczają tzw. poziom nasylenia.

W pracy oszacowano trzy typy modeli:

$$y = \alpha x^\beta + \varepsilon \quad (7)$$

$$y = \frac{\alpha x}{\beta+x} + \varepsilon \quad (8)$$

$$y = \exp\left(\alpha + \beta \frac{1}{x}\right) + \varepsilon \quad (9)$$

gdzie:

x – miesięczny realny dochód rozporządzalnych na osobę¹,

y – miesięczne spożycie owoców w kg na osobę,

α, β – parametry modelu podlegające estymacji,

ε – składnik losowy.

Estymację przeprowadzono w programie Gretl wykorzystując nieliniową metodę najmniejszych kwadratów [Cotrell, Lucchetti 2020]. W procesie optymalizacji nieliniowej zastosowano iteracyjny algorytm Levenberga–Marquardta.

WYNIKI I DYSKUSJA

W tabeli 1 przedstawiono oszacowania parametrów modeli, przy czym a i b oznaczają odpowiednio wartości estymatorów parametrów α i β , $S(a)$ i $S(b)$ – odpowiednio standardowe błędy szacunku parametrów α i β , R^2 – współczynnik determinacji.

¹ W celu uzyskania porównywalności danych z lat 2013-2018 dochody urealniono wyrażając je w cenach stałych z 2018 r.

Tabela 1. Wyniki estymacji modeli popytu na owoce

Grupa owoców	Model opisany funkcją:		
	potęgowa	Törnquista dla dóbr pierwszej potrzeby	wykładniczą z odwrotnością
owoce cytrusowe i banany	$a=0,025$ $b=0,537$ $S(a)= 0,003$ $S(b)= 0,017$ $R^2=0,975$	$a=2,900$ $b=1795,470$ $S(a)= 0,111$ $S(b)= 132,637$ $R^2 = 0,976$	$a=0,778$ $b=-692,338$ $S(a)= 0,034$ $S(b)= 45,398$ $R^2 = 0,921$
jabłka	$a=0,128$ $b=0,296$ $S(a)= 0,033$ $S(b)= 0,035$ $R^2=0,724$	$a=1,574$ $b=547,623$ $S(a)= 0,090$ $S(b)= 106,239$ $R^2=0,699$	$a=0,367$ $b=328,154$ $S(a)= 0,044$ $S(b)= 48,573$ $R^2=0,654$
owoce jagodowe	$a=0,006$ $b=0,593$ $S(a)=0,002$ $S(b)=0,043$ $R^2=0,885$	$a=1,221$ $b=2397,870$ $S(a)=0,140$ $S(b)=477,983$ $R^2=0,877$	$a=-0,183$ $b=-787,336$ $S(a)=0,061$ $S(b)=86,277$ $R^2= 0,804$
owoce suszone, mrożone, orzechy, przetwory owocowe	$a=0,0004$ $b=0,830$ $S(a)=0,0001$ $S(b)=0,025$ $R^2=0,979$	$a=1,206$ $b=9083,990$ $S(a)=0,192$ $S(b)=1788,600$ $R^2=0,978$	$a=-0,839$ $b=-1282,060$ $S(a)=0,058$ $S(b)=95,411$ $R^2=0,804$

Źródło: obliczenia własne na podstawie danych GUS

Analizując wartości współczynnika determinacji można zauważyć, że zdecydowanie najslabszym dopasowaniem do danych charakteryzował się model opisany funkcją wykładniczą z odwrotnością. Jakość dopasowania pozostałych dwóch typów modeli była zbliżona, zwłaszcza w odniesieniu do spożycia owoców cytrusowych i bananów oraz owoców suszonych, mrożonych, orzechów i przetworów owocowych.

Mając na uwadze bardzo dobre dopasowanie modeli opisanych funkcją potęgowa można wyciągnąć wnioski na temat dochodowej elastyczności popytu na daną grupę owoców. Na podstawie analizy oszacowań parametru β można stwierdzić, że najniższa elastyczność dochodowa odnosiła się do jabłek ($E(x) \approx 0,3$), następnie kolejno do owoców cytrusowych i bananów ($E(x) \approx 0,5$), owoców jagodowych ($E(x) \approx 0,6$) oraz owoców suszonych, mrożonych, orzechów i

przetworów owocowych ($E(x) \approx 0,8$)². Elastyczności te są miernikami intensywności reakcji konsumenta na zmiany dochodów. Analizując uzyskane wyniki można sądzić, że jabłka stanowiły najbardziej podstawową grupę owoców, gdyż produkty zajmujące ważną i stałą pozycję w codziennej racji żywnościowej wykazują mniejszą wrażliwość na zmiany dochodów [por. Kurzawa, Wysocki 2009].

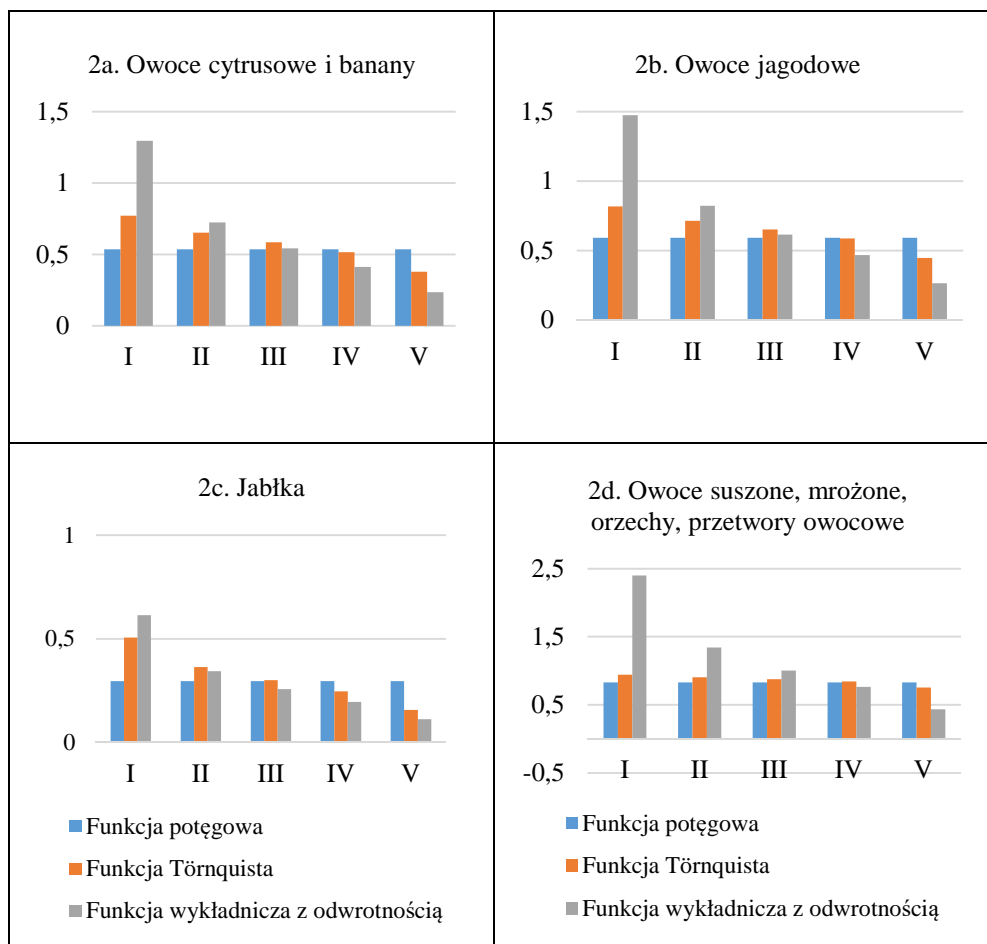
W celu wyznaczenia elastyczności w różnych grupach dochodowych wykorzystano wzory (5)-(6) odnoszące się do funkcji Törnquista dla dóbr pierwszej potrzeby oraz funkcji wykładniczej z odwrotnością. Wyniki w tym zakresie przedstawiono na rysunku 2.

Z informacji przedstawionych na rysunku 2 wynika, że elastyczność wyznaczona na podstawie funkcji potęgowej w przybliżeniu odpowiada elastycznościom w trzeciej lub czwartej grupie kwintylowej oszacowanym z wykorzystaniem funkcji Törnquista dla dóbr pierwszej potrzeby oraz funkcji wykładniczej z odwrotnością. Można zatem z pewną ostrożnością interpretować wyniki uzyskane z wykorzystaniem funkcji potęgowej jako reakcję popytu na relatywną zmianę przeciętnych dochodów gospodarstw domowych.

Należy podkreślić, że elastyczności wyznaczone na podstawie funkcji wykładniczej z odwrotnością charakteryzują się znacznie większą zmiennością niż elastyczności otrzymane na podstawie funkcji Törnquista dla dóbr pierwszej potrzeby. W szczególności, dla grupy owoców suszonych, mrożonych, orzechów i przetworów owocowych, biorąc pod uwagę pierwszą z wymienionych funkcji, elastyczności dochodowe zmieniały się od 2,43 (dla pierwszej grupy kwintylowej) do 0,43 (dla piątej grupy kwintylowej), natomiast wykorzystując drugą z funkcji – zakres zmienności wyniósł [0,75; 0,94]. Jednakże, niezależnie od wykorzystanego modelu, najniższe wartości dochodowej elastyczności odnoszą się do jabłek, natomiast najwyższe – do owoców suszonych, mrożonych, orzechów i przetworów owocowych.

² To oznacza, że jednoprocenowy wzrost dochodów powodował zwiększenie konsumpcji jabłek o 0,3%, owoców cytrusowych i bananów o 0,5%, owoców jagodowych o 0,6% oraz owoców suszonych, mrożonych, orzechów i przetworów owocowych o 0,8%.

Rysunek 2. Elastyczności dochodowe spożycia owoców wg grup dochodowych



Źródło: opracowanie własne

Uzyskane wyniki są w dużym stopniu zbieżne z rezultatami uzyskanymi przez innych autorów. W szczególności, zgodnie z wynikami przedstawionymi w opracowaniu [Gulbicka, Kwasek 2006] potwierdzono, że spożycie owoców zwiększało się wraz ze wzrostem dochodów gospodarstw domowych. Ponadto, podobnie jak w pracach [Borowska i in. 2002; Kurzawa, Wysocki 2009] stwierdzono nieco lepsze dopasowanie modeli opisanych przez funkcję Törnquista dla dóbr pierwszej potrzeby niż modeli wykorzystujących funkcję wykładniczą z odwrotnością.

PODSUMOWANIE

Na podstawie uzyskanych wyników można sądzić, że wszystkie rozpatrywane modele charakteryzowały się dość dobrym dopasowaniem do danych. Jednakże, modele opisane funkcją wykładniczą z odwrotnością „sprawdziły się” w tym względzie nieco gorzej niż modele opisane funkcją potęgową i Törnquista pierwszego rodzaju.

Stwierdzono, że niezależnie od wykorzystanego modelu, najniższe wartości dochodowej elastyczności dotyczą jabłek, natomiast najwyższe – owoców suszonych, mrożonych, orzechów i przetworów owocowych. Można zatem wnioskować, że jabłka okazały się najbardziej podstawowym dobrem spośród analizowanych grup owoców, za to na przeciwległym biegunie znalazły się owoce suszone, mrożone, orzechy i przetwory owocowe.

BIBLIOGRAFIA

- Borowska A., Dudek H., Szczesny W. (2002) Dobór postaci analitycznej i metod estymacji modeli zależności wydatków na żywność od dochodów. *Wiadomości Statystyczne*, 7, 20-29.
- Badania Budżetów Gospodarstw Domowych (2014-2019) Główny Urząd Statystyczny, Warszawa.
- Bąk I. (2017) Statystyczna analiza wydatków żywnościowych gospodarstw domowych 50+. *Śląski Przegląd Statystyczny*, 15(21), 251-264.
- Bywalec C. (2012) *Ekonomia i finanse gospodarstw domowych*. Wydawnictwo Naukowe PWN, Wyd. 2, Warszawa.
- Chudzian J., Chrzanowska M. (2014) Parametric and Non-parametric Regression Methods in Identifying an Impact of Components of Advertising on Consumers Behaviour. *Ekonometria*, 3(45), 56-70.
- Cotrell A., Lucchetti R. (2020) *Gretl User's Guide*. Dokument elektroniczny pobrany z <http://ricardo.ecn.wfu.edu/pub/gretl/gretl-guide.pdf> [dostęp w dniu 13.08.2020]
- Dudek H. (2008) Elastyczności cenowe popytu na żywność - analiza na podstawie modelu LA/AIDS. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, X(4), 62-67.
- Dudek H. (2011) Quantitative Analysis of the Household's Expenditure for Food. *Problems of World Agriculture*, 26, 23-30.
- Dudek H. (2012) Subiektywne skale ekwiwalentności - analiza na podstawie danych o satysfakcji z osiągniętych dochodów. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 232, 153-162.
- Głowicka-Wołoszyn R., Kurzawa I., Wołoszyn A. (2016) Ekonometryczne modelowanie popytu na turystykę zorganizowaną. *Zeszyty Naukowe Uniwersytetu Szczecińskiego, Ekonomiczne Problemy Turystyki*, 1(33), 43-53.
- Gostkowski M., Gajowniczek K., Jałowiecki P. (2014) Elastyczność dochodowa popytu na poszczególne grupy dóbr konsumpcyjnych w Polsce w latach 1999-2008. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, XVI(3), 87-91.

- Gulbicka B., Kwasek M. (2006) Wpływ dochodów na spożycie żywności - przesłanki dla polityki żywnościowej. *Zagadnienia Ekonomiki Rolnej*, 1, 19-33.
- Kozera A., Głowicka-Wołoszyn R., Stanisławska J. (2014) Niedobory konsumpcji w gospodarstwach domowych rolników po wstąpieniu Polski do Unii Europejskiej. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, 16(6), 274-280.
- Kubicová L., Lušňáková Z. (2010) Consumer foodstuffs demand and income standard development in the households of Slovakia. *Acta Univ. Agric. et Silv. Mendel. Brun. LVIII(3)*, 99-106.
- Kukuła K. (red.) (2009) *Wprowadzenie do ekonometrii*. Wydawnictwo Naukowe PWN, Warszawa.
- Kurzawa I., Wysocki F. (2009) Wybrane modele ekonometryczne w badaniach dochodowej elastyczności popytu konsumpcyjnego. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, Taksonomia* 16, 47, 70-78.
- Mikuła A. (2011) Poziom ubóstwa w Polsce w ujęciu regionalnym. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, XIII(3), 199-203.
- Olewnicki D., Jabłońska L., Dudek H. (2019) The demand for ornamental plants in Poland after its integration into the EU: a quantitative approach. *Bulgarian Journal of Agricultural Science* 25(5), 932-943.
- Sompolska-Rzechuła (2013) Jakość życia jako kategoria ekonomiczna. *Folia Pomeranae Universitatis Technologiae Stetinensis, Oeconomica*, 301(71), 127-140.
- Stanisławska J., Wysocki F. (2011) Dochodowa elastyczność wydatków na artykuły żywnościowe w gospodarstwach domowych rolników według grup dochodowych. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, 13(3), 315-319.
- Szwacka-Mokrzycka J. (2019) *Paradygmaty rozwoju konsumpcji żywności w Polsce*. Wydawnictwo Naukowe PWN, Warszawa.
- Utzig M. (2016) Zróżnicowanie struktury wydatków konsumpcyjnych gospodarstw domowych w Polsce według grup społeczno-ekonomicznych. *Nierówności Społeczne a Wzrost Gospodarczy*, 47, 454-465.
- Wołoszyn A., Głowicka-Wołoszyn R., Stanisławska J., Wysocki F. (2016) Identyfikacja uwarunkowań zakupów przez Internet przy zastosowaniu uogólnionego modelu tobitowego. *Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania Uniwersytetu Szczecińskiego*, 3(43), 459-469.
- Zalega T. (2012) *Konsumpcja. Determinanty. Teorie. Modele*. Polskie Wydawnictwo Ekonomiczne, Warszawa.

FRUIT CONSUMPTION IN POLAND – ANALYSIS ON THE BASIS OF SINGLE-EQUATION NONLINEAR ECONOMETRIC MODELS

Abstract: The article dealt with an assessment of the fit of typical non-linear demand models for fruit. The power function, the exponential function with the inverse and the Törnquist function for the basic goods were used in the econometric analysis. The models were estimated on the basis of income

quintile groups' data published by the Central Statistical Office. Based on the estimated model parameters, income elasticities of demand were calculated. It was found that the models described by the power function and the Törnquist function for basic goods were characterized by a better goodness of fit than the models using the exponential function with the inverse.

Keywords: consumption, fruit, households, nonlinear models

JEL classification: C20, D12

IDENTYFIKACJA CZYNNIKÓW WPŁYWAJĄCYCH NA OCENĘ RYZYKA KREDYTOWEGO MIKROPRZEDSIĘBIORSTW DZIAŁAJĄCYCH W REGIONIE ŁÓDZKIM

Bernard Kokczyński  <https://orcid.org/0000-0002-9379-0376>

B&B Co. Bernard Kokczyński

e-mail: bernard.jeremi@o2.pl

Dorota Witkowska  <https://orcid.org/0000-0001-9538-9589>

College of Business and Economics

University of Johannesburg

Republika Południowej Afryki

e-mail: mariaw@uj.ac.za

Streszczenie: Mikroprzedsiębiorstwa mają bardzo ograniczone źródła finansowania zewnętrznego, które zapewniane jest głównie przez banki. Celem badań jest identyfikacja czynników wpływających na ocenę ryzyka kredytowego małych firm. Badania prowadzono za pomocą regresji liniowej z wykorzystaniem rzeczywistych danych dotyczących mikroprzedsiębiorstw działających w województwie łódzkim, które uzyskały różne formy kredytowania w latach 2016-2018. Wyniki pokazują, że dochody i lata funkcjonowania na rynku to najważniejsze czynniki wpływające na poprawę wiarygodności kredytowej mikroprzedsiębiorstw.

Słowa kluczowe: mikroprzedsiębiorstwa, ocena ryzyka kredytowego, credit-scoring, regresja liniowa

JEL classification: G24, C21

WPROWADZENIE

Mikroprzedsiębiorstwo to podmiot zatrudniający mniej niż 10 pracowników, którego roczny obrót netto lub suma aktywów jego bilansu sporządzonego na koniec roku nie przekracza 2 mln euro¹.

¹ Ustawa z dnia 02.07.2004 r. o swobodzie działalności gospodarczej. Dziennik Ustaw nr 173 z 2004 r., poz. 1807.

Według [Tarnawa, Zadura-Lichota 2012, 68-85, Bubiak 2013]:

- mikroprzedsiębiorstwa to 95,7% przedsiębiorstw w Polsce;
- obserwowana jest znaczna różnorodność wśród małych firm wynikająca z dużego (tj. 72%) udziału tzw. „przymusowo zatrudnionych”;
- wśród działających mikroprzedsiębiorstw, 19,6% funkcjonuje krócej niż 11 miesięcy, a 9% krócej niż 5 miesięcy;
- systematycznie zwiększa się liczba pracujących w małych firmach, ale wciąż jest to mniej niż średnio jedna osoba na mikroprzedsiębiorstwo;
- wielkość generowanych przychodów zależy od branży w jakiej działa przedsiębiorstwo;
- w 2014 r. mikroprzedsiębiorstwa wygenerowały 41,1% polskiego PKB, wytworzonego przez wszystkie przedsiębiorstwa².

Jak zatem widać, te małe firmy odgrywają istotną rolę w polskiej gospodarce, a jednocześnie nie mają one tak wielu możliwości pozyskiwania środków finansowych na rozwój jak większe przedsiębiorstwa i głównym źródłem kapitału zewnętrznego są produkty bankowe dedykowane jednostkom gospodarczym działającym według różnych form prawnych.

Udzielanie kredytów wiąże się z ryzykiem, stąd badanie zdolności kredytowej potencjalnych kredytobiorców jest sprawą kluczową, bowiem znaczna liczba kredytów niespłaconych zwiększa rezerwy banku i może spowodować utratę płynności. Konkurencja na rynku finansowym powoduje, że banki łagodzą swoją politykę kredytową, czego konsekwencją jest przyznawanie kredytu jednostkom z podwyższonej grupy ryzyka [por. Chmielewski i in. 2007, s. 5]. Warto zauważyć, że po opublikowaniu przez Bazylejski Komitet ds. Nadzoru Bankowego wytycznych Nowej Umowy Kapitałowej, według których credit-scoring jest jednym z dopuszczalnych narzędzi kalkulacji ryzyka kredytowego, stał się on powszechnie stosowaną metodą oceny zdolności kredytowej [por. Schab 2005, Feruś 2006].

Badanie wiarygodności kredytowej przedsiębiorstw dokonywane jest na podstawie analizy sytuacji ekonomiczno-finansowej prowadzonej zazwyczaj w oparciu o sprawozdania finansowe. W przypadku mikroprzedsiębiorstw, informacje wykorzystywane w analizie sytuacyjnej firm są znacznie bardziej ubogie, a część z nich ma charakter jakościowy, podobnie jak w przypadku badania wiarygodności kredytowej klientów indywidualnych. Wykorzystuje się w tym celu różne metody badawcze, a wśród nich statystyczne metody klasyfikacji, tj. modele regresji binarnej, funkcje dyskryminacyjne, drzewa klasyfikacyjne oraz sztuczne sieci neuronowe [por. Altman 1968, Kuryłek 2000, Witkowska 2002, s.124-181, 2006, Chrzanowska, Witkowska 2007]. Kluczowym zagadnieniem jest więc identyfikacja czynników wpływających na ocenę ryzyka kredytowego.

² <https://twojewirtualnebiuro.pl/pl/mini-przeglad-sektora-msp-w-polsce/>

W związku z tym celem pracy jest identyfikacja czynników wpływających na ocenę scoringową przypisaną przedsiębiorcy przez bank. Analizy przeprowadzono na podstawie rzeczywistych danych dotyczących mikroprzedsiębiorstw działających w regionie łódzkim, które w latach 2016-2018 uzyskały kredyt w jednym z banków³.

OPIS DANYCH

Baza danych zawiera informacje dotyczące 81 mikroprzedsiębiorstw prowadzących działalność gospodarczą głównie na terenie województwa łódzkiego, które wniosły o jeden z trzech produktów kredytowych, tj. kredyt handlowy lub inwestycyjny, kartę kredytową typu biznes lub kredyt obrotowy w rachunku bieżącym. Wszyscy przedsiębiorcy to osoby fizyczne prowadzące działalność gospodarczą, które posiadają ksiązkę przychodów i rozchodów, a z urzędem skarbowym rozliczają się na zasadach ogólnych. Przygotowana baza danych zawiera następujące informacje dotyczące kredytobiorców, pochodzące z wniosków kredytowych:

- charakterystyka działalności, tj. kod PKD, miejscowość, gdzie prowadzona jest działalność gospodarcza oraz data jej rozpoczęcia;
- roczne wyniki finansowe, czyli: uzyskane przychody, poniesione koszty, uzyskany dochód, naliczone składki na ubezpieczenia społeczne oraz podatek dochodowy;
- charakterystyka produktów kredytowych tzn. liczba i rodzaj uzyskanych produktów kredytowych, wolumen uzyskanego finansowania, okres finansowania, klasa scoringowa przypisana kredytobiorcy przez bank oraz informacja na temat możliwego zabezpieczenia kredytu gwarancją bankową.

Należy zaznaczyć, że ocena scoringowa oznacza poziom ryzyka kredytowego przypisany przedsiębiorcy wnioskującemu o produkt kredytowy. Zatem im wyższa ocena scoringowa klienta banku tym mniej prawdopodobne jest terminowe uregulowanie zobowiązania względem banku. Wyższe ryzyko bank wynagradza sobie zwiększoną opłatą przygotowawczą oraz marżą. Oznacza to, że klient o najniższej ocenie scoringowej może liczyć na najtańszy kredyt.

Analizę struktury kredytobiorców dla wyróżnionych cech przedstawiono w tabelach 1-2. Kredytobiorcy zostali podzieleni na cztery klasy ze względu na wielkość miejscowości, w której przedsiębiorstwa prowadzą działalność gospodarczą (tabela 1). Najwięcej badanych przedsiębiorstw, prowadzi działalność gospodarczą w Łodzi (powyżej 100 tys. mieszkańców - 32%). Oni też uzyskali największą kwotę kredytu zarówno ogółem, jak i wartość średnią przyznanego finansowania przypadającą na jednego przedsiębiorcę.

³ Baza danych jest objęta tajemnicą przedsiębiorstwa oraz tajemnicą bankową. W związku z tym nie jest możliwym ujawnienie nazwy ani innych danych banku lub kredytobiorców.

Tabela 1. Struktura kredytobiorców ze względu na wielkość miejscowości, gdzie prowadzona jest działalność gospodarcza

Liczba mieszkańców (tys.)	Kredytobiorcy		Wartość kredytów		
	Liczba	Udział (%)	Ogółem (tys. PLN)	Średnia (tys. PLN)	Udział (%)
0-20	17	21	3 136,0	184,423	22
20-50	14	17	1 773,5	136,423	13
50-100	24	30	3 835,0	159,792	27
100 i więcej	26	32	5 427,0	201,741	38

Źródło: obliczenia własne

Biorąc pod uwagę rodzaj prowadzonej działalności gospodarczej kredytobiorcy zostali podzieleni według PKD na 6 głównych branż (tabela 2). Jak widać, najwięcej tj. 29 (36%) przedsiębiorców działa w handlu detalicznym i to ta branża uzyskała najwyższą łączną kwotę kredytu. Natomiast średnio na jednego kredytobiorcę najwyższe finansowanie przypadło na mikroprzedsiębiorstwa medyczne i zajmujące się handlem hurtowym.

Tabela 2. Struktura kredytobiorców ze względu na branżę

Branża	Kredytobiorcy		Wartość kredytów		
	Liczba	Udział (%)	Ogółem (tys. PLN)	Średnia (tys. PLN)	Udział (%)
produkcja	11	14	1 953,0	177,545	13
handel hurtowy	11	14	2 390,0	217,273	17
handel detaliczny	29	36	4 363,5	150,466	31
medyczna	10	11	2 214,0	221,400	16
usługi	13	16	2 440,0	187,692	17
inne	7	9	811,0	118,714	6

Źródło: opracowanie własne

Warto odnotować, że wśród badanych przedsiębiorców, liczba kobiet i mężczyzn była porównywalna, chociaż udział kobiet był zdecydowanie większy wśród zajmujących się handlem detalicznym (22 na 41 kobiet przedsiębiorców), natomiast najwięcej mężczyzn prowadziło działalność produkcyjną (9 na 40 mężczyzn). Interesująca jest również struktura kredytobiorców z punktu widzenia przeprowadzonej przez bank oceny ryzyka kredytowego, według której wyróżniono trzy klasy kredytobiorców, biorąc pod uwagę liczbę uzyskanych punktów scoringowych⁴:

⁴ W nomenklaturze banku są to tzw. klasy scoringowe.

1. klienci o największym prawdopodobieństwie spłaty (1-10 pkt.),
2. klienci o średnim prawdopodobieństwie spłaty (11-16 pkt.),
3. klienci o największym poziomie ryzyka (17-20 pkt.).

Z przeprowadzonych analiz wynika, że w klasie pierwszej znalazło się tylko siedmiu przedsiębiorców (9%), w drugiej – 31 (38%), a w ostatniej aż 44 (53%). Niewątpliwie jednym z podstawowych kryteriów oceny potencjalnych kredytobiorców jest uzyskiwany przez nich dochód, który mieścił się w przedziale od 7,94 do 304,29 tysięcy PLN, przy medianie równej 59,1. Przeprowadzona analiza pozwoliła zdiagnozować trzy obserwacje jako nietypowe, bowiem nie mieściły się w 3-sigmowym przedziale wokół średniej arytmetycznej. Dane dotyczące tych trzech przedsiębiorców zostały wykluczone z dalszych analiz.

ANALIZA ZALEŻNOŚCI MIĘDZY OCENĄ SCORINGOWĄ A BADANYMI CZYNNIKAMI

Jak można zauważyć, dane zawarte we wniosku kredytowym badanych przedsiębiorców zawierają relatywnie mało informacji. Co więcej dane opisujące mikroprzedsiębiorstwa są o charakterze ilościowym i jakościowym, które mogą być trudne do jednoczesnego uwzględnienia w procedurze identyfikacji czynników, wpływających na ocenę scoringową przedsiębiorcy. W związku z tym w analizach uwzględniono różny charakter zmiennych w modelach regresji, które estymowano MNK na podstawie 78 obserwacji. Innymi słowy oszacowano modele, opisujące liczbę punktów scoringowych przyznanych poszczególnym kredytobiorcom, uwzględniając w nich jedną lub kilka zmiennych objaśniających. W szczególności wyróżniono następujące zmienne objaśniające, opisujące cechy:

- ilościowe: przychody, dochód brutto, dochód netto [mln PLN], długość okresu funkcjonowania na rynku – staż na rynku [lata], wielkość miejscowości prowadzonej działalności gospodarczej, mierzona liczbą mieszkańców [tys.];
- porządkowe: wielkość miejscowości oznaczona jako klasa;
- nominalne: płeć (dwudzielna, zmienna referencyjna: męzczyzna), branża PKD (wielodzielna, zmienna referencyjna: produkcja), klasa miejscowości (wielodzielna, zmienna referencyjna: miejscowość powyżej 100 tys. mieszkańców tj. Łódź).

Uzyskane wyniki zamieszczono w tabelach, przyjmując następujące oznaczenia: β_i – oceny estymatorów parametrów modeli, $t(\beta_i)$ - wartości statystyki t-Studenta, R^2 - współczynnik determinacji, r – współczynnik korelacji liniowej Pearsona. Pogrubioną czcionką oznaczono parametry istotne różne od zera na poziomie istotności $\alpha = 0,05$.

Parametry modeli z jedną zmienną objaśniającą wskazują na niski poziom objaśnienia oceny ryzyka kredytowego wyrażonego w liczbie punktów scoringowych (R^2 znajduje się w przedziale od 0,007 do 0,491). Jednocześnie widoczne jest, że czym wyższe wartości przychodów, dochodów i okresu

działalności firmy na rynku tym ocena scoringowa niższa, a zatem klient oceniany jest przez bank jako bardziej wiarygodny (parametry ujemne i statystycznie istotne). Natomiast kobiety oceniane są jako mniej wiarygodne niż mężczyźni, bowiem parametr stojący przy zmiennej określającej płeć jest dodatni i statystycznie istotny. Można równocześnie odnotować, że wielkość miejscowości – niezależnie od sposobu kodowania tej zmiennej - nie ma istotnego znaczenia przy ocenie ryzyka kredytowego, dlatego w dalszych rozważaniach zmienna ta została pominięta. Oceniając siłę oddziaływania poszczególnych czynników na liczbę punktów scoringowych należy stwierdzić, że najsilniej oddziałuje staż na rynku, a dalej dochody, których wzrost o 10 tys. PLN przyczyni się do spadku liczby przyznanych punktów scoringowych o 0,29 i 0,36 odpowiednio w przypadku dochodów brutto i netto oraz o 0,45 dla dodatkowego roku działalności firmy.

Tabela 3. Oszacowania modeli regresji uwzględniającymi pojedyncze czynniki

Parametry	Przychody	Dochód brutto	Dochód netto	Staż na rynku	
β_0	16,2755	17,8921	17,7229	19,7138	
$t(\beta_0)$	38,6574	35,4719	36,346	37,4294	
β_1	-0,9624	-28,9990	-35,8374	-0,4496	
$t(\beta_1)$	-2,1713	-5,3351	-5,2145	-8,5619	
R^2	0,0563	0,26487	0,2561	0,4910	
r	-0,2373	-0,5147	-0,5060	-0,7007	
	Miejscowość – liczba mieszkańców		Klasa miejscowości	Płeć	
β_0		16,0760	17,1795	14,9737	
$t(\beta_0)$		32,941	17,2648	28,4901	
β_1		-0,9294	-0,486	1,7013	
$t(\beta_1)$		-0,7453	-1,4463	2,3181	
R^2		0,0073	0,0268	0,066	
r		-0,0852	-0,1636	0,257	
Klasa miejscowości kodowana binarnie					
	<i>constant</i>	Klasa 1	Klasa 2	Klasa 3	R^2
β_0	15,48	β_1 1,27	β_2 1,06	β_3 -0,23	0,0365
$t(\beta_0)$	23,21	$t(\beta_1)$ 1,19	$t(\beta_2)$ 0,93	$t(\beta_3)$ -0,24	

Źródło: obliczenia własne

Biorąc pod uwagę branże, w jakich działają analizowane mikroprzedsiębiorstwa, to należy zauważyć istotny wzrost wiarygodności kredytowej wśród przedsiębiorców z branży medycznej i pozostałych w porównaniu z firmami produkcyjnymi (tabela 4). Przy czym poziom objaśnienia liczby przyznanych punktów scoringowych za pomocą tego modelu jest relatywnie wysoki w stosunku do pozostałych analizowanych czynników.

Tabela 4. Oszacowanie modelu uwzględniającego branżę gospodarczą

Parametry	<i>constant</i>	PKD 2	PKD 3	PKD 4	PKD 5	PKD 6	R ²
β	16,3636	-0,4546	1,0157	-4,3636	-0,5175	-2,9351	0,3008
t(β)	19,1294	-0,3757	1,011	-3,5201	-0,4452	-2,1397	

Źródło: opracowanie własne

Dalsze analizy prowadzono na podstawie modelu, w którym uwzględniano jednocześnie wszystkie czynniki, które w poprzednio analizowanych modelach wykazywały statystycznie istotny wpływ na zmienną objaśnianą (tabela 5). Wyniki przedstawione w tabeli 5 jednoznacznie potwierdzają wcześniejsze wnioski wskazując na istotny wpływ okresu funkcjonowania firmy na rynku i dochodu netto, których wzrost przyczynia się do zmniejszenia oceny ryzyka kredytowego. Podobnie jak w poprzednio omawianym modelu, działalność w branży medycznej i pozostałych oceniana jest jako mniej ryzykowna niż w przemyśle. Natomiast płeć przestała mieć istotnie znaczenie na ocenę wiarygodności kredytowej przedsiębiorcy.

Tabela 5. Oszacowanie modelu uwzględniającego wszystkie czynniki

Parametry	PKD 2	PKD 3	PKD 4	PKD 5	PKD 6
β	-1,1524	-0,6091	-2,8913	-1,2089	-3,5229
t(β)	-1,2451	-0,7412	-2,8270	-1,3593	-3,3359
	Płeć	Staż na rynku	Dochód netto	<i>constant</i>	R ²
β	0,4860	-0,3589	-15,8049	20,7172	0,6768
t(β)	0,9552	-7,1146	-1,9287	25,4192	

Źródło: opracowanie własne

PODSUMOWANIE

Niniejsza praca poświęcona została identyfikacji czynników wpływających na ocenę ryzyka kredytowego mikroprzedsiębiorstw prowadzących działalność w regionie łódzkim, z których jedna trzecia działa na terenie Łodzi. W badaniu uwzględniono 78 firm, które uzyskały kredyt (w różnej formie) w kilku oddziałach pewnego banku działających w województwie. Najwięcej mikroprzedsiębiorstw (ponad jedna trzecia) to firmy zajmujące się handlem detalicznym. Pozostałe badane jednostki gospodarcze (w zbliżonej liczbie firm) prowadzą działalność produkcyjną, usługową i handel hurtowy, a także zostały zaklasyfikowane do branży medycznej i „pozostałe”. Wśród właścicieli badanych firm było 40 kobiet i 38 mężczyzn. Średni uzyskany dochód dla badanej grupy mikroprzedsiębiorstw wyniósł 72 tys. PLN. Natomiast ryzyko kredytowe ponad połowy badanych kredytobiorców oceniono w przedziale 17-20 punktów scoringowych, co oznacza najwyższy możliwy poziom ryzyka, a tylko siedmiu przedsiębiorców znalazło się

w klasie o najmniejszym ryzyku kredytowym (tj. w przedziale między 1, a 10 punktów scoringowych).

Na podstawie przeprowadzonych analiz regresji stwierdzono, że największy statystycznie istotny wpływ na ryzyko kredytowe ma okres funkcjonowania firmy i uzyskiwany przez nią dochód, których wzrost powoduje lepszą ocenę wiarygodności kredytobiorcy. Istotna jest też branża w jakiej działa mikroprzedsiębiorstwo, stwierdzono bowiem statystycznie istotne obniżenie ryzyka kredytowego w branży medycznej i branży sklasyfikowanej jako: pozostałe w stosunku do przedsiębiorstw o charakterze przemysłowym. W przypadku badania wpływu płci właściciela firmy, można zaobserwować, że kobiety oceniane są jako mniej wiarygodne niż mężczyźni, ale ten czynnik okazał się statystycznie istotny jedynie w modelu, w którym uwzględniono jedynie tę zmienną. Jednocześnie stwierdzono brak wpływu lokalizacji mikroprzedsiębiorstw na ocenę ryzyka kredytowego.

Zamrożenie działalności gospodarczej (great lockdown) spowodowane pandemią istotnie wpłynęło na funkcjonowanie mikro- i małych przedsiębiorstw. Wnukowski i Wąsiński (2020) twierdzą, że w najbliższej przyszłości dojdzie do znaczącego zmniejszenia się konsumpcji i recesji, czego skutkiem będzie załamanie się światowej gospodarki. Aktualna sytuacja gospodarcza w Polsce (mimo rządowej pomocy) daje słabe podstawy do przetrwania, zwłaszcza tym przedsiębiorstwom, które nie posiadają własnego kapitału. Chociaż są takie branże, które odczują znaczny wzrost obrotów np. branże związane z usługami zdalnymi, firmy kurierskie czy medyczne.

BIBLIOGRAFIA

- Altman E. (1968) Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23(4), 589-603.
- Bubiak B. (2013) Rola nauk o zarządzaniu w rozwoju mikroprzedsiębiorstw. *Przedsiębiorczość i Zarządzanie*, 14(12/1), 287-298.
- Chmielewski T., Głogowski A. (red), Gołajewska M. (red), Grabowski Sz., Grodzicki M., Hałaj G., Imielska M., Kozak S., Maliszewski K., Tymoczko D. (red), Zajączkowski S. (2007) Przegląd stabilności systemu finansowego I półrocze 2007. NBP, Warszawa.
- Chrzanowska M., Witkowska D. (2007) Zastosowanie wybranych metod klasyfikacji do rozpoznawania indywidualnych kredytobiorców, [w:] Jajuga K., Walesiak M. (red.) *Taksonomia 14, Klasyfikacja i analiza danych – teoria i zastosowania*, 108-114.
- Feruś A. (2006) Zastosowanie metody DEA do określenia poziomu ryzyka kredytowego przedsiębiorstw. *Bank i Kredyt*, 7, 44-59.
- Kuryłek W. (2000) Credit scoring – podejście statystyczne. *Bank i Kredyt*, 6, 72-77.

- Schab I. (2005) Ocena ryzyka kredytowego w ramach wewnętrznych systemów ratingowych – charakterystyka podejścia oraz podstawowych wymogów. *Bezpieczny Bank*, 1 (26), 89-102.
- Tarnawa A., Zadura-Lichota P. (red.) (2012) Raport o stanie sektora małych i średnich przedsiębiorstw w latach 2010–2011. PARP, Warszawa.
- Witkowska D. (2002) Sztuczne sieci neuronowe i metody statystyczne. Wybrane zagadnienia finansowe. Wydawnictwo C.H. Beck, Warszawa.
- Witkowska D. (2006) Discrete Choice Model Application to the Credit Risk Evaluation. *International Advances in Economic Research*, 12, 33-42.
- Wnukowski D., Wąsiński M. (2020) The Great Lockdown: Impact of the COVID-19 Pandemic on the Global Economy. *Bulletin PISM*, 84 (1514).
<https://twojewirtualnebiuro.pl/pl/mini-przeglad-sektora-msp-w-polsce/>
[dostęp 10.08.2020].

IDENTIFICATION OF FACTORS INFLUENCING CREDIT RISK EVALUATION OF MICROENTERPRISES FUNCTIONING IN LODZ REGION

Abstract: Microenterprises have limited sources of external financing that is provided mostly by banks. The aim of our research is to identify factors influencing credit risk evaluation of small firms. Investigation is provided applying linear regression using real data concerning microenterprises functioning in Lodz region which obtained different form of credits in years 2016-2018. The results show that earnings and years of operating on the market are the most important factors which affect the improvement of credibility.

Keywords: microenterprises, credit risk evaluation, credit scoring, linear regression

JEL classification: G24, C21

ANALIZA DEBIUTÓW NA RYNKU GŁÓWNYM GPW W LATACH 2005-2019

Kacper Pacutkowski  <https://orcid.org/0000-0001-7408-6060>

Konrad Furmańczyk  <https://orcid.org/0000-0002-7683-4787>

Instytut Informatyki Technicznej

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mail: kacper_pacutkowski@sggw.edu.pl; konrad_furmanczyk@sggw.edu.pl

Streszczenie: Niniejsze opracowanie traktuje o debiutach giełdowych na Rynku Głównym GPW w latach 2005 – 2019. W tym kontekście przywołano na początku informacje dotyczące tematyki debiutu oraz zagadnień procesowo- organizacyjnych z tym związanych. W dalszej części zaprezentowano przegląd danych udostępnianych w serwisach GPW dotyczących badanej dziedziny problemowej oraz dokonano na ich podstawie analizy wybranych statystyk ilościowych. Ważną część poświęcono zbadaniu scenariuszy zmiany ceny w pierwszym dniu notowań i przeprowadzeniu opartego na nich wnioskowania statystycznego.

Słowa kluczowe: debiut giełdowy, Giełda Papierów Wartościowych, GPW, stopa zwrotu, rynek pierwotny akcji, efektywność inwestycji

JEL classification: C02, C12, G10, G23

WSTĘP

Rozważania dotyczące inwestycji na rynku kapitałowym są istotnym elementem badań wielu dziedzin nauki. Tematyka ta jest obecna w opracowaniach związanych z metodami ilościowymi, a w szczególności z prognozowaniem, inżynierią finansową oraz teoriami portfelowymi. Co oczywiste – jednymi z najczęściej podejmowanych tematów są te dotyczące lub obejmujące elementy odpowiedzi na kilka kluczowych pytań: „W co inwestować?”, „W jaki sposób inwestować?”, „Kiedy inwestować?”, czy w końcu: „Czym kierować się przy podejmowaniu decyzji inwestycyjnych?”. Naturalnym jest, że te pytania mają

<https://doi.org/10.22630/MIBE.2020.21.1.3>

charakter bardzo otwarty i odpowiedzi na nie są kwestią wyznaczenia generycznych odpowiedzi. Zależą one wszakże bardzo silnie od szeregu różnorodnych czynników – począwszy od danej sytuacji rynkowej a na profilu inwestora skończywszy.

Odwracając jednak nieco jedną z tych fundamentalnych kwestii, możemy zapytać: „W co inwestują inwestorzy?”. Co jasne - na tak postawione pytanie odpowiedź jest jedynie sprawą przytoczenia odpowiedniego faktu, opartego między innymi na badaniu obrotów i danych zagregowanych o rachunkach maklerskich. Otóż, najpopularniejszym na świecie instrumentem finansowym bezsprzecznie są akcje [Krzywdą 2007]. Z pewnością nie bez wpływu pozostaje to, że są one instrumentami stosunkowo prostymi w konstrukcji i szeroko dostępnymi do obrotu także dla inwestorów indywidualnych.

W kontekście walorów jakimi są akcje jedną z ważniejszych spraw wartych bliższemu przyjrzeniu jest to, w jaki sposób trafiają one do obrotu oraz to, jakie uwagi, związane z tym procesem, można formułować. W związku z tym, w niniejszym opracowaniu podjęto się przybliżenia tematyki debiutu giełdowego i analizy dostępnych danych w tym zakresie z dotychczasowej działalności Giełdy Papierów Wartościowych. Przeprowadzone działania miały na celu w szczególności: ilościowe scharakteryzowanie debiutu giełdowego w zadanym okresie badawczym oraz ocenę zjawiska debiutu w kontekście potencjalnego zysku dla inwestora. Publikacja składa się z dwóch części. W początkowej części pracy przedstawiono charakterystykę badanej dziedziny problemowej, kolejne natomiast stanowią rozważania metodologiczno-aplikacyjne. W podsumowaniu, poza wnioskami płynącymi z przeprowadzonej analizy, zawarto propozycje kolejnych działań badawczych dotyczących poruszonej tematyki. Zawarte w treści publikacji rozważania, analizy i wnioski w większości pochodzą z badań podjętych w pracy dyplomowej [Pacutkowski 2020].

CHARAKTERYSTYKA RYNKU GIEŁDOWEGO AKCJI

Pojęcie akcji jako udziału w spółce

Przedmiotem dalszych rozważań będą akcje – jedne z walorów podlegających obrotowi na giełdach papierów wartościowych. Poprzez pojęcie „papieru wartościowego” rozumiemy dokument ściśle wyrażający prawo majątkowe jego posiadacza [Górniak 1995]. Termin „akcja” pochodzi od łacińskiego słowa „actio”, które w bezpośrednim tłumaczeniu oznacza działanie lub czynność. Znaczenie tego słowa jest odzwierciedlone wyraźnie w charakterze waloru – co obrazowo zaprezentowano na rysunku 1. Wyróżnikiem tego, że dany papier wartościowy określamy mianem akcji jest jego wskazanie na potwierdzenie udziału okaziciela w kapitale zakładowym konkretnej spółki. Co więcej – za posiadaniem akcji idzie prawo decydowania o spółce – poprzez głosowanie w walnym zgromadzeniu akcjonariuszy. Innym aspektem jest uczestnictwo w zyskach spółki realizowane w postaci dywidendy (łac. *divido* – podzielić), która jest wypłatą z zysku spółki.

Rysunek 1. Charakterystyki akcji



Źródło: opracowanie własne na podstawie [Dębski 2014]

Pierwszy na świecie debiut papierów udziałowych odbył się w XVII w. na giełdzie w Amsterdamie. Zaczęto wówczas notować udziały m.in. Holenderskiej Kompanii Wschodnioindyjskiej [Czerniawski 1992]. Za pierwszy polski akcent tego typu wymienia się natomiast emisję akcji przez Kompanię Manufaktur Wełnianych w XVIII wieku [Zientara i in. 1973].

Cena akcji

Na rynku kapitałowym doskonale działają i widoczne są reguły gry podaży i popytu. W tym kontekście cenę postrzegamy odpowiednio jako wartość pieniężną za jaką spółka jest gotowa oddać część swoich udziałów inwestorom (cena emisyjna) lub jako wartość pieniężną za jaką jeden inwestor jest gotów sprzedać innemu posiadany przez siebie papier wartościowy (cena rynkowa).

Inną klasyfikacją cen jest nomenklatura podawana przez okologiczne serwisy informacyjne. W swoich zestawieniach prezentują one następujące wartości:

- cena otwarcia – cena zawarcia pierwszej transakcji na akcjach danej spółki w danej sesji giełdowej,
- cena najwyższa – maksimum z cen transakcyjnych danego dnia,
- cena najniższa – minimum z cen transakcyjnych danego dnia,
- cena zamknięcia – cena ostatniej transakcji w danym dniu.

Spośród tych wyróżnionych cen, szerzej wykorzystamy dwie – cenę emisyjną oraz cenę rynkową zamknięcia w pierwszym dniu notowań. W rezultacie przedmiotem analizy będą scenariusze zmiany ceny i powiązane z nimi stopy zwrotu z inwestycji, której pozycja zamyka się na zamknięciu pierwszej sesji notowań dla danej spółki.

Zasady emisji akcji

Jako obszar zainteresowań niniejszego badania obrano debiuty nowych walorów na Rynku Głównym GPW. W tym przypadku ważnym elementem rozpatrywanego procesu subskrypcji oferty publicznej jest przygotowanie prospektu

emisyjnego. Jego kompletność jest zatwierdzana zawsze przez regulatora – Komisję Nadzoru Finansowego. Za jego pomocą można pozyskać informacje o takich kluczowych aspektach jak: wielkość i cena emisji, powiązania prawno-kapitałowe z innymi podmiotami oraz niekiedy także o przewidywanych założeniach przyszłej działalności [Bień 2008]. Rola tego dokumentu jest szczególnie istotna – głównie z uwagi na to, że przede wszystkim dzięki niemu drobni inwestorzy mają możliwość zapoznania się z wiarygodnością emitenta oraz na jego podstawie mogą oceniać rzetelność założeń emisji. Bez wątpienia wpływa on wobec tego na te rozważane tutaj pośrednio „początkowe” decyzje inwestycyjne i co za tym idzie – kształtowanie się cen. Jest to równocześnie ważne dla oceny spółki w rozumieniu aspektów analizy fundamentalnej.

Oczywiście istnieją także odrębne, inaczej sformalizowane okoliczności, w wyniku których spółka staje się spółką podlegającą notowaniom na Rynku Głównym GPW. Jednymi z takich sytuacji są przejścia spółek z alternatywnego systemu obrotu (ASO) – realizowanego w ramach rynku NewConnect – na rynek regulowany, czyli właśnie Rynek Główny GPW. Zauważmy jednak, że wówczas koncepcyjnie nie mamy do czynienia z zupełnie nowym walorem, o którym nic a priori nie wiemy. Jest to raczej niejako „przedłużenie” szeregu cen, którego proces mógł być już obserwowany w poprzednim systemie notowań. Z tego względu podjęto decyzję o wyłączeniu ich z procedury badawczej.

STATYSTYKI ILOŚCIOWE DEBIUTÓW GIEŁDOWYCH

Przeprowadzony zostanie przegląd wybranych zagadnień na podstawie danych, które są ewidencjonowane w oficjalnych giełdowych serwisach informacyjnych [GPW 2020]. Zbiór danych podlegający poniższej analizie traktuje o debiutach w piętnastoletnim okresie od 2005 roku do 2019 roku. Ze względu na znaczny zakres danych zostanie on zaprezentowany w sposób zagregowany.

W badanym okresie odnotowano nieco ponad 400 debiutów spółek na Rynku Głównym. Zgodnie z uwagami z poprzednich części wśród nich są też spółki, które wcześniej notowane były w alternatywnym systemie obrotu. Występują one w liczności stanowiącej niecałe 17% debiutów z tego okresu.

Rysunek 2. Struktura debiutów na GPW



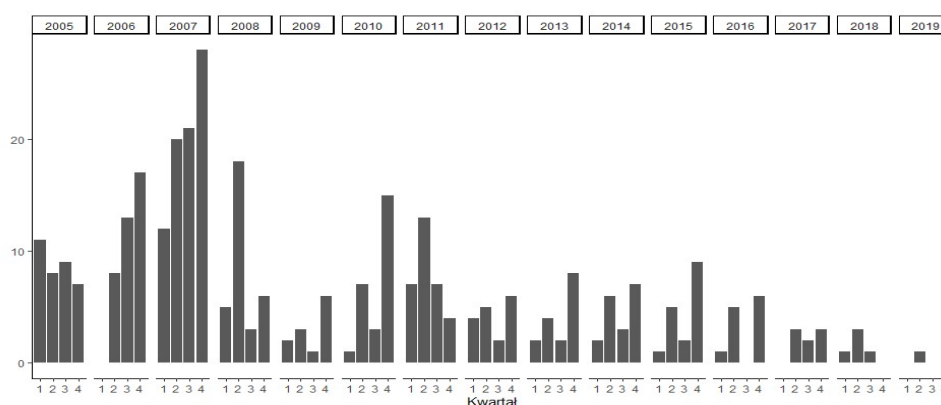
Źródło: opracowanie własne

W przypadku rynku NewConnect - widoczny jest już na tym etapie stosunkowo niewielki ilościowy udział takich przypadków w całkowitej liczbie debiutów giełdowych na Rynku Głównym. Porządkując wobec tego dotychczas wyartykułowane uwagi podkreślić należy, że dalsza część szczegółowych analiz dotyczy debiutów, które miały miejsce bezpośrednio i pierwotnie na Rynku Głównym GPW.

Liczba debiutów na przestrzeni lat

Rysunek 3. przedstawia szereg czasowy liczby debiutów giełdowych w kolejnych kwartałach.

Rysunek 3. Liczba debiutów w poszczególnych kwartałach lat 2005-2019



Źródło: opracowanie własne na podstawie danych z [GPW 2020]

Łatwo zauważyć, że liczba debiutów w poszczególnych okresach była dosyć mocno zróżnicowana – przy czym widać wyraźnie dużą koncentrację debiutów w początku badanego okresu i silne, skokowe ograniczenie liczebności w okolicy roku 2009. Bez wątplenia jego przyczyn należy szukać w globalnym kryzysie finansowym zapoczątkowanym upadkiem amerykańskiego banku inwestycyjnego Lehman Brothers Holdings w 2008. Z kolei największa liczba debiutów odbyła się w roku 2007. Dokonując dalszego opisu posłużono się kilkoma statystykami opisowymi zebranych w tabeli 1. Zawiera ona obliczone wartości średniej arytmetycznej, odchylenia standardowego i współczynnika zmienności dla liczby debiutów w badanym okresie w ujęciach odpowiednio: kwartalnym i rocznym.

Tabela 1. Wybrane statystyki opisowe dotyczące liczby debiutów giełdowych

Statystyka		Średnia	Odchylenie standardowe	Współczynnik zmienności
Ujęcie	Kwartałne	5,73	5,58	97%
	Roczne	22,93	18,76	82%

Źródło: obliczenia własne

Wartości statystyk dosyć wyraźnie ustalają uwagę na bardzo duże zróżnicowanie wyznaczanych podokresów – zarówno kwartałów jak i poszczególnych lat – w kontekście liczby debiutujących spółek. Ciekawym spostrzeżeniem jest natomiast fakt, że w tym kontekście kwartały były zróżnicowane między sobą nieco bardziej. Dla ogólnego rozeznania skali odbywania się debiutów giełdowych wyznaczono także wartości średnie, które wskazują że przeciętnie możemy mówić o około 6 debiutach w ciągu kwartału lub w ujęciu rocznym – około 23 w ciągu roku. Biorąc jednak pod uwagę stwierdzoną dużą dyspersję, należy wartości te traktować czysto pogładowo.

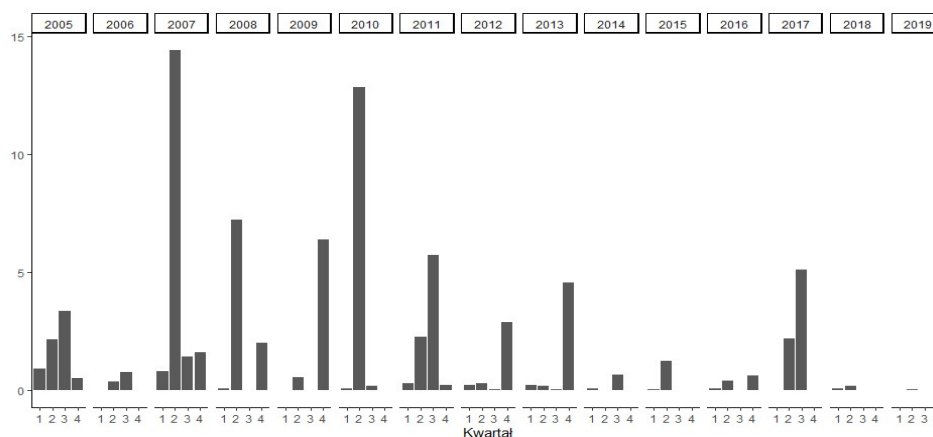
Przyjrzyjmy się jeszcze dosyć widocznemu zmniejszaniu się liczby debiutujących spółek w kolejnych latach badanego okresu. Ilościowy charakter tego zjawiska można opisać za pomocą indeksów łańcuchowych. Na bazie obliczonej na ich podstawie średniej geometrycznej możemy wyznaczyć syntetyczne średniokresowe (tutaj: średnioroczne) tempo zmian. Wartość tej statystyki wynosi w rozpatrywanym okresie w ujęciu rocznym: – 22.4%. Oznacza to, że rokrocznie liczba debiutów malała przeciętnie o ponad jedną piątą. Zjawisko to jednak zdaje się być częściowo wytłumaczalne choćby przez występowanie zwiększonej liczby debiutów spółek o udziale Skarbu Państwa w początkach XXI wieku.

Wartości ofert debiutów na przestrzeni lat

Oczywiście liczba walorów jakie pojawiają się w obrocie jest niezwykle istotna dla jego funkcjonowania. Rozszerzanie w ten sposób „oferty” giełdowej wpływa bezpośrednio na warunki inwestycyjne. Pojawiające się walory pozwalają inwestorom wybierać interesujące dla nich propozycje zarówno ze względu na aspekty fundamentalne – jak na przykład branża i obszar działalności spółki, ale też wpływa to na możliwości optymalizacji i doboru preferowanego portfela inwestycyjnego w rozumieniu teorii portfelowych, w szczególności przy badaniu oczekiwanej stopy zwrotu, oczekiwanego ryzyka jak i poziomu korelacji między walorami.

Nie mniej ważne jest również to, ile nowego kapitału pojawia się w obrocie giełdowym. Wartość tą dla każdej spółki odzwierciedla wartość oferty debiutu danej spółki. Poprzez to pojęcie rozumiana jest suma następujących kwot: wartości emisji nowych akcji oraz wartości sprzedaży akcji posiadanych przez dotychczasowych akcjonariuszy, które nie były wcześniej dopuszczone do obrotu. Analizę tego elementu przeprowadzimy na danych uzyskanych ze źródła wskazywanego już poprzednio [GPW 2020].

Rysunek 4. Wartość ofert debiutów giełdowych [mld] w poszczególnych kwartałach lat 2005-2019



Źródło: opracowanie własne na podstawie danych z [GPW 2020]

Omawiając rysunek 4. nie sposób nie zwrócić uwagi na nieporównywalnie większe wartości ofert debiutów odpowiednio w drugich kwartałach w roku 2007 i 2010. W przypadku roku 2007 jest to w dużej mierze wynik kwotowo rekordowego debiutu zagranicznej spółki Immoeast na polskim parkiecie. Rok 2010 to natomiast kluczowe wartościowo debiuty spółek Skarbu Państwa – w szczególności: Powszechnego Zakładu Ubezpieczeń i Tauronu. Koresponduje to wydatnie ze spostrzeżeniem sformułowanym we wcześniejszej części. Dla ilościowego ujęcia omawianego zjawiska obliczono, paralelnie do danych z tabeli 1., wartości kilku statystyk opisowych i zebrano je w tabeli 2.

Tabela 2. Wybrane statystyki opisowe dotyczące wartości debiutów giełdowych

Statystyka		Średnia [mld PLN]	Odchylenie standardowe [mld PLN]	Współczynnik zmienności
Ujęcie	Kwartałne	1,4	2,74	195%
	Roczne	5,62	5,37	96%

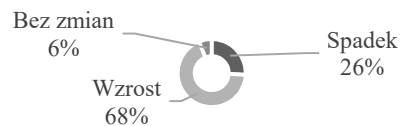
Źródło: obliczenia własne

Ważnym charakterystyką jest bardzo duża zmienność wartości w poszczególnych podokresach. Zarówno analiza rysunku 4. jak i odczytanie obliczonej wartości współczynnika zmienności z tabeli 2., pozwala sformułować wniosek, że podokresy zarówno w ujęciu kwartalnym jak i rocznym różniły się między sobą dużo bardziej gdy rozpatrujemy wartość debiutów niż wtedy gdy rozpatrywana jest liczba debiutów. W związku z tak dużą dyspersją podobnie tutaj wyznaczone wartości średnie mają użyteczność jedynie poglądową. Inaczej jednak niż poprzednio – nie obserwujemy w danych zdecydowanego, jawnego trendu malejącego.

Wynik dnia debiutu

Inną ciekawą i nierzadko rozpatrywaną charakterystyką jest wynik dnia debiutu, który wiąże się bezpośrednio z ceną zamknięcia w pierwszym dniu notowań. Innymi słowy chodzi tutaj o sprawdzenie tego, czy spółka w ciągu pierwszego dnia notowań pozwoliła zarobić inwestorom, którzy nabyli akcje w pierwszej ofercie publicznej. W celu zbadania tej cechy dokonuje się porównania ceny emisyjnej z ceną rynkową na zamknięcie w dniu debiutu giełdowego. Frakcje scenariuszy przebiegu ceny z badanego okresu zaprezentowano na rysunku 5.

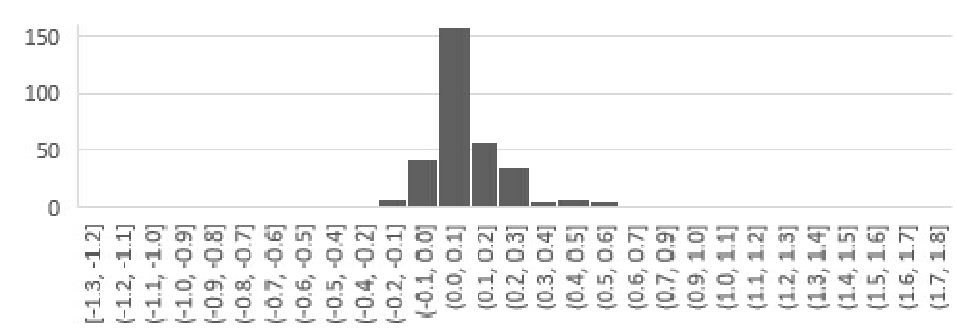
Rysunek 5. Frakcje scenariuszy zmian ceny w dniu debiutu



Źródło: opracowanie własne

Wyznaczone wartości poszczególnych frakcji pozwalają wyraźnie zauważyć, że w zadanym okresie (lata 2005 – 2019) pierwszy dzień na giełdzie kończył się dla pierwszych posiadaczy akcji najczęściej korzystnym scenariuszem. Wartość akcji nowo debiutujących spółek wzrastała w pierwszym dniu w ponad dwa razy większej ilości przypadków niż malała. Zauważalne były także niezbyt liczne sytuacje, w których cena akcji po pierwszym dniu notowań była równa cenie emisyjnej.

Rysunek 6. Histogram logarytmicznej stopy zwrotu pierwszego dnia notowań



Źródło: opracowanie własne

W celu dokładniejszej analizy zmian ceny wyznaczono logarytmiczne stopy zwrotu i ich wartości zaprezentowano w formie histogramu na rysunku 6. Warto podkreślić, że te wartości o małych frakcjach to raczej wartości wyższe od wartości przeciętnej niż te od niej niższe.

Jeżeli potraktujemy wyznaczone historyczne wartości zmiany ceny jako wartości próby, to możemy przeprowadzić test statystyczny weryfikujący

populacyjne zjawisko zmiany ceny, innymi słowy – weryfikację tego czy przeciętnie inwestor nabywający akcje w ramach emisji traci czy zyskuje w pierwszym dniu notowań. W tym celu przeprowadzimy procedurę testowania istotności dla wartości średniej logarytmicznej stopy zwrotu z pierwszego dnia notowań. Formułujemy hipotezy postaci:

$$H_0: E(y) = 0 \quad H_1: E(y) > 0$$

gdzie: $E(*)$ – wartość oczekiwana, y – logarytmiczna stopa zwrotu.

Odpowiednia statystyka testowa [Witkowska i in. 2012] ma dla dużych liczebności standaryzowany rozkład normalny i przybiera ona postać:

$$u = \frac{\bar{y}}{S_y} \sqrt{N} \quad (1)$$

gdzie \bar{y} to średnia stopa zwrotu z historycznych pierwszych dni notowania, S_y – estymator odchylenia standardowego stopy zwrotu, N – liczba spółek w historycznej próbie.

Odpowiedni zestaw danych obliczeniowych niezbędnych do przeprowadzenia testu istotności średniej dla naszych danych zawarto w tabeli 3.

Tabela 3. Zestawienie wartości wykorzystanych w teście istotności dla wartości średniej

Wyszczególnienie	\bar{y}	S_y	N	u	$\Phi(0,95)$
Wartość statystyki	0,08	0,22	345	6,99	1,64

Źródło: obliczenia własne

Ponieważ uzyskana wartość statystyki $u = 6,99$ jest większa od odpowiedniej wartości dystrybucyjnej standaryzowanego rozkładu normalnego $\Phi(0,95) = 1,64$, to możemy na poziomie istotności odrzucić hipotezę zerową zakładającą, że wartość oczekiwana jest równa zero. W związku z przeprowadzoną procedurą statystyczną można sformułować wniosek o tym, że przeciętnie cena akcji wzrasta w dniu debiutu. Co za tym idzie – inwestor, który nabywa je w pierwszej ofercie publicznej, statystycznie zyskuje w tym dniu.

PODSUMOWANIE

Debiut giełdowy jest ważnym momentem w funkcjonowaniu danej organizacji, ale także jest bardzo istotny dla funkcjonowania rynku kapitałowego. W końcu także – debiut giełdowy to moment, który umożliwia inwestorowi uzyskanie dostępu do nowego sposobu lokowania kapitału. W kontekście takich okazji inwestycyjnych okazuje się, że choć liczba debiutów ma tendencję malejącą to jednak nie można tego stwierdzić jednoznacznie o ich wartości.

Przeprowadzone badanie wykazało, że entuzjazm inwestorów utrzymuje się także bezpośrednio po debiucie, a większość spółek po pierwszym dniu notowań jest warta więcej niż w momencie emisji. Co więcej – badanie rozkładu stopy zwrotu pokazało, że istotność tej prawidłowości daje się potwierdzić również korzystając z wnioskowania statystycznego.

Należy mieć jednocześnie na uwadze, że horyzont inwestycji wielu inwestorów może być jednak dłuższy niż zbadany w prezentowanej pracy. W takich przypadkach nieodzownym będzie zastosowanie narzędzi wykorzystujących analizę szeregów czasowych. Uzasadnionym wobec tego wydaje się podjęcie dalszych wysiłków dotyczących tego jak na danym etapie życia waloru konstruować prognozy. Materializacją działań tego typu może być szukanie okresu po debiucie, w którym klasa modelu (np. ARIMA) przyjmuje taką samą postać przez odpowiednio długi czas.

BIBLIOGRAFIA

- Bień W. (2008) Rynek papierów wartościowych. Centrum Doradztwa i Informacji Difin, Warszawa.
- Czerniawski R. (1992) Giełdy. Wydawnictwo Park, Bielsko-Biała.
- Dębski W. (2014) Rynek finansowy i jego mechanizmy. Podstawy teorii i praktyki. Wydawnictwo Naukowe PWN, Warszawa.
- Giełda Papierów Wartościowych w Warszawie (2020) Główny Rynek GPW – Debiuty. <https://www.gpw.pl/debiuty> [dostęp. 08.02.2020].
- Górnjak T. (1995) Przegląd papierów wartościowych. Galeria Akcji, Warszawa.
- Krzywda M. (2007) GPW – Giełda Papierów Wartościowych w praktyce. Internetowe Wydawnictwo „Złote Myśli”, Gliwice.
- Pacutkowski K. (2020) Stabilizacja postaci modeli ARIMA dla notowań po debiucie giełdowym. Praca dyplomowa magisterska. WZiIM SGGW, Warszawa.
- Witkowska D., Matuszewska-Janica A., Kompa K. (2012) Wprowadzenie do ekonometrii dynamicznej i finansowej. Wydawnictwo SGGW, Warszawa.
- Zientara B., Mączak A., Ichnatowicz I., Landau Z. (1973) Dzieje gospodarcze Polski do 1939. Wyd. Wiedza Powszechna, Warszawa.

ANALYSIS OF DEBT ON THE WSE MAIN MARKET IN THE YEARS 2005-2019

Abstract: This study deals with debuts on the WSE Main Market in 2005-2019. In this context, information on the subject of the debut as well as the related procedural and organizational requirements were referred to at the beginning. The following section presents an overview of the data available on the WSE websites relating to the problem area under study, and an analysis of selected quantitative statistics on their basis. An important part was devoted to examining the scenarios of price changes on the first trading day and carrying out statistical application based on them.

Keywords: stock exchange debut, Warsaw Stock Exchange, WSE, rate of return, primary stock market, investment efficiency

ANALIZA ZMIAN POZIOMU ŻYCIA W POWIATACH WOJEWÓDZTWA MAŁOPOLSKIEGO¹

Jacek Wolak  <https://orcid.org/0000-0001-7158-2851>

Wydział Zarządzania

AGH Akademia Górniczo-Hutnicza w Krakowie

e-mail: jwolak@agh.edu.pl

Streszczenie: Celem opracowania jest przestrzenna analiza zmian poziomu życia mieszkańców powiatów województwa małopolskiego. W ramach jego realizacji, z wykorzystaniem wskaźnika syntetycznego wyznaczonego dla danych z lat 2011-2017 na podstawie metody TOPSIS, utworzono rankingi powiatów. W świetle wyników badań okazało się, że najwyższym poziomem życia cieszą się mieszkańcy Krakowa i jego najbliższych okolic. Zachodzące zmiany nie są gwałtowne, a malejący współczynnik zmienności wskaźnika syntetycznego sugeruje, że przestrzenne zróżnicowanie poziomu życia w Małopolsce się zmniejsza.

Słowa kluczowe: poziom życia, porządkowanie liniowe, metoda TOPSIS

JEL classification: I31, O1, C19

WSTĘP

Tematyka poziomu życia i jego przestrzennego zróżnicowania należy do istotnych zagadnień debaty publicznej. W szczególności znajduje się ona w obszarze zainteresowania władz państwowych i samorządowych, które w celu m.in. wyrównania szans, wdrażają regionalne programy rozwoju. Temat oceny zróżnicowania poziomu życia jest również chętnie eksplorowany w opracowaniach naukowych. W polskiej literaturze przedmiotu pojawiają się badania porównawcze, które dotyczą zarówno obszaru np. Unii Europejskiej lub Polski w podziale na kraje, regiony i podregiony (zob. [Bal-Domańska, Wilk 2007; Kuc 2015; Kuc 2017;

¹ Publikacja finansowana przez AGH Akademię Górniczo-Hutniczą im. Stanisława Staszica w Krakowie (dotacja podmiotowa na utrzymanie potencjału badawczego).

Sompolska-Rzechuła 2013]), jak i liczne analizy przeprowadzane dla wybranych, mniejszych jednostek administracyjnej (zob. [Malinowski 2017; Sompolska-Rzechuła, Machowska-Szewczyk 2018; Strózik, Strózik 2014] lub [Łuczak, Wysocki 2013]).

Mimo bogactwa opracowań naukowych dotyczących rozważanego w opracowaniu tematu, brakuje opracowań analizujących obszar województwa małopolskiego. Znane w literaturze badania (zob. [Berbeka 2002 i Berbeka 2005]) koncentrują się na jakości życia. Zdaniem autora uzasadnia to potrzebę przeprowadzenia analizy koncentrującej się na przestrzennym zróżnicowaniu poziomu życia mieszkańców województwa małopolskiego na poziomie powiatów.

Choć pojęcie poziomu życia jest szeroko stosowane, to w piśmiennictwie nie ma zgody co do formalnej definicji tego pojęcia. Nierzadko jest ono utożsamiane z innymi określeniami², tj. jakością życia i warunki życia, ale wielu badaczy podkreśla istniejące różnice (zob. [Gotowska, 2018] oraz [Sompolska-Rzechuła, Machowska-Szewczyk, 2018]).

W niniejszym opracowaniu przyjęto definicję, którą zaproponowali Bywalec i Wydymus [1992]. Określa ona poziom życia jako „stopień zaspokojenia potrzeb ludności wynikający z konsumpcji wytworzonych przez człowieka dóbr materialnych i usług oraz wykorzystania walorów środowiska naturalnego i społecznego”.

WYKORZYSTANE DANE

Rozważa się wiele zmiennych, które są brane pod uwagę w kontekście badań nad poziomem życia w jednostkach terytorialnych. Wstępna analiza przydatności zmiennych w niniejszej pracy objęła 51 zmiennych diagnostycznych, z których – po sprawdzeniu kompletności i wiarygodności danych oraz przeprowadzeniu analizy merytoryczno-formalnej (zob. [Bąk, 2017]) – finalnie wyłoniono 17 cech. Źródłem danych statystycznych jest Bank Danych Lokalnych GUS.

Zmienne, które posłużyły do budowy wskaźnika syntetycznego oceny poziomu życia, można podzielić na siedem głównych grup:

- ludność:

X_1 – saldo migracji ogółem na 1000 ludności [os.] (stymulanta),

X_2 – współczynnik obciążenia demograficznego osobami starszymi [%] (destymulanta);

- edukacja:

X_3 – absolwenci liceów ogólnokształcących w ogólnej liczbie absolwentów szkół ponadgimnazjalnych [%] (stymulanta),

X_4 – odsetek dzieci w wieku 3-5 lat objętych wychowaniem przedszkolnym [%] (stymulanta),

² Encyklopedia PWN (data dostępu: 12.06.2019)

- X_5 – odsetek osób zdających maturę wśród wszystkich maturzystów [%] (stymulanta);
- warunki mieszkaniowe:
- X_6 – mieszkania oddane do użytku w przeliczeniu na 10 000 ludności [szt.] (stymulanta),
- X_7 – odsetek mieszkań zaopatrzonych w instalacje [%] (stymulanta);
- ochrona zdrowia:
- X_8 – liczba mieszkańców przypadająca na aptekę ogólnodostępną [os.] (stymulanta),
- X_9 – liczba porad w ramach podstawowej opieki zdrowotnej [wizyty/os.] (destymulanta);
- transport:
- X_{10} – samochody osobowe w przeliczeniu na 1000 mieszkańców powiatu [szt.] (stymulanta),
- X_{11} – długość dróg gminnych i powiatowych o twardej nawierzchni w przeliczeniu na 100 km² powierzchni powiatu [km] (stymulanta),
- X_{12} – liczba wypadków drogowych w przeliczeniu na 100 000 osób [szt.] (destymulanta);
- rynek pracy:
- X_{13} – nowo zarejestrowane podmioty gospodarki narodowej w rejestrze REGON na 10 000 osób [szt.] (stymulanta),
- X_{14} – stopa bezrobocia rejestrowanego [%] (destymulanta),
- X_{15} – udział bezrobotnych pozostających bez pracy dłużej niż rok wśród wszystkich zarejestrowanych bezrobotnych [%] (destymulanta);
- kultura, rekreacja, czas wolny:
- X_{16} – liczba mieszkańców na 1 bibliotekę [os.] (destymulanta),
- X_{17} – koła/kluby/sekcje w ramach działalności centrów, domów, ośrodków kultury, klubów i świetlic w przeliczeniu na 1000 osób [szt.] (stymulanta).

METODA BADAWCZA

Budowa miary poziomu życia w oparciu o narzędzia wielowymiarowej analizy statystycznej, realizuje się poprzez techniki porządkowania liniowego. Pozwalają one na stworzenie wskaźnika syntetycznego i umożliwiają w ten sposób posortowanie rozważanych obiektów (zob. [Bąk, 2016], [Kukuła, Luty, 2017]). W literaturze przedmiotu wykorzystywane są dwa główne rodzaje procedur, które są podstawą do tworzenia rankingu obiektów: metody wzorcowe i bezwzorcowe. Metody wzorcowe wyznaczają wartość wskaźnika syntetycznego na podstawie odległości badanych obiektów od pewnych obiektów modelowych, a metody bezwzorcowe sumują lub uśredniają wartości unormowanych cech wykorzystanych w analizie problemu.

W badaniu empirycznym wykorzystano metodę TOPSIS, której autorami są Hwang i Yoon [1981]. Wywodzi się ona z teorii podejmowania decyzji i z powodzeniem jest stosowana zarówno w analizie przestrzennego zróżnicowania poziomu życia (zob. [Łuczak, Wysocki 2013]), jak i wielu innych praktycznych zastosowaniach porządkowania liniowego. (zob. [Zalewski 2012 i Bąk 2018]).

Jest to metoda wzorcowa, której ideą jest porównanie obiektów na podstawie miernika preferującego punkty położone możliwie najdalej od hipotetycznego rozwiązania antyidealnego i jednocześnie położone możliwie blisko hipotetycznego rozwiązania idealnego.

Algorytm metody TOPSIS prezentuje się następująco:

1. ustalenie charakteru zmiennych (stymulanta, destymulanta lub nominanta) oraz ew. przekształcenie nominant w stymulanty,
2. normalizacja wyjściowych zmiennych, która następuje poprzez przekształcenie ilorazowe, tj.

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}} \quad (1)$$

gdzie x_{ij} to pomiar dla i -tego obiektu w j -tej zmiennej,

3. wyznaczenie współrzędnych wzorca (z_0^+) i antywzorca (z_0^-), tj.

$$z_{0j}^+ = \begin{cases} \max_i(x_{ij}) & \text{zmienna } x_j \text{ to stymulanta} \\ \min_i(x_{ij}) & \text{zmienna } x_j \text{ to destymulanta} \end{cases} \quad (2)$$

$$z_{0j}^- = \begin{cases} \min_i(x_{ij}) & \text{zmienna } x_j \text{ to stymulanta} \\ \max_i(x_{ij}) & \text{zmienna } x_j \text{ to destymulanta} \end{cases} \quad (3)$$

4. obliczenie odległości obiektów od wzorca (d_{i0}^+), i antywzorca (d_{i0}^-), tj.

$$d_{i0}^+ = \sqrt{\sum_{j=1}^n (z_{ij} - z_{0j}^+)^2} \quad \text{oraz} \quad d_{i0}^- = \sqrt{\sum_{j=1}^n (z_{ij} - z_{0j}^-)^2} \quad (4)$$

5. wyznaczenie zmiennej syntetycznej (q_i), tj.

$$q_i = \frac{d_{i0}^-}{d_{i0}^- + d_{i0}^+} \quad (5)$$

WYNIKI BADAŃ

W wyniku przeprowadzonych obliczeń, w tabeli 1 zamieszczono wartości wskaźnika syntetycznego określającego poziom życia, a także określono ranking małopolskich powiatów na podstawie danych GUS z lat 2011, 2013, 2015 i 2017. Można zauważyć, że zgodnie z oczekiwaniami, klasyfikację otwiera miasto Kraków. Stolica województwa jest uznanym ośrodkiem akademickim, który w bardzo wysokim stopniu zaskapaja potrzeby mieszkańców w zakresie edukacji. Rosnące zainteresowanie międzynarodowych korporacji zakładających w jego granicach swoje siedziby, wpływa natomiast na wysoką ocenę rynku pracy.

Z bliskości geograficznej Krakowa w znacznym stopniu korzystają znajdujące się od 2013 roku na drugiej i trzeciej lokacie w rankingu powiaty wielicki i krakowski. Obok ponadprzeciętnych zarobków i dobrej sytuacji na rynku pracy, charakteryzują się one wysokim saldem migracji, który w znacznym stopniu wynika z efektu suburbanizacji.

Poza okolicami Krakowa, stosunkowo wysoki poziom życia w latach 2011-2017 jest notowany w dwóch miastach na prawie powiatów, w Nowym Sączu i Tarnowie. Co ciekawe, w tym przypadku powiaty ziemskie położone wokół nich, nie korzystają z ich bliskości geograficznej i znajdują się na odległych miejscach rankingu.

Tabela 1: Wartość zmiennej syntetycznej oraz ranking powiatów województwa małopolskiego ze względu na poziom życia

Powiat	Rok 2011		Rok 2013		Rok 2015		Rok 2017	
	MS	ranga	MS	Ranga	MS	ranga	MS	Ranga
m. Kraków	68,85%	1	67,69%	1	66,77%	1	63,69%	1
wielicki	58,89%	2	54,95%	2	57,04%	2	53,56%	2
krakowski	51,46%	6	53,56%	3	55,43%	3	50,42%	3
m. Tarnów	57,73%	3	51,50%	5	53,91%	4	48,62%	4
m. Nowy Sącz	56,25%	4	53,51%	4	51,50%	5	47,76%	5
myślenicki	40,16%	15	41,05%	12	45,29%	10	47,49%	6
oświęcimski	50,06%	8	49,40%	7	50,96%	6	46,12%	7
wadowicki	46,84%	9	47,68%	8	48,19%	8	44,76%	8
bocheński	50,23%	7	47,16%	9	49,55%	7	44,50%	9
chrzanowski	52,80%	5	50,32%	6	47,15%	9	43,63%	10
olkuski	45,02%	10	42,78%	10	43,28%	12	43,62%	11
suski	40,20%	14	40,08%	13	42,61%	13	42,63%	12
proszowicki	44,83%	11	41,20%	11	43,59%	11	41,56%	13
brzeski	40,55%	13	38,28%	15	39,76%	16	39,41%	14
limanowski	33,75%	21	34,78%	20	37,80%	19	38,23%	15
nowotarski	36,03%	18	37,69%	17	42,30%	14	37,89%	16
dąbrowski	38,42%	16	38,96%	14	42,22%	15	37,57%	17
tatrzański	43,85%	12	37,66%	19	38,32%	18	37,15%	18
gorlicki	34,81%	20	37,66%	18	37,61%	20	36,73%	19
tarnowski	36,54%	17	37,69%	16	38,95%	17	35,84%	20
miechowski	35,10%	19	34,43%	21	34,12%	21	33,25%	21
nowosądecki	27,33%	22	29,71%	22	30,52%	22	30,82%	22

Źródło: opracowanie własne

Do oceny skali zmian poziomu życia zdecydowano się na wyznaczenie macierzy korelacji rangowej Spearmana (tabela 2).

Tabela 2. Wartości współczynnika korelacji rangowej Spearmana dla rankingów poziomu życia w powiatach województwa małopolskiego w latach 2011-2017

Rok	2011	2013	2015	2017
2011	1,00	0,94	0,93	0,87
2013	0,94	1,00	0,98	0,94
2015	0,93	0,98	1,00	0,96
2017	0,87	0,94	0,96	1,00

Źródło: opracowanie własne

Analizując wyniki przedstawione w tabeli 1 i tabeli 2 można zauważyć, że miejsca rankingowe powiatów zmieniają się w czasie. I choć w perspektywie dwuletniej różnice nie są znaczące (współczynnik korelacji Spearmana między kolejnymi punktami czasowymi badania jest bardzo wysoki i wynosi od 0,94 do 0,98), to tylko trzy powiaty zajmujące skrajne lokaty, nie zmieniły swojego położenia w ciągu badanych sześciu lat. Są to miasto Kraków i powiat wielicki, które znajdują się na czele rankingu oraz regularnie zamykający klasyfikację powiat nowosądecki.

Wśród stale podnoszących poziom życia swoich mieszkańców, znajdują się powiat myślenicki (awans z 15 miejsca w roku 2011 na miejsce 6 w roku 2017) i limanowski (awans z 21 pozycji w 2011 roku na 15 pozycję w 2017 roku). W tym samym okresie największe spadki w rankingu zanotowały: powiat tatrzański (o sześć miejsc z pozycji 12 na 18) oraz chrzanowski (o pięć miejsc, z pozycji 5 na lokatę 10).

W celu dokładniejszego zaobserwowania przestrzennej dynamiki zmian poziomu życia, zgodnie z zasadą jednego odchylenia od średniej, zdecydowano się podzielić rozważane obiekty na 4 grupy (patrz Rysunek 1):

- grupa I (najwyższy wskaźnik poziom życia), jeśli

$$MS_i > \overline{MS} + S_{MS} \quad (6)$$

- grupa II (ponadprzeciętny wskaźnik poziom życia), jeśli

$$MS_i \in (\overline{MS}, \overline{MS} + S_{MS}] \quad (7)$$

- grupa III (przeciętny poziom życia), jeśli

$$MS_i \in (\overline{MS} - S_{MS}, \overline{MS}] \quad (8)$$

- grupa IV (niski poziom życia), jeśli

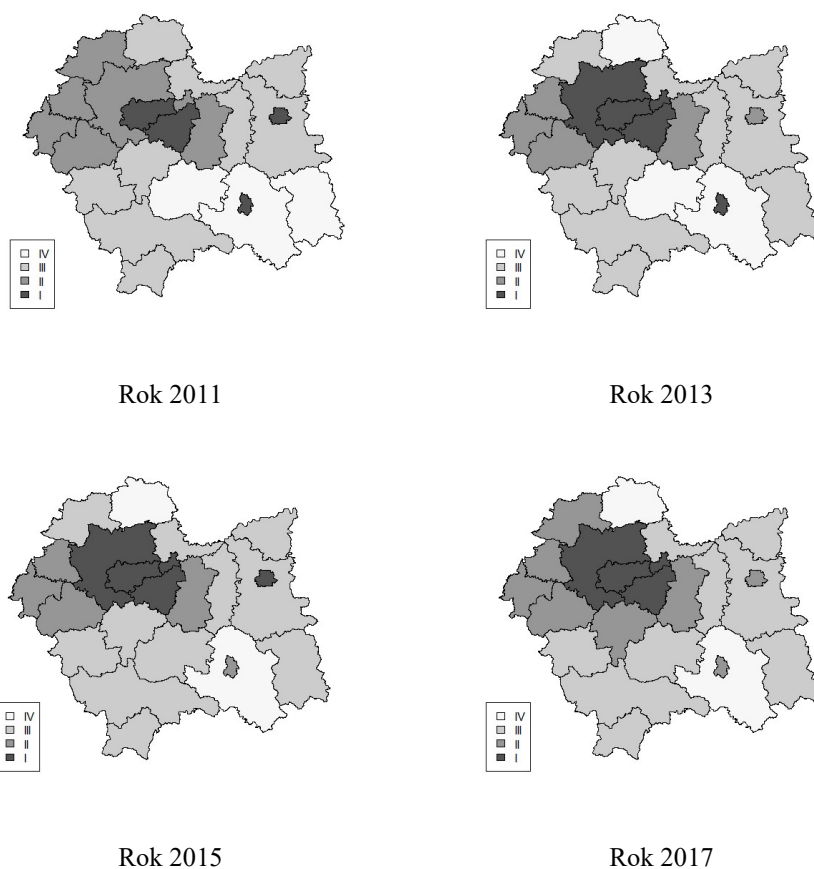
$$MS_i < \overline{MS} - S_{MS} \quad (9)$$

gdzie \overline{MS} i S_{MS} to odpowiednio: średnia i odchylenie standardowe wartości wskaźnika syntetycznego w rozważanym roku.

Uzyskane wyniki przedstawiono na Rysunku 1. Jak można zauważyć, w 2011 roku powiaty o najwyższym wskaźniku poziom życia koncentrują się wokół miast na prawie powiatów. Wraz z upływem lat sytuacja zmieniła się na niekorzyść Tarnowa i Nowego Sącza. Aktualnie (wg danych na 2017 rok) w grupie I znajdują się tylko

miasto Kraków oraz bezpośrednio z nim sąsiadujące powiaty wielicki i krakowski. W skład grupy II wchodzi przede wszystkim powiaty położone nieopodal (głównie na wschód i zachód) od stolicy województwa małopolskiego, a także coraz więcej oferujący swoim mieszkańcom, powiat myślenicki. Najniższy poziom życia (grupa IV) jest udziałem mieszkańców powiatu miechowskiego i nowosądeckiego.

Rysunek 1. Przestrzenne zróżnicowanie życia w powiatach województwa małopolskiego w latach 2011-2017



Źródło: opracowanie własne

W trakcie badań zauważono także, że współczynnik zmienności dla syntetycznego wskaźnika poziomu życia nieznacznie maleje w czasie (od 22,3% w 2011 roku, przez 20,1% w roku 2013, 18,6% w roku 2015 do 17,2% w 2017 roku). Może to sugerować zmniejszanie się różnic w poziomie życia mieszkańców Małopolski.

PODSUMOWANIE

Celem badania była analiza przestrzenna zmian poziomu życia mieszkańców powiatów województwa małopolskiego. Do realizacji problemu wykorzystano wzorcową metodę porządkowania liniowego (tj. metodę TOPSIS), w efekcie zastosowania której uzyskano syntetyczne wskaźniki poziomu życia Małopolan w podziale na powiaty w latach 2011, 2013, 2015 i 2017.

Zaprezentowane wyniki dowodzą, że poziom życia w omawianym regionie wykazuje przestrzenne zróżnicowanie, a widoczne w czasie niewielkie zmiany wskazują raczej na zmniejszanie się różnic. Obszarami o najwyższym poziomie życia, są powiaty grodzkie (miasta Kraków, Nowy Sącz i Tarnów), a także powiaty sąsiadujące z Krakowem (tj. powiat krakowski i powiat wielicki).

BIBLIOGRAFIA

- Bal-Domańska B., Wilk J. (2011) Wielowymiarowa analiza statystyczna poziomu życia w podregionach w roku 2009, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 163, 58-69.
- Bank Danych Lokalnych <https://bdl.stat.gov.pl/BDL/start> (dostęp 22.06.2019).
- Bąk A. (2016) Porządkowanie liniowe obiektów metodą Hellwiga i TOPSIS – analiza porównawcza, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 426, 22-31.
- Bąk A. (2017) Statystyczne metody doboru zmiennych w porządkowaniu liniowym, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 468, 29-37.
- Bąk A. (2018) Zastosowanie metod wielowymiarowej analizy porównawczej do oceny stanu środowiska w województwie dolnośląskim, *Wiadomości Statystyczne*, 63(1), 7-20.
- Berbeka J. (2002) Poziom życia w ocenie gospodarstw domowych województwa małopolskiego (na podstawie wyników badań ankietowych), *Zeszyty Naukowe/Akademia Ekonomiczna w Krakowie*, 612, 17-29.
- Berbeka J. (2005) Jakość życia ludności w województwie małopolskim : ocena subiektywna, *Zeszyty Naukowe / Akademia Ekonomiczna w Krakowie*, 697, 17-28.
- Bywalec C. Wydymus S. (1992) Poziom życia ludności Polski w porównaniu z krajami Europejskiej Wspólnoty Gospodarczej, *Ekonomista*, 5-6, 669-687.
- Gotowska M (2018) Poziom życia w krajach Unii Europejskiej, *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, 4, 89-93.
- Hwang C.L., Yoon K. (1981) *Multiple Attribute Decision Making: Methods and Applications*, Springer-Verlag, New York.
- Kuc M. (2015) Wpływ sposobu definiowania macierzy wag przestrzennych na wynik porządkowania liniowego państw Unii Europejskiej pod względem poziomu życia ludności, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 384, 163-170.
- Kuc M. (2017) Poziom życia w regionach państw nordyckich z uwzględnieniem zależności przestrzennych, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 469, 100-108.

- Kukuła K., Luty L. (2017) Jeszcze o procedurze wyboru metody porządkowania liniowego, *Przegląd Statystyczny*, 64 (2), 163-176.
- Łuczak A., Wysocki F. (2013) Zastosowanie mediany przestrzennej Webera i metody TOPSIS w ujęciu pozycyjnym do konstrukcji syntetycznego miernika poziomu życia, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*. 278, 63-73.
- Malinowski M. (2017) Przestrzenne zróżnicowanie poziomu życia ludności w ujęciu powiatów, *Wiadomości Statystyczne*, 62(2), 52-71.
- Sompolska-Rzechuła A. (2013) Zastosowanie miar pozycyjnych do porządkowania liniowego województw Polski ze względu na poziom jakości życia, *Przegląd Statystyczny*, 60(4), 523-538.
- Sompolska-Rzechuła A., Machowska-Szewczyk, M. (2018) Warunki życia w powiatach ziemskich województwa zachodniopomorskiego w latach 2002-2015 w świetle zmian demograficznych, *Przegląd Statystyczny*, 65(1), 115-136.
- Strózik D., Strózik T. (2014) Przestrzenne zróżnicowanie poziomu życia w województwie wielkopolskim, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 328, 335-342.
- Zalewski W. (2012), Zastosowanie metody TOPSIS do oceny kondycji finansowej spółek dystrybucyjnych energii elektrycznej, *Ekonomia i Zarządzanie*, 4, 137-145.

ANALYSIS OF CHANGES IN STANDARD OF LIVING IN POWIATS OF THE LESSER POLAND VOIVODESHIP

Abstract: The aim of the study is a spatial analysis of changes in the standard of living of the poviats of the Lesser Poland voivodeship. During the study, the synthetic measure has been used which is based on 2011-2017 data. Finally with TOPSIS method the rankings of poviats has been created. It is shown that Krakow and its closest neighborhoods citizens enjoy the highest standard of living. The changes in synthetic measure are rather slow. Coefficient of variation suggests that the spatial diversity of standard of living in Lesser Poland is decreasing.

Keywords: standard of living, linear ordering, TOPSIS method

JEL classification: I31, O1, C19

ANALIZA PORÓWNAWCZA STANU SŁUŻBY ZDROWIA W WOJEWÓDZTWACH

Marta Wójcik  <https://orcid.org/0000-0003-3654-386X>

Uniwersytet Ekonomiczny we Wrocławiu

e-mail: 165382@student.ue.wroc.pl

Streszczenie: Artykuł porównuje stan służby zdrowia w wybranych obszarach tematycznych w województwach. Skupia się na 2 głównych aspektach, takich jak: wydatki na służbę zdrowia i liczba specjalistów. Analizę przeprowadzono na podstawie danych pochodzących z lat 2018-2019. Jej celem było pogrupowanie województw ze względu na wybrane zmienne dotyczące stanu służby zdrowia. Przeanalizowano ich umiejscowienie geograficzne, oraz demografię. Wskazano rozbieżność pomiędzy stanem służby zdrowia w aglomeracjach miejskich i obszarach wiejskich. W badaniu wykorzystano metodę k-średnich.

Słowa kluczowe: analiza skupień, metoda k-średnich, służba zdrowia, kadra medyczna, finansowanie służby zdrowia

JEL classification: C38

WSTĘP

Służba zdrowia w dzisiejszych czasach zyskuje na istotności. Społeczeństwo w szybkim tempie starzeje się, a osoby po 65 roku życia wymagają znacznie bardziej kosztownej i częstszej pomocy medycznej. W ostatnich latach głośno zrobiło się o tak zwanych „ruchach antyszczepionkowych”, promujących brak szczepień u dzieci. Jest to niepokojące zjawisko, które może wpłynąć na większą liczbę zachorowań na choroby, które wydawałoby się, że są tylko relikami przeszłości.

Ogromnym testem dla służby zdrowia okazała się pandemia koronawirusa. Kraje z wysoką liczbą zachorowań cierpiały na brak odpowiedniej ilości łóżek szpitalnych i sprzętu. W Polsce, w momencie pisania artykułu, liczba zachorowań utrzymuje się na względnie stałym poziomie. Pozwala ona na dostarczenie pomocy

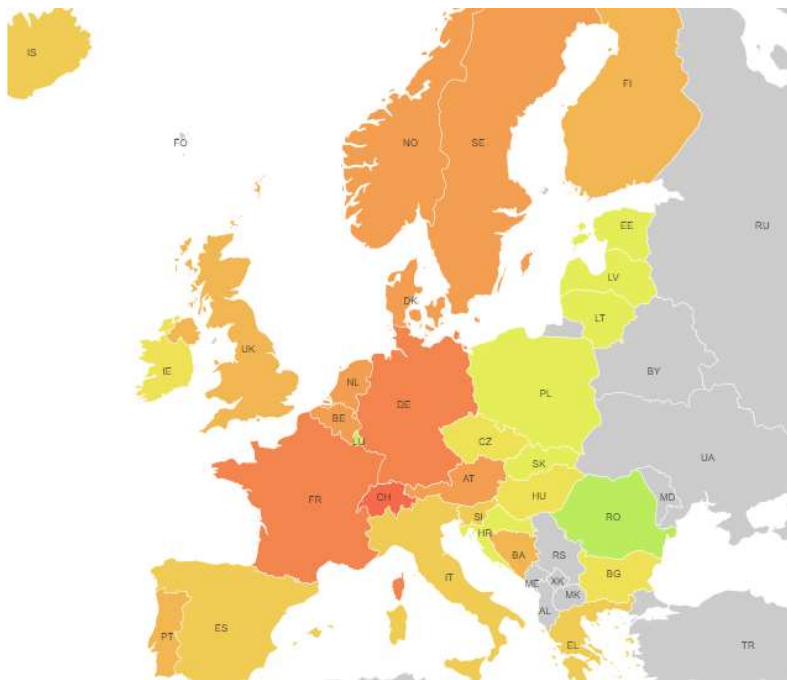
<https://doi.org/10.22630/MIBE.2020.21.1.5>

najbardziej jej potrzebującym. Nie ma jednak pewności kiedy pandemia ustąpi, ani czy nie pojawi się na jej miejsce kolejna.

Służba zdrowia w Polsce jest utrzymywana z obowiązkowych ubezpieczeń zdrowotnych i dotowana przez budżet państwa, oraz samorząd terytorialny [Dercz, Izdebski 2001]. Zadaniem samorządu jest zakładanie, zarządzanie i likwidowanie zakładów opieki zdrowotnej. Dodatkowo w jego obowiązkach leży profilaktyka chorób i wsparcie niepełnosprawnych [Golinowska i in. 2012]. Ośrodki zdrowotne dzielą się na wojewódzkie, powiatowe i gminne. Nie są sterowane odgórnie, przez państwo, a samorządowo. Sprawia to, że, potencjalnie, sytuacja w poszczególnych województwach może się istotnie różnić. Analiza pozwoli na przedstawienie i porównanie między sobą danych dotyczących województw.

Trzech na czterech Polaków jest niezadowolonych z tego jak funkcjonuje opieka zdrowotna. Wskazuje się dwa główne powody takiego stanu rzeczy. Pierwszy to organizacja służby zdrowia, a drugi to poziom jej finansowania. Odbiega on znacznie od poziomu prezentowanego przez inne kraje europejskie [Libura i in. 2018].

Rysunek 1. Mapa odsetka PKB przeznaczanego przez wybrane państwa europejskie na służbę zdrowia w 2016 roku



Źródło: www.ec.europa.eu/health/indicators_data/indicators_en

Polska posiada zdecydowanie zbyt niską liczbę lekarzy i pielęgniarek na 1000 osób. Jest to liczba najniższa w Europie, wynosi odpowiednio 2,3 i 5,2 na

1000 osób. Problemem jest między innymi odpływ wykształconej kadry poza granice kraju, głównie ze względu na zarobki [Ibidem].

Zasoby kadrowe stanowią istotny element w funkcjonowaniu służby zdrowia. Liczba specjalistów przypadających na osobę zwykle stanowi o tym jak efektywnie mogą działać. To samo tyczy się wydatków publicznych, które przeznaczone są na służbę zdrowia. Im są wyższe, tym pozwalają na lepszy sprzęt, który ułatwia i przyspiesza wykonywaną pracę, a często wręcz umożliwia wykonanie określonego zabiegu.

Z powyższych względów zdecydowano się na przeprowadzenie analizy, której celem było pogrupowanie województw ze względu na takie aspekty jak: kadra i wydatki na służbę zdrowia. Inne wymiary jakie można analizować w kontekście służby zdrowia to między innymi punkt widzenia pacjenta. Perspektywa ta, została przedstawiona na przykład w Indeksie Sprawności Ochrony Zdrowia prowadzonym przez PwC. Sprawdza województwa pod 5-ciomą względami: prawa pacjenta i informacja, dostępność, wyniki, przyjazność konsumentom i farmaceutyki. Innym popularnym wskaźnikiem szeregującym kraje europejskie jest Europejski Konsumentki Indeks Zdrowia. Skupia się on na wskaźnikach takich jak: prawa pacjenta, dostępność (czas oczekiwania na leczenie), wyniki leczenia, zakres i zasięg oferowanych świadczeń, profilaktyka, środki farmaceutyczne. W badaniu w 2016 roku Polska zajęła 31 pierwsze miejsce na 35 krajów biorących udział w badaniu. Autorzy wyróżnili dobry i rozbudowany system kształcenia medycznego. Stwierdzono, że powód niskiej lokaty nie jest jednoznaczny, jednak może, według nich, mieć związek z prowadzoną przez Polskę polityką.

WYKORZYSTANE DANE I METODA BADAWCZA

W celu wykonania analizy zebrano dane dotyczące służby zdrowia. Uwzględniono najnowsze dostępne dane. Zmienne zostały dobrane w sposób subiektywny. W danych zawarto liczbę przychodni i łóżek w zakładach opiekuńczo-leczniczych, jako przedstawicieli podmiotów medycznych, ze względu na to, że ich głównymi użytkownikami są osoby starsze. W procesie starzenia się społeczeństwa zyskują one na istotności. Wybrano głównych przedstawicieli kadry medycznej i wydatki gmin, powiatów, oraz województw na służbę zdrowia. W trakcie wyboru uwzględniono, by dane posiadały zmienność nie mniejszą niż 0,1, oraz nie były ze sobą silnie skorelowane [Ostasiewicz 1998]. Zmienne wybrane do analizy to:

- X_1 – przychodnie podległe samorządowi terytorialnemu na 10 tys. osób (2019),
- X_2 – farmaceuci na 10 tys. osób (2018),
- X_3 – lekarze posiadający prawo wykonywania zawodu medycznego na 10 tys. ludności (2018),

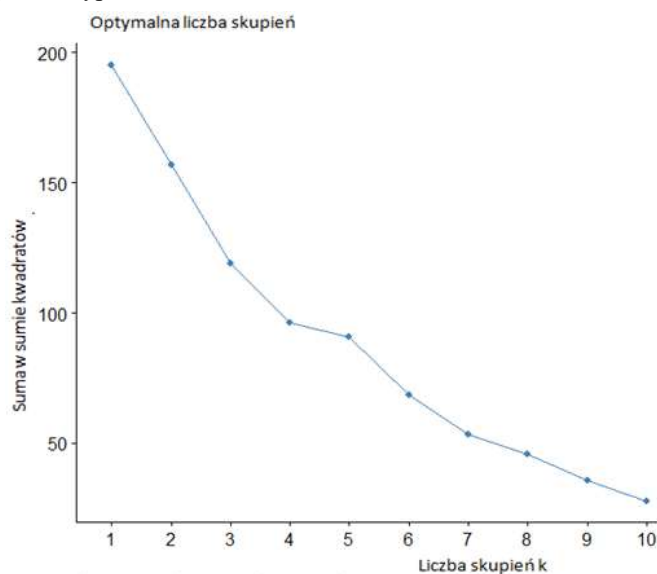
- X_4 – pielęgniarki i położne na 10 tys. osób (2018),
- X_5 – łóżka w hospicjach, zakładach opiekuńczo-leczniczych i pielęgnacyjno-opiekuńczych na 100 tys. ludności (2018),
- X_6 – wydatki na ochronę zdrowia przez gminy łącznie z miastami na prawach powiatu na 10 tys. osób (2018),
- X_7 – wydatki na lecznictwo ambulatoryjne przez gminy łącznie z miastami na prawach powiatu na 10 tys. osób (2019),
- X_8 – wydatki na przeciwdziałanie alkoholizmowi przez gminy łącznie z miastami na prawach powiatu na 10 tys. osób (2019),
- X_9 – wydatki budżetów powiatu na służbę zdrowia na 10 tys. osób (2019),
- X_{10} – wydatki budżetów województw na służbę zdrowia na 10 tys. osób (2019),
- X_{11} – wydatki budżetów województw na medycynę pracy na 10 tys. osób (2019),
- X_{12} – łóżka w szpitalach ogólnych na 10 tys. osób (2018),
- X_{13} – liczba szpitali ogólnych na 10 tys. osób (2018).

Do przeprowadzenia analizy wykorzystano analizę skupień. Jest to jedna z metod uczenia nienadzorowanego. Jej celem jest pogrupowanie obserwacji najbardziej podobnych do siebie. Każdy ze zbiorów jest niepusty i rozłączny. Obiekty badane są jako punkty wielowymiarowej przestrzeni, a jej wymiar określany jest liczbą zmiennych, które opisują obiekty [Lasek 2002]. W przypadku tej metody istotną kwestią jest normalizacja pierwotnych wartości. Jej celem jest umożliwienie porównywalności pomiędzy zmiennymi [Stanimir 2006]. Po transformacji wartości posiadają ujednolicone rzędy wielkości oraz są pozbawione mian [Walesiak 2014]. W artykule skorzystano ze standaryzacji, jako jednej z metod normalizacji danych. Bazuje ona na wyrażeniu wartości cechy w postaci odchylenia standardowego. W celu znormalizowania zmiennych wykonano standaryzację, w efekcie otrzymano zmienne o średniej 0 oraz odchyleniu standardowym równym 1 [Balicki 2013]. Wybrano grupowanie metodą k-średnich należąca do grupy niehierarchicznych metod klasteryzacji. Aby ją przeprowadzić należy uprzednio wskazać liczbę skupień a-priori [Stanisz 2007]. Wybór liczby skupień następuje arbitralnie lub za pomocą wybranego kryterium, takiego jak wykres osypiska, wskaźnik Silhouette, oraz kryterium Calińskiego-Harabasz. W analizie jako kryterium przyjęto wykres osypiska. Wykres osypiska przedstawia wartości całkowitej sumy kwadratów wewnątrzgrupowych w zależności od liczby skupień. Celem metody k-średnich jest zminimalizowanie zmienności w zbiorze i jej maksymalizacja poza nim. Badanie przeprowadzono z wykorzystaniem programu R. Wizualizacje wykonano w programie QGIS.

WYNIKI BADAŃ

W celu wyznaczenia liczby skupień na jaką podzielić zbiór skorzystano z wykresu osypiska. Analizując rysunek 2. zdecydowano się na podział obiektów na 4 skupienia.

Rysunek 2. Wykres osypiska



Źródło: opracowanie własne z zastosowaniem programu R

W efekcie uzyskano podział zaprezentowany w tabeli 1.:

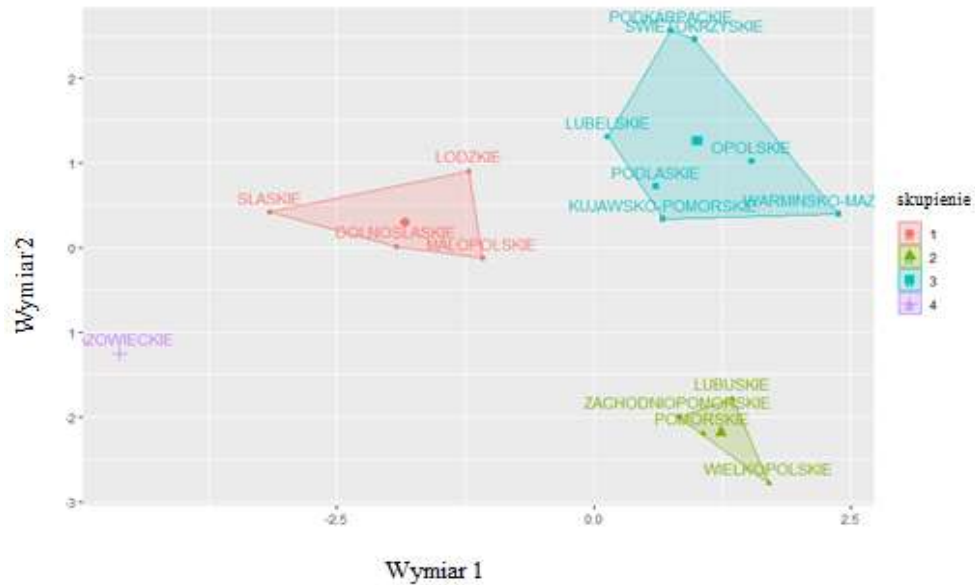
Tabela 1. Podział skupień

Nr skupienia			
1	2	3	4
Śląskie	Zachodniopomorskie	Kujawsko-Pomorskie	Mazowieckie
Łódzkie	Wielkopolskie	Świętokrzyskie	
Małopolskie	Pomorskie	Warmińsko-Mazurskie	
Dolnośląskie	Lubuskie	Podkarpackie	
		Opolskie	
		Lubelskie	
		Podlaskie	

Źródło: opracowanie własne

W celu sprawdzenia, czy 4 skupienia połączyły najbardziej zbliżone obiekty wykonano wizualizację przedstawioną w rysunku 3. prezentującą podział.

Rysunek 3. Wykres skupień

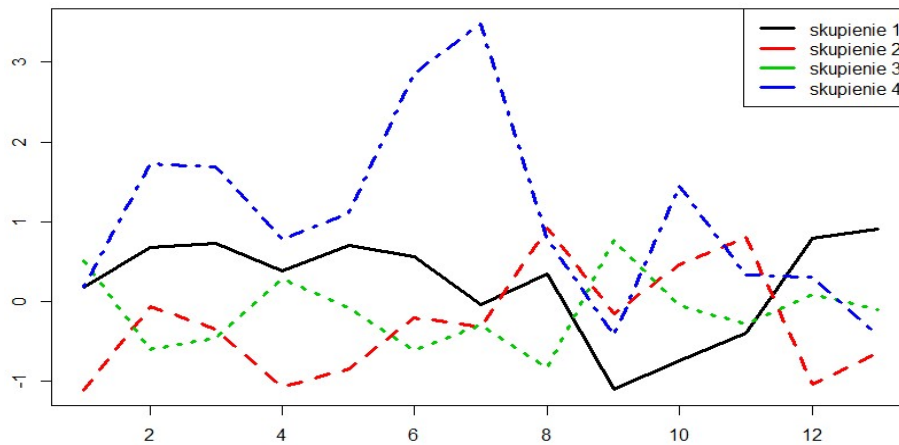


Źródło: opracowanie własne z zastosowaniem programu R

Widoczne jest, że województwo mazowieckie znacznie odbiega od pozostałych i pozostało skupieniem jednoelementowym.

Wykonano również wykres średnich przedstawiony na rysunku 4. Na osi x uwzględniono numery poszczególnych zmiennych, a na y – wartości średnich w tych skupieniach.

Rysunek 4. Wykres średnich skupień



Źródło: opracowanie własne z zastosowaniem programu R

Skupienie nr 1 to województwa dość silnie zurbanizowane. Posiadają wysoką liczbę przychodni, łóżek w hospicjach, zakładach opiekuńczo-leczniczych i pielęgnacyjno-opiekuńczych na 100 tys. ludności i kadry medycznej na 10 tys. mieszkańców. Zgrupowane obszary mają jednak jedne z najniższych wydatków powiatów i województw na służbę zdrowia. Wysokie wartości osiągnięto dla liczby łóżek szpitalnych i liczby szpitali.

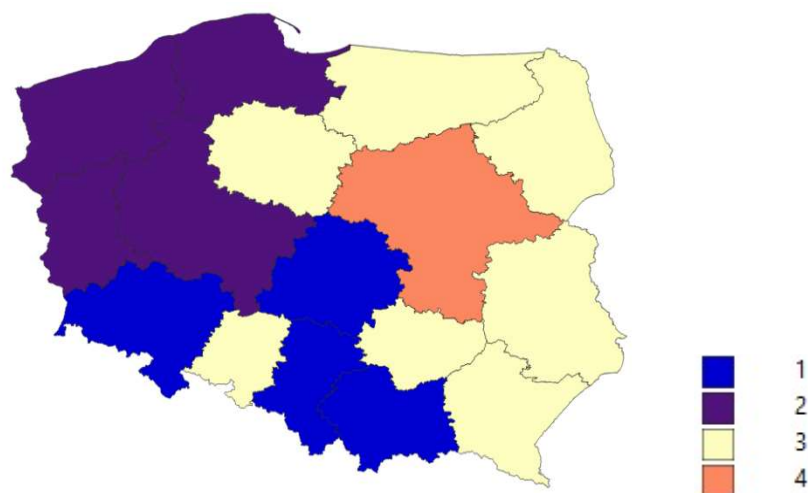
Skupienie nr 2 to obiekty o niewielkiej liczbie specjalistów, ale o dość wysokich wydatkach na służbę zdrowia. Szczególnie niewiele jest tam przychodni, pielęgniarek, łóżek szpitalnych i szpitali.

Skupienie nr 3 są to województwa cechujące się najwyższym odsetkiem ludności wiejskiej w ludności ogółem, według danych GUSu na 2019 rok. Posiadają najniższe liczby lekarzy i farmaceutów na 10 tys. osób, oraz najniższą wartość wydatków w gminach. Jednocześnie mają najwyższą liczbę przychodni na 10 tys. osób, oraz wydatków w powiatach.

Skupienie nr 4 to skupienie silnie zurbanizowane, znacznie odbiegające od pozostałych. Województwo mazowieckie w nim zawarte ma bardzo wysokie liczby lekarzy, farmaceutów i pielęgniarek na 10 tys. osób, dużą liczbę łóżek w hospicjach, zakładach opiekuńczo-leczniczych i pielęgnacyjno-opiekuńczych na 100 tys. ludności i bardzo wysokie wartości wydatków na służbę zdrowia w gminach.

Widoczne jest, że skupienie nr 4 posiada istotnie wyższe wartości od pozostałych, skupienie nr 1 cechuje się dość wysokimi wskaźnikami w zakresie kadr, liczby przychodni i łóżek, ale jest niedofinansowane. Skupienia 2 i 3 są do siebie zbliżone, ich wartości w przypadku większości zmiennych prezentują najniższe wartości w zbiorze.

Rysunek 5. Wizualizacja podziału skupień na mapie Polski



Źródło: opracowanie własne z zastosowaniem programu QGIS

Zgrupowane zostały województwa na wschodniej i zachodniej ścianie kraju, skupienia określono jako słabsze względem pozostałych w zbiorze. Istotnie różna jest sytuacja w województwie mazowieckim. Skupienie nr 1 zgrupowało województwa posiadające duże aglomeracje miejskie: Wrocław, Kraków, Łódź, gęsto zaludnione miasta Śląska. Skupienia 1 i 4 wyróżniają się wysokimi wartościami dotyczącymi kadry medycznej, łóżek w ośrodkach leczniczych, oraz liczbie przychodni. Wynika to w dużej mierze z występujących najludniejszych miast Polski, które są skupiskiem szpitali wojewódzkich i specjalistycznych. Ponadto miasta te, na uczelniach medycznych, kształcą większość polskiej kadry medycznej. Na podstawie wyników metody k-średnich umieszczono w skupieniu nr 3 województwa o najniższym odsetku urbanizacji. Cechują się one bardzo niską liczbą specjalistów. Według WHO główne problemy z jakimi borykają się mieszkańcy obszarów wiejskich to utrudniony dostęp do specjalistycznych usług zdrowotnych, niedobór pracowników służby zdrowia, odległość od centrów medycznych, gorsze wyposażenie oddziałów medycznych [WHO 2010], co potwierdza wyżej przedstawiona wizualizacja. Czynnikiem wpływającym na niższy poziom świadczeń u mieszkańców wsi jest mniejsza mobilność ludzi. Powodowana gorszym stanem dróg i mniej rozwiniętą komunikacją zbiorową [Ucieklak-Jeż, Bem 2017]. Liczba aptek ogólnodostępnych przypada na obszarach miejskich na 3 tys. osób, wiejskich – na 7 tys. osób. Z tego względu otwierane są punkty apteczne. Są to punkty obrotu detalicznego gotowych produktów leczniczych. [GUS 2017]. Punkty te mają na celu ułatwić dostępność do farmaceutyków mieszkańcom wsi. Głównym celem NFZ na lata 2019-2022 jest wzrost poziomu zaufania społecznego. Poprawy wymaga dostępność i jakość świadczeń, oraz rzetelna informacja o tym jak poruszać się w systemie ochrony zdrowia. NFZ ma zamiar podjąć działania służące wzrostowi pozytywnych doświadczeń pacjentów w związku z systemem opieki zdrowotnej, poprawić jakość i dostępność świadczeń, oraz upowszechnić wiedzę o zdrowiu. Wprowadzone mają zostać badania zadowolenia pacjentów mające dostarczyć informację o oczekiwaniach względem służby zdrowia. Jak można przeczytać w strategii rozwoju NFZ, planowane jest wdrożenie e-usług i popularyzacja profilaktyki [NFZ 2019]. Może to w dłuższym okresie zmniejszyć, w pewnych obszarach ochrony zdrowia, wykluczenie mieszkańców terenów wiejskich. Każde z badanych województw potrzebuje większego finansowania i poprawy jakości, a przede wszystkim dostępności usług. Pozytywną informacją jest to, że NFZ w swoich założeniach rozwojowych uwzględnia opinię pacjenta, chce sugerować się jego zdaniem. Wydaje się jednak, że sytuacja nie poprawi się bez wzrostu liczebności kadry medycznej i finansowania służby zdrowia.

BIBLIOGRAFIA

- Balicki A. (2013) Statystyczna analiza wielowymiarowa i jej zastosowania społecznoekonomiczne. Wydawnictwo Uniwersytetu Gdańskiego, 27-29.

- Dercz M., Izdebski H. (2001) Organizacja ochrony zdrowia w Rzeczypospolitej Polskiej w świetle obowiązującego ustawodawstwa. *Iuris*, Poznań, 18-22.
- Dobosz M. (2004) Wspomagana komputerowo statystyczna analiza wyników badań. *Akademicka Oficyna Wydawnicza EXIT*, 344-346.
- Europejski Konsumencki Indeks Zdrowia. (2016), *Health Consumer Powerhouse*. 18.
- Golinowska, S. (2012) Zarys systemu ochrony zdrowia. *Drukarnia Open Eyes*, 31-46, 81-112.
- Indeks Sprawności Ochrony Zdrowia. (2019) *PwC*, dane-i-analizy.pl, 16-32.
- Libura M. i in. (2018) Finansowanie i organizacja systemu ochrony zdrowia w oczach pacjentów. 22-44.
- Lasek M. (2002) Data Mining. Zastosowania w analizach i ocenach klientów bankowych. *Oficyna Wydawnicza „Zarządzanie i Finanse”*, Warszawa, 125-126.
- Magda I., Szczygielski K. (2011) Ocena możliwości poprawy działania polskiego systemu ochrony zdrowia. *Ernst & Young*, 12-24.
- Ostasiewicz W. (red.) (1998) *Statystyczne metody analizy danych*. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, 96-116.
- Raport: system ochrony zdrowia w polsce – stan obecny i pożądane kierunki zmian. (2018), 49.
- Stanimir A. (2006) *Analiza danych marketingowych. Problemy, metody, przykłady*. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, 121.
- Stanisz A. (2007) *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 3. Analizy wielowymiarowe*. Statsoft, 118-127.
- Strategia Narodowego Funduszu Zdrowia na lata 2019-2022. (2019), *NFZ*, 1-27.
- Ucieklak-Jeż P., Bem A. (2017) Dostępność opieki zdrowotnej na obszarach wiejskich w Polsce. 1-15.
- Ustawa z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych. (Dz. U. z 2018 r. poz. 1510, ze zm.).
- Ustawa z dnia 23 marca 2017 r. o zmianie ustawy o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych. (Dz. U. poz. 844).
- Ustawa z dnia 22 lipca 2014 r. o zmianie ustawy o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych oraz niektórych innych ustaw. (Dz. U. poz. 1138, ze zm.).
- Ustawa z dnia 27 sierpnia 2009 r. o finansach publicznych. (Dz. U. z 2019 r. poz. 869).
- World Health Organization (2010) *World Health Report, 2010: Health Systems Financing the Path to Universal Coverage*. 18-20.
- Zdrowie i ochrona zdrowia w 2016 r. (2017) *GUS*, 30-111.
- www.ec.europa.eu/health/indicators_data/indicators_en [dostęp: 02.08.2020].
- www.stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT# [dostęp: 02.08.2020].

COMPARATIVE ANALYSIS OF THE STATE OF HEALTH CARE IN VOIVODESHIPS

Abstract: : The article compares the state of health care in selected thematic areas in voivodeships. It focuses on 2 main aspects, such as: health care expenses and the number of specialists. The analysis was carried out on data from 2018-2019. Its aim was to group voivodships according to selected variables regarding the state of health care. Their geographical location and demography were analyzed. The difference between the state of health care in urban and rural areas was featured.

Keywords: cluster analysis, k-means method, medical staff, health care, rural areas

JEL classification: C38

NIEPEWNOŚĆ NAUKI A DZIAŁANIE, CZYLI – PROBLEMY Z EVIDENCE – BASED POLICY

Grzegorz M. Malinowski  <https://orcid.org/0000-0002-4152-1625>

Akademia Leona Koźmińskiego w Warszawie

e-mail: gmalinowski@kozminski.edu.pl

Streszczenie: Celem niniejszego artykułu jest przede wszystkim wprowadzenie tematyki niepewności naukowej do szerszego kontekstu ekonomii i zarządzania. Naukowa niepewność jest jednym z przejawów nieredukowalnej niepewności i refleksja nad nią powinna umożliwić podejmowanie lepszych decyzji. Podmiot, który opiera swoje działanie na aktualnych badaniach naukowych, które z czasem ulegają deprecjacji i prowadzą ostatecznie do błędnych decyzji, zostaje określony mianem „fajera”. W tekście zostają wskazane oszacowania dotyczące potencjalnej skali tego problemu uzupełnione o zarys socjologicznych trudności, zidentyfikowanych w ramach analizy procesu budowania twierdzeń naukowych. Artykuł zakończony jest szkicem odpowiedzi na pytanie „jak działać w kontekście naukowej niepewności?”

Słowa kluczowe: niepewność nauki, nieredukowalna niepewność, problem fajera, błędy statystyczne, zasada ostrożności

JEL classification: D01, D81, H12

WPROWADZENIE

Fenomen nauki jest niewątpliwie jednym z największych osiągnięć ludzkości. Postęp technologiczny oraz poprawiające się warunki życia powodują, że nie tylko same wytwory naukowych przedsięwzięć, lecz także, a może nawet przede wszystkim naukowy sposób myślenia i, tak zwana metoda naukowa stały się standardem, któremu starają się sprostać różne grupy społeczne - badacze, myśliciele, politycy czy przedsiębiorcy.

Upowszechnienie metody naukowej trwa nieprzerwanie od czasów oświecenia i zajmuje coraz to nowe obszary zdominowane wcześniej przez mity, religie, wyobrażenia ludowe i tradycje. W konsekwencji wykrystalizował się

<https://doi.org/10.22630/MIBE.2020.21.1.6>

fenomen **działania opartego na faktach**, którego przejawem są, chociażby *evidence – based medicine* lub *evidence – based policy*. Podstawowym założeniem tego podejścia jest twierdzenie, że działanie musi opierać się na wiarygodnych teoriach naukowych, których skuteczność i bezpieczeństwo zostało potwierdzone odpowiednimi badaniami empirycznymi.

Powstaje jednak pewna trudność. Mianowicie metodologiczna specyfika „podejścia naukowego” wiąże się bezpośrednio z jego zasadniczym celem, którym jest „poznanie prawdy”, nie zaś rozwiązywanie konkretnych problemów. Innymi słowy, metoda naukowa nakierowana jest na ustalenie statusu logicznego (prawda / fałsz) stawianych hipotez badawczych, a nie na skuteczność działania.

Istota tej różnicy staje się łatwa do zrozumienia wówczas, gdy rozpatruje się ją przez pryzmat „**naukowej niepewności**”, czyli zmienności i ewolucji samych teorii naukowych. W ujęciu historycznym teorie naukowe podlegają przeobrażeniom, czasami dość znacznym, a niekiedy zostają w całości bądź w części sfalsyfikowane. Wynika z tego, że dysponujemy niepełnym obrazem świata, i chociaż wiemy o nim coraz więcej to jednak z punktu widzenia podmiotu podejmującego decyzje to, czego jeszcze nie wiemy, może mieć o wiele istotniejsze znaczenie od tego, co już udało nam się ustalić.

Prowadzi to do osobliwości, którą określam mianem „**problemu frajera**”. Otóż frajer, to ktoś, kto DZIAŁA w oparciu o pewną teorię, koncepcję lub model naukowy, która z czasem okazuje się błędna lub niekompletna, natomiast skutki działań frajera są dla niego opłakane. Najlepszą, choć nieco populistyczną ilustracją tego fenomenu jest przykład frankowiczów - a więc kredytobiorców, którzy zadłużyli się we frankach szwajcarskich, gdyż modele szacujące ryzyko wskazywały, że takie rozwiązanie jest bezpieczne (a zarazem nie brakowało ekspertów podkreślających optymalność takiego działania), a następnie spotkał ich spory zawód oraz olbrzymie nieszczęście finansowe.

W tym kontekście warto zwrócić uwagę na nie lubiany przez metodologów podział nauki na *nauki podstawowe* oraz *nauki stosowane*. Istotą tej klasyfikacji jest konstatacja, że fundamentalnym celem jednych dyscyplin naukowych jest odkrywanie prawdy a innych - przeobrażanie rzeczywistości w pożądanym kierunku.

Należy zauważyć, że „problem frajera” praktycznie nie dotyczy nauk podstawowych! Gdyby bowiem dotyczył, wówczas na przykład zwolennicy „teorii płaskości Ziemi”, „kreacjonizmu”, tezy, że ziemia ma najwyżej kilka tysięcy lat, lub kontestatorzy „teorii grawitacji”, „mechaniki kwantowej” czy „dryfu kontynentalnego” – na gruncie swych ewidentnie błędnych przekonań, podejmowałiby fundamentalnie błędne decyzje. Tak natomiast nie jest. Osoby takie zakładają „normalne” rodziny i zarabiają „normalne pieniądze”. Ich ignorancja naukowa wcale nie przekłada się na ignorancję w działaniu.

Inaczej sprawa się przedstawia w naukach stosowanych. Koncepcje powstałe na gruncie takich dyscyplin jak: psychologia, medycyna, dietetyka, socjologia, ekonomia, zarządzanie czy finanse, są podstawą konkretnych działań rządów

państw, inwestorów, polityków gospodarczych, instytucji finansowych oraz zwykłych konsumentów. W tych obszarach „problem frajera” potencjalnie uwidacznia się w całej rozciągłości. Należy bowiem mieć na uwadze fakt, że podstawa naukowa tych decyzji wraz z upływem czasu może się zmienić lub zostać odrzucona a ponadto w przyszłości mogą uwidocznic się skutki uboczne związane z danym postępowaniem, o których wcześniej nie wiedzano.

Powyższe ustalenia prowadzą do konkluzji, że wielkie teorie naukowe, powstałe na gruncie nauk podstawowych, jak teoria ewolucji, czy mechanika kwantowa rozwijają naszą wiedzę o świecie i są bezpośrednim lub pośrednim źródłem wielu innowacji poprawiających ludzką egzystencję, ale nie są podstawą codziennych decyzji – takich, od których uzależnione jest przetrwanie rozumiane zarówno w kategoriach biologicznych, jak również biznesowych - gospodarczych.

Powstaje zatem istotny problem, z którym musi borykać się każdy podmiot podejmujący autonomiczne decyzje. Z jednej strony powszechnie uważa się, że podejmowane decyzje są tym skuteczniejsze, im lepsza jest ich podstawa naukowa, a z drugiej strony pojawia się naukowa niepewność, którą decydent musi brać pod uwagę.

Celem niniejszego artykułu jest przede wszystkim wprowadzenie tematyki niepewności naukowej do szerszego kontekstu ekonomii i zarządzania. Naukowa niepewność jest jednym z przejawów nieredukowalnej niepewności [Kozłowski, 2008] i refleksja nad nią powinna umożliwić podejmowanie lepszych decyzji. Z tego względu drugim celem artykułu jest próba naszkicowania odpowiedzi na pytanie „jak działać w kontekście naukowej niepewności?”.

Pierwsza część artykułu wprowadza do zagadnienia niepewności naukowej. W paragrafie pierwszym opisana zostaje istota tego zagadnienia na poziomie statystycznym, której zwieńczeniem jest wstępne oszacowanie wskaźnika wartości predykcyjnej netto. W dalszej części wskazane zostają podstawowe ograniczenia i słabości podejmowania decyzji na podstawie koncepcji naukowych (*evidence based policy*). Na końcu tej części tekstu znalazła się próba uwzględnienia tych ograniczeń w ramach ilościowej analizy wiarygodności badań empirycznych. W drugiej części pracy wątek naukowej niepewności zostaje umieszczony w kontekście dwóch innych fenomenów, a mianowicie „zależności od ścieżki” oraz „ryzyka katastrofy”, co umożliwia wyciągnięcie ostrożnego wniosku odnośnie do normatywnych zaleceń związanych z podejmowaniem decyzji w warunkach naukowej niepewności.

DLACZEGO NIE MOŻNA OPIERAĆ SIĘ NA „NAJNOWSZYCH WYNIKACH BADAŃ”?

Wyobraźmy sobie, że próbujemy zweryfikować prawdziwość 1000 hipotez badawczych. Nie jest istotna ani treść tych hipotez, ani też problematyka, jakiej dotyczą. Załóżmy również, że skądinąd wiadomo, iż wśród tych hipotez jedynie 100

jest prawdziwych. Pozostałe zaś 900 to hipotezy fałszywe [The Economist 2013]. Procedura weryfikacji hipotez oparta jest na wykorzystaniu narzędzi statystycznych. W procesie tym badacz narażony jest na popełnienie dwóch rodzajów błędów. Może popełnić **błąd I rodzaju** (α), polegający na tym, że fałszywa hipoteza zostanie uznana za prawdziwą¹ (tak zwany *false positive*). Może także uznać prawdziwą hipotezę za fałszywą i tym samym popełnić **błąd II rodzaju** (β)² (tak zwany *false negative*).

Tabela 1. Typy błędów statystycznych

Typy błędów statystycznych		Hipoteza H_0 jest:	
		Prawdziwa	Fałszywa
Decyzja odnośnie do hipotezy H_0	Odrzucić	Błąd I rodzaju	Wniosek prawidłowy
	Nie odrzucać	Wniosek prawidłowy	Błąd II rodzaju

Źródło: opracowanie własne

Wśród osób, które profesjonalnie nie zajmują się prowadzeniem badań naukowych powszechny jest pogląd, że weryfikacja hipotez naukowych nakierowana jest na równomierną minimalizację prawdopodobieństwa popełnienia obu tych błędów. Tak jednak nie jest. Prawdopodobieństwo popełnienia błędu I rodzaju oraz prawdopodobieństwo popełnienia błędu II rodzaju są z sobą powiązane matematycznie w taki sposób, że zmniejszenie prawdopodobieństwa popełnienia jednego typu błędu automatycznie powoduje zwiększenie prawdopodobieństwa błędu drugiego typu. A zatem minimalizując prawdopodobieństwo popełnienia błędu I rodzaju, zwiększamy prawdopodobieństwo popełnienia błędu II rodzaju i odwrotnie.

Wprawdzie w ramach każdej procedury badawczej pragnie się uniknąć obu typów pomyłek, to jednak pierwszorzędnym celem każdego przedsięwzięcia naukowego jest poznanie prawdy i unikanie fałszu. Brzmi to wprawdzie nieco górnolotnie, ale wyraża się w prostym postępowaniu, mianowicie każdy rzetelny badacz najbardziej obawia się popełnić błąd I rodzaju. Jest tak dlatego, ponieważ badacz pragnący rozbudowywać wiedzę o świecie znacznie bardziej obawia się sytuacji, w której do istniejącego gmachu twierdzeń naukowych doda jakąś cegielkę – hipotezę, która z czasem okaże się fałszywa niż sytuacji, w której przeoczy prawdziwość jakiejś hipotezy. Innymi słowy, lepiej jest nie dodawać nic, aniżeli dorzucić coś, co jest nieprawdziwe. Takie postawienie sprawy jest zupełnie racjonalne, jeżeli założymy, że nadrzędnym celem jest poszerzenie wiedzy o świecie. W dalszej części artykułu zostanie jednak wykazane, że racjonalność

¹ Dzieje się tak wtedy, gdy zostanie odrzucona prawdziwa hipoteza H_0 stwierdzająca brak zależności między badanymi zjawiskami.

² Dzieje się tak wtedy, gdy fałszywa hipoteza H_0 stwierdzająca brak zależności między badanymi zjawiskami nie zostaje odrzucona.

podejmowania decyzji pozostaje w pewnej sprzeczności z tak rozumianą racjonalnością naukową.

Swego rodzaju standardem jest przyjęcie wartości 5%, jako maksymalnego, dopuszczalnego prawdopodobieństwa popełnienia błędu I rodzaju. Z drugiej zaś strony maksymalne, dopuszczalne prawdopodobieństwo popełnienia błędu II rodzaju wynosi 20%. Te statystyczne warunki brzegowe sprawiają, że procedura testowania 1000 hipotez może wygenerować 45 hipotez nieprawdziwych, które uznane zostaną za prawdziwe (*false positives*) oraz 20 hipotez prawdziwych, które zostaną uznane za fałszywe (*false negatives*). W rezultacie pozytywnie zweryfikowanych zostanie 125 hipotez, w oparciu o które zwolennicy działania opartego na wiedzy naukowej będą podejmowali decyzje. Pamiętać jednak należy, że wśród tych 125 hipotez aż 45 (36%) może być fałszywych.

Wyobraźmy sobie, że testujemy 2400 substancji pod kątem kancerogenności. Załóżmy, że 40 procent z nich ma właściwości kancerogenne, 36 procent nie ma takich właściwości, natomiast co do pozostałych 24 procent nie mamy pewności. Otóż stosując się do omówionych standardów statystycznych, pracując na odpowiednio licznej próbie, badacz uzna 43 substancje za kancerogenne, chociaż takie nie są, a 192 substancje uzna za bezpieczne, chociaż w rzeczywistości są kancerogenne. Skutek: w użyciu pojawi się prawie 200 niebezpiecznych substancji, a 43 będą niesłusznie niedopuszczone do użytku [Lemons, Shrader, Cranor 1997].

Czy te spostrzeżenia deprecjonują osiągnięcia nauki? Z całą mocą należy stwierdzić, że nie. Nie ulega jednak wątpliwości, że opisana prawidłowość wskazuje na to, iż w samym fundamencie badania naukowego (posługującego się wnioskowaniem statystycznym) tkwi pewna nieusuwalna, nieredukowalna niepewność [Weiss 2003, Lo 2009], a identyfikacja tego faktu prowadzi do kilku wniosków:

- Hipotezy naukowe muszą podlegać weryfikacji wielokrotnie;
- Czas jest najbardziej wiarygodnym kryterium ustalania prawdziwości i fałszywości hipotez badawczych (ciągłe powtarzanie badań ostatecznie prowadzi do eliminacji fałszywych hipotez);
- Podejmowanie decyzji oparte na wiedzy naukowej nie może być jedynym kryterium skutecznego działania.

DLACZEGO NIE MOŻNA UFAĆ WYNIKOM BADAŃ EMPIRYCZNYCH?

Ponieważ większość opublikowanych odkryć naukowych opartych na badaniach empirycznych jest fałszywa. Do takiego, zdumiewającego wniosku dochodzi J. Ioannidis [2005] w odniesieniu do nauk medycznych. Zasugerowany przez niego sposób rozumowania można jednak także przenieść na grunt nauk społecznych.

Kiedy uczony na gruncie badań stwierdza istnienie jakiejś zależności w ramach reprezentowanej przez siebie dziedziny, to to, czy taka relacja rzeczywiście występuje, jest uzależnione od:

- A. **Prawdopodobieństwa *a priori*** (przed faktem) – to prawdopodobieństwo odzwierciedlające wiedzę badacza odnośnie do analizowanej rzeczywistości jeszcze przed wykonaniem docelowych eksperymentów lub obserwacji;
- B. **Mocy testu statystycznego** – to prawdopodobieństwo niepopelnienia błędu II rodzaju. Moc jest zatem dopełnieniem prawdopodobieństwa popelnienia błędu II rodzaju, czyli $1 - \beta$;
- C. **Poziomu istotności statystycznej** – to maksymalne dopuszczalne prawdopodobieństwo popelnienia błędu I rodzaju.

Aby dowieść, że większość badań empirycznych jest fałszywa, należy wprowadzić współczynnik R , który jest ilorazem odkryć prawdziwych i odkryć nieprawdziwych.

$$R = \frac{\text{odkrycia prawdziwe } (T)}{\text{odkrycia nieprawdziwe } (F)}$$

Warto także podkreślić, że współczynnik R jest charakterystyczny dla danej dyscypliny, i może przybierać różne wartości. Zasadniczo jednak postęp nauki zakłada, że $R \leq 1$, a zatem ilość odkryć prawdziwych jest mniejsza od ilości odkryć fałszywych.

Tabela 2. Badania empiryczne – odkrycia prawdziwe

Werdykt badacza	Rzeczywistość		Suma
	Hipoteza jest prawdziwa	Hipoteza jest nieprawdziwa	
Hipoteza jest prawdziwa	$c(1 - \beta)R/(R + 1)$	$c\alpha/(R + 1)$	$c(R + \alpha - \beta R)/(R + 1)$
Hipoteza jest nieprawdziwa	$c\beta R/(R + 1)$	$c(1 - \alpha)/(R + 1)$	$c(1 - \alpha + \beta R)/(R + 1)$
Suma	$cR/(R + 1)$	$c/(R + 1)$	c

Źródło: opracowanie własne na podstawie Ioannidis [2005]

Na gruncie powyższych ustaleń, a także zakładając, że parametr „ c ” oznacza ilość nowych odkryć zawartych w publikacjach naukowych, możliwe jest wskazanie następujących zależności (Tabela 2):

$R/(R + 1)$ – Prawdopodobieństwo *a priori* tego, że badana hipoteza jest prawdziwa³.

$(1 - \beta)$ – Moc testu.

α – prawdopodobieństwo popelnienia błędu I rodzaju.

³ Jeżeli $R = \frac{T}{F}$, to prawdopodobieństwo, że dana hipoteza jest prawdziwa $P(T)$ jest równe $P(T) = \frac{T}{T+F}$, co po przekształceniu daje wyrażenie $P(T) = \frac{R}{(R+1)}$

$c(1 - \beta)R/(R + 1)$ – liczba dokonanych odkryć naukowych, które są prawdziwe.
 $c\alpha/(R + 1)$ – liczba dokonanych odkryć naukowych, które są nieprawdziwe (błędy I rodzaju).

$c\beta R/(R + 1)$ – liczba przeoczonych odkryć naukowych (błędy II rodzaju).

$c(1 - \alpha)/(R + 1)$ – liczba odrzuconych hipotez, które w istocie są nieprawdziwe.

Z punktu widzenia celu niniejszego artykułu najbardziej istotne są przypadki błędów I rodzaju, a więc stwierdzenie istnienia pewnej zależności tam, gdzie ona nie istnieje. Z tego punktu widzenia należy zbadać stosunek: $\frac{D'}{D'+E'}$. Wyrażenie należy interpretować jako stosunek prawidłowo rozpoznanych zależności do wszystkich poczynionych (opublikowanych prawdziwych i fałszywych) odkryć naukowych. Jest to zatem wskaźnik mierzący wartość predykcijną danego odkrycia. W dalszej części określać go będą mianem wskaźnika Wartości Predykcyjnej Netto (WPN). Uzyskujemy:

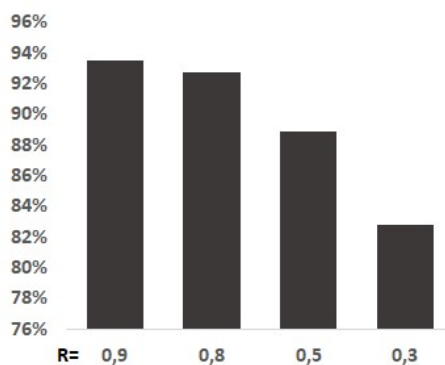
$$WPN = \frac{c(1 - \beta)R/(R + 1)}{c(R + \alpha - \beta R)/(R + 1)}$$

A po przekształceniu:

$$WPN = \frac{(1 - \beta)R}{(1 - \beta)R + \alpha}$$

Zasadnicze pytanie, jakie należy sobie postawić dotyczy „typowej” wartości wskaźnika WPN w badaniach z zakresu nauk społecznych. Otóż swego rodzaju standardem jest przyjęcie, że $\alpha = 0,05$ a $\beta = 0,2$. W zależności od przyjętej wartości współczynnika R, wartość WPN kształtuje się na poziomie między 94% a 83% (Tabela 3).

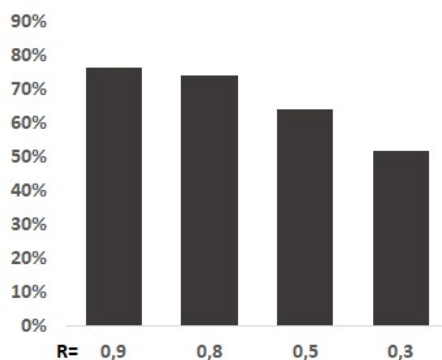
Rysunek 1. Wartość WPN przy różnych poziomach R.



Źródło: opracowanie własne

Należy jednak zaznaczyć, że wartość β często nie jest weryfikowana przez badaczy. Analiza 159 meta – analiz artykułów z zakresu ekonomii prowadzi do wniosku, że najbardziej typowy poziom mocy statystycznej, która je charakteryzuje nie przekracza 18% [Ioannidis, Stanley, Doucouliagos 2017]. Uwzględnienie tego faktu powoduje, że w zależności od przyjętej wartości współczynnika R, wartość WPN oscylują między 76% a 52%.

Rysunek 2. Wartość WPN przy różnych poziomach R ($1 - \beta = 18\%$)



Źródło: Opracowanie własne

Na tym etapie rozważań można więc wyciągnąć wniosek, że prawdopodobieństwo tego, że stwierdzona, istotna statystycznie zależność jest prawdziwa może wahać się pomiędzy 95% a 50%. W dalszej części artykułu zostaną jeszcze wskazane okoliczności, które skutkują znaczącym zmniejszeniem tej wartości.

PROBLEMY Z NAUKĄ

Naukowa niepewność oznacza dla decydenta zasadniczy problem, ponieważ rzuca cień na skuteczność działania opartego na faktach, a więc na dowodach i koncepcjach naukowych, przez co jest bezpośrednio powiązana z „problemem frajera”.

Nie tylko jednak względy statystyczne sprawiają, że podejmowaniu decyzji bazujących na dowodach naukowych towarzyszy pewna fundamentalna, nieredukowalna niepewność.

Otóż podmiot podejmujący decyzje, który pragnie oprzeć swoje postępowanie na wiedzy pochodzącej z opublikowanych badań, musi zmierzyć się z następującymi trudnościami:

- a. Nauka potrzebuje czasu – dowód naukowy często pojawia się za późno;

- b. Racjonalność badacza – naukowiec najbardziej boi się nie tego, czego obawia się decydecja;
- c. System zachęt i publikacje – brak racji może naruszyć „ego” badacza, brak publikacji spycha go w niebyt;
- d. Stronniczość badań – podmiot finansujący badania spodziewa się określonych wyników.

Pierwszą kwestią (a) jest związek czasu oraz dowodu naukowego. Postęp naukowy, w ujęciu falsyfikacjonizmu odbywa się dzięki dwóm czynnikom. (1) Stawianiu śmiałych hipotez naukowych oraz (2) falsyfikacji tychże. Te dwa równoległe procesy toczą się w czasie którego upływ staje się pewnego rodzaju gwarancją tego, że hipotezy błędne zostaną odrzucone.

W tym kontekście należy stwierdzić, że niemal wszelkiego rodzaju działaniom towarzyszą różne skutki uboczne, których istnienie oraz natężenie są trudne do natychmiastowego stwierdzenia. Często potrzeba wielu lat, aby je zaobserwować. Ponadto, może pojawić się sytuacja, w której badany system jest unikatowy i nie jest możliwe utworzenie niezbędnej próby badawczej. Jeśli na przykład pojawia się hipoteza, że nadmierna emisja dwutlenku węgla może doprowadzić do katastrofalnych zmian klimatycznych, to należy stwierdzić, iż jej prawdziwość jest niemożliwa do udowodnienia, gdyż klimat możliwy do badania jest tylko jeden, a dane historyczne, jakimi dysponujemy, są zbyt skąpe, aby dostarczyć jednoznacznych konkluzji. Podkreślenia wymaga fakt, że obiektywnym, empirycznym dowodem będziemy dysponowali dopiero wówczas, gdy nastąpi klimatyczna katastrofa. Dowód naukowy pojawi się więc za późno.

Zagadnienie to, staje się szczególnie istotne w kontekście rozważanego „problemu frajera”. Historyczna refleksja nad przypadkami, w których brak działań prewencyjnych usprawiedliwiany był brakiem dowodów naukowych, bądź istniejące dowody naukowe sugerowały, że ostrożność jest nieusprawiedliwiona - prowadzi do wniosku, że podmiot podejmujący decyzje powinien zawsze stawiać pytania „co, jeśli dana teoria okaże się błędna?” oraz „czy skutki działania opartego na błędnej teorii mogą być gorsze od niestosowania się do rekomendowanych?”, i niejednokrotnie musi podejmować decyzje pomimo braku dowodów naukowych, lub nawet wbrew dowodom naukowym, jeśli nie są one dostatecznie pewne [Ahteensuu 2013; Malinowski 2018].

Przykładów jest wiele: margaryna, Talidomid, azbest, glifosat, promieniowanie rentgenowskie a ostatnio nawet kwestia badań mammograficznych. Jeszcze do niedawna medycyna dysponowała licznymi dowodami naukowymi wskazującymi na neutralny lub nawet pozytywny wpływ tych substancji / działań na zdrowie, wraz z upływem czasu pojawiło się jednak wiele świadectw obalających wcześniejsze ustalenia. Podobne trudności wiążą się współcześnie z udowodnieniem szkodliwości palenia e-papierosów. Są one dostępne na rynku stosunkowo od niedawna, zatem ewentualny dowód na ich szkodliwość poznamy za kilkadziesiąt lat, kiedy badacze będą dysponowali odpowiednim zapisem historycznym.

Druga kwestia (b) dotyczy racjonalności badacza. Otóż należy stwierdzić, że działalności naukowej towarzyszy nieco inny typ motywacji, niż to ma miejsce w podejmowaniu decyzji w codziennym życiu. Ma to istotne znaczenie z punktu widzenia realizacji celu danego działania.

Załóżmy, że w niedalekiej okolicy osiedla mieszkaniowego powstało przedsiębiorstwo – fabryka, które emituje do atmosfery pewien związek chemiczny „X”. Dodajmy także, że związek ten jeszcze dotychczas nigdy nie był stosowany na tak dużą skalę. Mieszkańcy osiedla są narażeni na wdychanie oparów tego specyfiku, dlatego też zostaje podjęta decyzja o przebadaniu substancji „X” pod kątem wpływu na zdrowie ludzkie.

Badacze, którzy podjęli się tego działania, postępują zgodnie z procedurą badawczą:

1. Stawiają hipotezę H_0 , która brzmi „substancja X jest obojętna dla organizmu ludzkiego”;
2. Formułują hipotezę alternatywną H_A , która brzmi „substancja X jest szkodliwa dla organizmu ludzkiego”.

Następnie rozpoczyna się proces testowania hipotez, w którym hipoteza H_0 może zostać odrzucona lub nie. Wiadomo już jednak, że procedura ta związana jest z prawdopodobieństwem popełnienia błędu I i II rodzaju.

Tutaj właśnie pojawia się sedno problemu. Podstawowym celem badaczy jest poznanie prawdy, a nie bezpieczeństwo mieszkańców. Te dwa cele zazwyczaj są ze sobą zgodne, ale w rozważanej sytuacji pojawia się między nimi pewne napięcie. Otóż naukowcy przede wszystkim nie chcą popełnić błędu I rodzaju i uznać za prawdę coś, co prawdą nie jest. Z tego względu akceptują wyższe prawdopodobieństwo błędu II rodzaju. W analizowanym przykładzie oznacza to, że **badacze najbardziej obawiają się uznania substancji „X” za niebezpieczną w sytuacji, gdy w rzeczywistości jest ona obojętna dla zdrowia.** Inaczej jednak wygląda perspektywa mieszkańców osiedla. Wprawdzie mogą oni także wysoko cenić prawdę naukową, ale najprawdopodobniej wyżej stawiają swoje bezpieczeństwo i wolą być ostrożni. A skoro tak, to oznacza, że **najbardziej obawiają się popełnienia błędu II rodzaju (a nie I), czyli uznania substancji „X” za bezpieczną w sytuacji, gdy taką ona nie jest.** Dlatego zapewne zaakceptowałyby oni wyższe prawdopodobieństwo błędu I rodzaju w zamian za minimalizację prawdopodobieństwa błędu II rodzaju.

Tabela 3. Błędy statystyczne – przykład

Decyzja:	Stan problemu w rzeczywistości:	
	H_0 jest prawdziwa (substancja jest obojętna)	H_0 jest fałszywa H_A jest prawdziwa (substancja jest szkodliwa)
Odrzucenie H_0 Uznanie, że substancja jest szkodliwa	Błąd I-ego rodzaju (<i>false positive</i>) Odrzucenie hipotezy zerowej H_0 , która w rzeczywistości jest prawdziwa Przyjęcie za prawdziwą hipotezy alternatywnej H_A , która w rzeczywistości jest fałszywa	decyzja słuszna
Przyjęcie H_0 Uznanie, że substancja jest obojętna	decyzja słuszna	Błąd II-ego rodzaju (<i>false negative</i>) Przyjęcie za prawdziwą hipotezy zerowej H_0 , która w rzeczywistości jest fałszywa

Błąd II-ego rodzaju - poważniejszy dla mieszkańców

Błąd I-ego rodzaju - ważniejszy dla badaczy

Źródło: opracowanie własne

W tym kontekście wydaje się uprawnione stwierdzenie, że badacz stosuje w praktyce nieco inną racjonalność od osoby, dla której błędny rezultat badania wiąże się z poważnym niebezpieczeństwem. Badacz jest tropicielem prawdy, a nie stróżem bezpieczeństwa. Tymczasem działający człowiek zazwyczaj nad wszystko inne przedkłada bezpieczeństwo własne lub wspólnoty, do której należy.

Procedurę badawczą cechuje nacisk na restrykcyjność testu kosztem wrażliwości testu. Z tego właśnie względu słabe lub wczesne sygnały ostrzegawcze zostają uznane za „nieistotne statystycznie” i dopiero wraz z upływem czasu może okazać się, że zdrowie publiczne zostało narażone, ponieważ metody statystyczne nie były w stanie zidentyfikować zagrożenia odpowiednio wcześniej. W dodatku, jeśli przyjmemy, że typowa moc statystyczna badań empirycznych publikowanych przez ekonomistów wynosi 18%, wówczas prawdopodobieństwo popełnienia błędu II rodzaju okaże się gigantyczne, a podejmowanie decyzji na podstawie takich badań staje się przedsięwzięciem niezwykle ryzykownym.

Kolejna kwestia (c) dotyczy samego procesu badawczego. Proces projektowania, wykonywania i interpretowania badań naukowych, który wypracowany został w ramach tak zwanej „metody naukowej”, jest jednym z największych osiągnięć ludzkości i jednocześnie gwarantem obiektywizmu i rzetelności. Tak rozumiana nauka jest jednak także fenomenem socjologicznym, i jako taka – może być analizowana pod kątem wiarygodności czy tendencyjności badań. W ten sposób rozumiana meta – analiza pozwala uchwycić kilka

problematicznych aspektów funkcjonowania szeroko rozumianego „środowiska naukowego”.

Zacząć należy od tego, że system zachęt stosowany do oceny dorobku naukowego poszczególnych uczonych na samym szczycie hierarchii umieszcza publikacje naukowe – liczy się zarówno ich ilość, jak również ich jakość, której miarą jest chociażby miejsce na tzw. liście filadelfijskiej. Otóż zaobserwowano, że żurnale naukowe znacznie chętniej publikują badania będące weryfikacją hipotez nowych i oryginalnych, aniżeli analizy i powtórzenia istniejących już hipotez, koncepcji itp. Innymi słowy, czasopismo troszczące się o swoją renomę pragnie prezentować najnowsze i najciekawsze odkrycia. Takie podejście rodzi jednak co najmniej jeden problem. Mianowicie generuje silny bodziec do znajdowania prawidłowości i związków przyczynowych nawet tam, gdzie ich nie ma.

W konsekwencji już na etapie projektowania badania, uczony jest wystawiony na silną pokusę. Może bowiem w taki sposób skonstruować swoją eksplorację, aby odznaczała się zwiększonym prawdopodobieństwem wystąpienia istotnych statystycznie - błędów I rodzaju. W ten sposób uzyskuje wynik oryginalny i nadający się do publikacji. A to się liczy w środowisku, którym rządzi zasada *publish or perish*.

Medialną egzemplifikacją tego typu działania jest słynna prowokacja, która zyskała duży rozgłos, a która ma szczególnie istotne znaczenie w kontekście rozważanego w tym tekście „problemu frajera”. Otóż w 2015 roku światową prasę obiegła informacja, że „czekolada pomaga schudnąć” [Bohannon et al. 2015]⁴. Wniosek ten został wyprowadzony na podstawie badań wykonanych przez „Instytut Zdrowia i Diety”. Badanie przeprowadzone zostało na trzech grupach. Jedna grupa przez pewien okres poddana była diecie niskowęglowodanowej, druga grupa stosowała tę samą dietę z tą jednak różnicą, że każdy uczestnik miał codziennie zjadać tabliczkę czekolady. Trzecia grupa – grupa kontrolna składała się z osób, które miały trwać przy dotychczasowych nawykach żywieniowych. Wyniki badania były niezwykle interesujące, gdyż okazało się, że czekolada nie tylko przyspiesza proces utraty wagi, ale również pozytywnie wpływa na poziom cholesterolu i satysfakcję życiową.

Badanie było jednak zwykłą manipulacją. Jego autor w sposób świadomy w taki sposób zaprojektował badanie, aby dało ono rezultat istotny statystycznie i sprawiało wrażenie rzetelnej nauki. Wnikliwa analiza tekstu wzbudza jednak wątpliwości:

- Posłużono się małą próbą badawczą (każda grupa uczestników składała się z 5 osób);

⁴ Publikacja ta była w swej istocie prowokacją naukową. Artykuł pojawił się na stronach internetowych prestiżowego czasopisma medycznego. Po ujawnieniu motywów autorów tekstu – artykuł został usunięty.

- U każdego uczestnika eksperymentu weryfikowano aż 18 różnych zmiennych min. poziom cholesterolu, parametry krwi, jakość snu, satysfakcję z życia itp.

Jeżeli niewielka próba badawcza zostanie oceniona pod kątem 18 różnych kryteriów, wówczas prawdopodobieństwo uzyskania statystycznie istotnego błędu I rodzaju wynosi około 60%. Przy tak małej próbie badawczej i przy tak licznych kryteriach zawsze bowiem łatwo o jakąś przypadkową korelację.

Wprawdzie zdecydowana większość badań naukowych jest przeprowadzana w dobrej wierze, to jednak naukowcy często nieświadomie wykorzystują narzędzia tak zwanego *p-hackingu*⁵, w celu uzyskania wyniku statystycznie istotnego. Jest to wielki problem, który w połączeniu ze wzmiankowaną wcześniej preferencją czasopism naukowych do publikowania koncepcji nowych i oryginalnych - powoduje, że po pierwsze podejmowanie decyzji na podstawie wyników badań naukowych może sprowadzać się do działania na podstawie błędów I rodzaju, a więc bezpośrednio prowadzi do „problemu frajera”. Po drugie, sprawia, że przestaje funkcjonować jedna z fundamentalnych właściwości badań naukowych, mianowicie *powtarzalność*⁶. Wykazano na przykład, że badania medyczne z obszaru nowotworów są powtarzalne tylko w 10% przypadków⁷ [Begley 2013], dyscypliny psychologiczne mogą pochwalić się powtarzalnością na poziomie 33% [Aarts 2015]. Nie lepiej jest w naukach ekonomicznych, Chang i Li [2015] stwierdzili replikowalność badań makroekonomicznych na poziomie 48% należy jednak zaznaczyć, że wartość ta odnosi się jedynie do publikacji, w przypadku których udostępniano komplet danych, a właśnie ta dostępność jest osobnym problemem w naukach społecznych.

Wątek ten zakończę przykładem badania, które doskonale ilustruje problem analizy danych, błędów statystycznych, preferencji czasopism naukowych i niepewności nauki. 29 zespołom analityków (łącznie 61 badaczy) dostarczono te same dane i poproszono o udzielenie odpowiedzi na pytanie badawcze: czy czarnoskórzy piłkarze częściej dostają czerwone kartki? Wyniki były (1) bardzo różne, (2) dwadzieścia drużyn stwierdziło statystycznie istotny, pozytywny związek między kolorem skóry a ilością otrzymywanych czerwonych kartek, (3) dziewięć drużyn nie znalazło statystycznie istotnej zależności, (4) dwie drużyny oszacowały, że czarnoskórzy piłkarze otrzymują czerwone kartki aż 3 razy częściej [Silberzahn 2017]. Problem polega na tym, że pomimo faktu, iż wniosek (4) jest poglądem niszowym, to jednak ze względu na swą oryginalność, wyrazistość i prowokacyjność – to właśnie on ma największą szansę publikacji.

⁵ P-hacking to praktyka polegająca na tym, że badacz podejmuje decyzję odnośnie do sposobu analizowania danych na podstawie obserwacji samych danych, a nie na etapie projektowania badania.

⁶ Chodzi o to, czy powtarzanie danego badania naukowego daje takie same rezultaty.

⁷ Warto zaznaczyć, że analiza dotyczyła 53 badań opublikowanych przez wiodące czasopisma naukowe.

Ostatnią kwestią (d) jest zagadnienie stronniczości badań. Tradycyjne wyobrażenie o nauce karze ją postrzegać, jako proces budowania wiedzy o rzeczywistości, którego kołem zamachowym jest ciekawość badacza popychająca go do bezkompromisowej konfrontacji z naturą. Ta pełna patosu wizja bardzo różni się od realiów dzisiejszej nauki [Ravetz 2003]. Przede wszystkim współczesny naukowiec do prowadzenia badań potrzebuje zaawansowanych i kosztownych narzędzi, na które z reguły go nie stać. Pojawia się zatem konieczność finansowania badań. Oznacza to, że cele badań nie biorą się z czystej ciekawości badawczej, ale są ustalane przez rozmaite grupy interesu, którym zależy nie tyle na obiektywnej procedurze badawczej, ile na określonym wyniku. Nauka z konieczności ulega więc swego rodzaju komercjalizacji, czego wyrazem jest fakt, że coraz więcej projektów badawczych finansowanych jest przez prywatne firmy działające w danej branży, a nie przez państwo. Warto sobie uświadomić, jak bardzo niebezpieczne jest to zjawisko z punktu widzenia relacji: badacz – badana rzeczywistość. Finansujące dany projekt grupy kapitałowe zwykle są zainteresowane konkretnym wynikiem badań skorelowanym z celem ich działalności biznesowej. Rodzi to nieuniknioną presję na badacza. W takich warunkach może dochodzić zarówno do naginania faktów przez badaczy chcących wygenerować wynik zgodny z oczekiwaniami zleceniodawców, jak i do wybiórczego finansowania badań (w zależności od oczekiwanego wyniku), co daje ten sam efekt. Jako przykład można przywołać oczywiste manipulacje dokonywane przez przemysł tytoniowy oraz przemysł cukrowniczy, które w latach 60. sponsorowały badania dowodzące nieszkodliwości stosowania tytoniu oraz cukru, albo niedawną aferę związaną z glifosatem, którego szkodliwość była skutecznie ukrywana przez koncern Monsanto.

W naukach przyrodniczych wszelkie manipulacje stosunkowo łatwo można ustalić i udowodnić. Znacznie gorzej na ich tle wypadają nauki społeczne. O ile bowiem w naukach przyrodniczych finansowanie konkretnego badania nie daje gwarancji uzyskania pomyślnych (z punktu widzenia sponsorującej firmy) wyników, o tyle w naukach społecznych dofinansowanie danego projektu badawczego, powstającego w ramach konkretnej „szkoły ekonomicznej” daje często przewidywalne rezultaty [Malinowski 2018].

WIARYGODNOŚĆ BADAŃ EMPIRYCZNYCH – JESZCZE WIĘKSZE PROBLEMY

W paragrafie **Dlaczego nie można ufać wynikom badań empirycznych?** zostało wykazane, że wiarygodność badań empirycznych jest wartością z przedziału między 50 a 95 procent. Warto jednak podkreślić, że wniosek ten dotyczy badań nieobciążonych niektórymi problemami opisanymi w paragrafie **Problemy z nauką**. Ioannidis [2005] idzie jednak o krok dalej i sugeruje, aby przyjąć, że pewna część badań, która w warunkach idealnych zostałaby odrzucona – trafia do publikacji i staje się częścią nauki. Innymi słowy, zaleca on uznanie, że część badań obarczona

jest stronniczością (*bias*) z powodu: projektu badań, specyfiki weryfikacji hipotez, powszechnie panującego paradygmatu w danej dyscyplinie, uzależnienia od sponsorów itp. Potraktujmy zatem „*u*” jako fałszywe obserwacje, które zostają uznane za prawdziwe.

Tabela 4. Badania empiryczne – z uwzględnieniem problemu stronniczości

Werdykt badacza	Rzeczywistość		Suma
	Hipoteza jest prawdziwa	Hipoteza jest nieprawdziwa	
Hipoteza jest prawdziwa	$(c[1 - \beta]R + uc\beta R)/(R + 1)$	$c\alpha + uc(1 - \alpha)/(R + 1)$	$c(R + \alpha - \beta R + u - u\alpha + u\beta R)/(R + 1)$
Hipoteza jest nieprawdziwa	$(1 - u)c\beta R/(R + 1)$	$(1 - u)c(1 - \alpha)/(R + 1)$	$c(1 - u)(1 - \alpha + \beta R)/(R + 1)$
Suma	$cR/(R + 1)$	$c/(R + 1)$	c

Źródło: opracowanie własne na podstawie Ioannidis [2005]

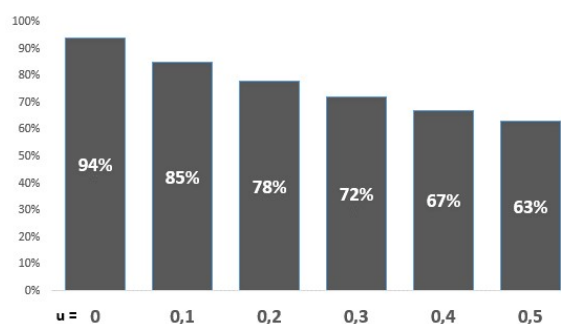
Stosując analogiczne kalkulacje do tych, z paragrafu 3 należy stwierdzić, że:

- $c\alpha + uc(1 - \alpha)/(R + 1)$ – to liczba odkryć naukowych, które są nieprawdziwe z uwagi na prawdopodobieństwo popełnienia błędu I rodzaju oraz z uwagi na stronniczość badań;
- $(c[1 - \beta]R + uc\beta R)/(R + 1)$ – to liczba dokonanych odkryć naukowych, które są prawdziwe + badania, które błędnie uznane zostałyby za nieprawdziwe, ale z uwagi na stronniczość badań, zostały potraktowane jako prawdziwe⁸.

$$WPN = \frac{(c[1 - \beta]R + uc\beta R)/(R + 1)}{c(R + \alpha - \beta R + u - u\alpha + u\beta R)/(R + 1)} = \frac{(1 - \beta)R + u\beta R}{(1 - \beta)R + \alpha + u(\beta R + 1 - \alpha)}$$

Istotne jest to, jak reaguje wskaźnik Wartości Predykcyjnej Netto (WPN) na zmiany wartości parametru „*u*”. W idealnym świecie $R=1$, $\alpha = 0,05$, $1 - \beta = 0,8$ i nie pojawia się zjawisko stronniczości badań ($u = 0$). W takiej rzeczywistości wskaźnik WPN osiąga poziom 94%.

Rysunek 3. Wskaźnik WPN (przy różnych poziomach parametru „*u*”)



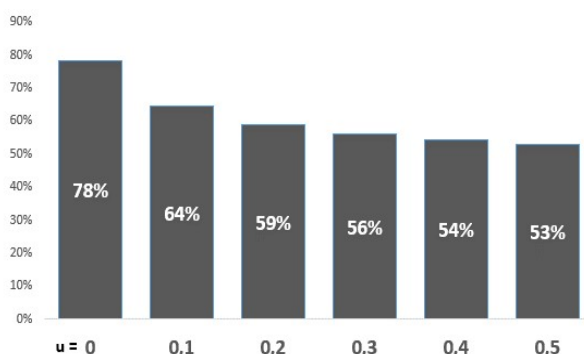
Źródło: opracowanie własne

⁸ Jest to więc pozytywny „skutek uboczny” stronniczości badań.

Jeżeli jednak parametr „u” zacznie przybierać wartości wyższe od zera (*ceteris paribus*), wówczas siła predykcyjna istotnie spada. Zapewne parametr „u” przybiera różne wartości w zależności od dyscypliny badań, ale nawet nadanie mu wartości 0,1 skutkuje spadkiem siły predykcyjnej o 10 punktów procentowych.

Jeżeli zaś w działaniu posłużymy się typową wartością mocy statystycznej charakteryzującej publikacje z zakresu ekonomii ($1-\beta = 18\%$), to nawet przy 10 – procentowym poziomie stronniczości, siła predykcyjna spada do 78% i do 56%, gdy $u=0,3$ (rys. 4).

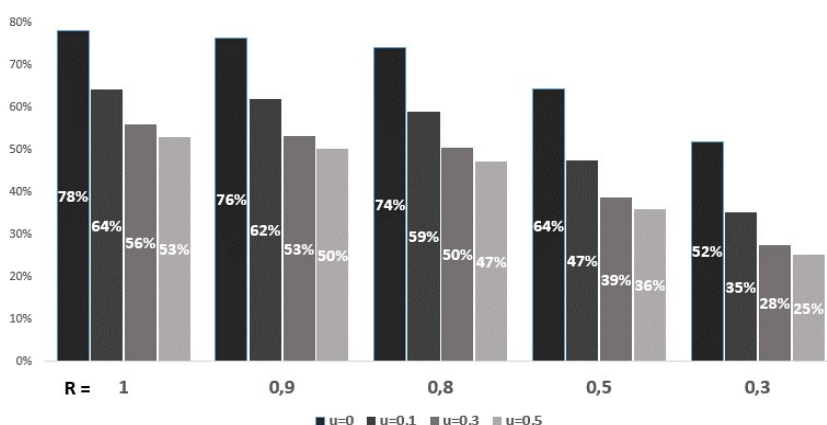
Rysunek 4. Wskaźnik WPN (przy różnych poziomach parametru „u”, gdy $1-\beta = 18\%$)



Źródło: opracowanie własne

Jednak największy niepokój niesie z sobą połączenie 18 – procentowej mocy statystycznej, różnych wartości „u” oraz „R” (rysunek 5), gdyż siła predykcyjna spada istotnie i wynosi 78% w scenariuszu najbardziej optymistycznym oraz 25% w scenariuszu najbardziej pesymistycznym.

Rysunek 5. Wskaźnik WPN (przy różnych poziomach współczynnika „R” oraz parametru „u”, gdy $1-\beta = 18\%$)



Źródło: opracowanie własne

Należy też zaznaczyć, że wartości $u=0,5$ oraz $R=0,3$, które w niniejszym artykule prezentowane są jako skrajne, wcale skrajnymi nie są. W zależności od przykład upolitycznienia badań – poziom parametru „u” może być znacznie wyższy od 0,5 a z kolei współczynnik „R”, może przybierać wartości bardzo niskie ($R < 0,1$). Dotyczy to zazwyczaj badań, w których kontekst odkrycia odgrywa szczególnie istotną rolę.

Zaprezentowane rozważania prowadzą do wniosku, że dyscypliny naukowe, w których dowodzenie odbywa się za pomocą narzędzi statystycznych, obciążone są znacznym poziomem niepewności naukowej. Oczywiście stopniowo, wraz z upływem czasu – błędne hipotezy są eliminowane, ale dla praktycznie zorientowanego podmiotu, podejmowanie decyzji na podstawie badań naukowych może się okazać bardzo ryzykownym przedsięwzięciem.

NIEPEWNOŚĆ NAUKI W KONTEKŚCIE „ZALEŻNOŚCI OD ŚCIEŻKI” ORAZ „RYZYKA KATASTROFY”

Nie byłoby „problemu frajera”, gdyby nie wyjątkowy status, jakim cieszy się nauka we współczesnym świecie. Niestety niepewność nauki często nie jest w ogóle brana pod uwagę przy podejmowaniu istotnych decyzji, zwłaszcza o charakterze ekonomicznym / biznesowym. Innymi słowy, decydenci zakładają zwykle, że opinie ekspertów, które wszakże opierają się na aktualnych badaniach naukowych – są rzeczywistością obiektywną i pewną, i podchodzą do nich bezkrytycznie, albo też z niewystarczającym poziomem krytycyzmu. Mając zaś na uwadze wcześniejsze ustalenia dotyczące siły predykcyjnej badań empirycznych, wydaje się, że zachowanie szczególnej ostrożności w odniesieniu do wiarygodności badań jest postępowaniem ze wszech miar bardziej racjonalnym.

Można jednak stwierdzić, że lepiej jest działać w oparciu o „niepewną naukę” gdyż alternatywą jest odwoływanie się do przesłanek o metodologicznie niejasnym charakterze – jak intuicja lub heurystyki. Takie podejście należałoby jednak skonfrontować ze znanym pytaniem retorycznym postawionym przez Nassima Taleba „czy lepiej jest iść przez las z błędną mapą, czy bez mapy?”

Ponadto, warto zaznaczyć, że „problem frajera” nabiera szczególnego sensu w kontekście:

- zależności od ścieżki;
- ryzyka katastrofy.

Pierwsze zjawisko podkreśla rolę sekwencji zdarzeń. Osoba grająca w ruletkę może najpierw wygrać milion dolarów, a następnie zbankrutować. Nie może jednak najpierw zbankrutować, a następnie wygrać miliona dolarów. Oznacza to, że ostateczny wynik jakiegoś procesu (może nim być zarówno rozwój gospodarczy kraju, jak również i funkcjonowanie przedsiębiorstwa w czasie) nie jest prostą sumą następujących po sobie wyników cząstkowych, gdyż ich kolejność może (choć nie musi) mieć decydujące znaczenie [Taleb 2016].

Drugi termin odnosi się do czegoś, co można także określić mianem „nieakceptowalnej katastrofy” a więc jakiegoś nieodwracalnego zniszczenia o charakterze systemowym. Ryzyko katastrofy jest nieakceptowalne, ponieważ jego materializacja wiąże się z nieodwracalną klęską, bankructwem, porażką itp. Warto zwrócić uwagę na to, że podejmowanie decyzji związanych z tym rodzajem ryzyka znacząco zmienia kryteria racjonalnego postępowania. Okazuje się bowiem, że precyzyjna, skwantyfikowana wiedza dotycząca prawdopodobieństwa materializacji tego rodzaju zagrożenia – wcale nie zmienia podstaw podejmowania decyzji. Można wręcz stwierdzić, że w przypadku ryzyka katastrofy, podmioty podejmujące decyzje często postępują tak, jakby poziom prawdopodobieństwa materializacji katastrofalnych scenariuszy wcale nie miał znaczenia [Jablonowski 2006].

W kontekście tych dwóch zjawisk, problem naukowej niepewności nabiera szczególnego charakteru z tego względu, że istnieją pewne rodzaje błędów, których organizacja nie może popełnić, gdyż wiążą się one z definitywnym wypadnięciem z gry. Zazwyczaj rozumieją to przedsiębiorcy, których decyzje są z całą bezwzględnością weryfikowane przez rynek i są zazwyczaj pierwszymi, którzy ponoszą konsekwencje błędnych decyzji.

Pojawia się w tym miejscu ciekawa refleksja, której istota odbiega od problematyki tego artykułu, ale która jest warta dalszego pogłębienia. Warto mianowicie zwrócić uwagę na fakt, że biorąc pod uwagę triadę: państwo – przedsiębiorstwa – jednostki, to na poziomie państwa ryzyko katastrofy jest ryzykiem nieakceptowalnym. Innymi słowy – państwo nie może popełnić egzystencjonalnego błędu, ponieważ jego fundamentalnym celem jest przetrwanie. Przetrwanie natomiast bezpośrednio związane jest z unikaniem potencjalnie katastrofalnych wyborów. W dłuższej perspektywie bowiem stała, nawet minimalna ekspozycja na katastrofalne ryzyko w końcu się zmaterializuje.

Zupełnie zaś inaczej jest z przedsiębiorstwami. Przedsiębiorstwa dotyka swoista „kreatywna destrukcja”, co oznacza, że są one zmuszane przez szeroko rozumiane społeczeństwo do podejmowania bardzo ryzykownych decyzji, gdyż to jest właśnie zgodne z interesem społecznym. W konsekwencji jednak – przedsiębiorstwa mają stosunkowo krótki czas istnienia, gdyż stale są zastępowane przez inne, których ryzykowne decyzje przyniosły korzyść, a potencjalnie zabójcze ryzyka się (jeszcze) nie zmaterializowały. W kontekście „problemu frajera” można więc stwierdzić, że wprawdzie przedsiębiorca nie chce być frajerem, to jednak często musi nim być, gdyż w przeciwnym razie inne przedsiębiorstwo podejmie bardzo nieostrożną decyzję, która w wyniku szczęśliwego zbiegu okoliczności nie doprowadzi do katastrofy, ale wręcz przeciwnie – do sukcesu i do osiągnięcia czasowej przewagi.

Jak zatem podejmować decyzje w warunkach naukowej niepewności? Zasygnalizowany w tekście problem związany jest z nieredukowalną niepewnością, w kontekście której trudno udzielać uniwersalnych, metodologicznie uzasadnionych rekomendacji. Wydaje się, że najważniejsza w tym kontekście jest sama świadomość istnienia problemu „naukowej niepewności”. Podmiot podejmujący decyzje musi w

swych algorytmach decyzyjnych uwzględnić fakt, że opierania działania na bieżących badaniach naukowych niesie z sobą pewne ryzyko, które pomimo przedstawionych w niniejszym tekście szacunków – jest niekwantyfikowalne.

Być może swoistym kompasem w świecie niepewności jest roztropne korzystanie z ukształtowanych przez wieki „wskazań mądrościowych” będących w swej istocie zakumulowaną, wielopokoleniową intuicją [Gigerenzer 2014]. W ramach przykładu można przytoczyć chińską maksymę, wokół której zresztą budowana była chińska strategia rozwoju gospodarczego zwieńczona tak ogromnym sukcesem [Góralczyk 2018]. Sentencja stwierdza, że należy „kroczyć przez rzekę, czując kamienie pod stopami”. Interpretacja tych słów nakazuje zwrócić uwagę na dwa elementy. (1) świadomość wyraźnie wytyczonego celu i (2) ostrożność.

Wszakże przez rzekę przejść trzeba, żeby dostać się na drugi brzeg, ale należy to robić stopniowo, w razie konieczności trzeba się przesunąć w bok albo cofnąć i nieustannie należy z pełną podejrzliwością badać dno, po którym stąpamy. Nie możemy pozwolić sobie na błąd, gdyż oznacza on katastrofę.

Idea podejmowania decyzji na podstawie nieskomplikowanych, przystępnych poznawczo reguł znajduje teoretyczne wsparcie w nurcie tak zwanej racjonalności adaptacyjnej. Badacze reprezentujący to podejście dowodzą, że w warunkach nieredukowalnej niepewności lepsze decyzje są efektem stosowania suboptymalnych heurystyk, nie zaś optymalnych modeli czy algorytmów [Potocki, Opolski 2015].

ZAKOŃCZENIE

W ramach zakończenia należy zwrócić uwagę na metodologiczną specyfikę nauk o zarządzaniu jak i ekonomicznych, która ma istotny wpływ na status „problem frajera”. Otóż Herbert Simon podzielił dyscypliny naukowe na teoretyczne oraz inżynierskie. Te pierwsze dążą do poznania prawdy zaś te drugie nastawione są na przeobrażanie rzeczywistości. Takie dyscypliny jak medycyna, zarządzanie czy ekonomia Simon zaklasyfikował właśnie do tej drugiej kategorii. Koźmiński i Latusek-Jurczak [2011] posługując się tą klasyfikacją rozważają, jaki jest stosunek twierdzeń powstających na gruncie teorii zarządzania (a więc także i ekonomii) do rzeczywistości. Dochodzą do wniosku, że z całą pewnością rozstrzygnięcia teoretyczne nie nadają się do stosowania w konkretnych sytuacjach. Rolą teorii jest raczej dostarczanie opisów i wyjaśnień składających się na bardziej dogłębne zrozumienie rzeczywistości a ponadto, koncepcje teoretyczne mają na celu przybliżenie praktykom możliwych rozwiązań danych sytuacji problemowych. Ostatecznie zatem, Koźmiński dochodzi do wniosku, że podstawową funkcją badań w ramach nauk ekonomicznych, w stosunku do praktyków jest inspiracja.

Warto zwrócić uwagę, że przyjęcie optyki Koźmińskiego i Latusek-Jurczaka prowadzi do istotnej neutralizacji „problemu frajera”. Jest tak dlatego, ponieważ praktyk „jedynie” inspiruje się koncepcjami teoretycznymi powstałymi na gruncie nauk ekonomicznych. Ustalenia tych dziedzin wcale nie determinują jego poczyną

a więc kwestia ostatecznej efektywności działań zostaje niejako wyłączona z naukowej analizy i pozostawiona takim czynnikom jak intuicja czy doświadczenie. Innymi słowy – naukowa siła przekonywania danej koncepcji / teorii wcale nie „zmusza” praktyka do kierowania się nią w procesie podejmowania decyzji.

Trudno się jednak oprzeć wrażeniu, że współcześnie status badań naukowych jest czymś więcej niż tylko „inspiracją”. Wszakże rozpowszechnione w medycynie podejście *evidence – based medicine* nie pozostawia lekarskiej intuicji zbyt dużo miejsca, a niewątpliwy sukces nauk medycznych powoduje, że także inne dyscypliny wdrażają podobny sposób myślenia po to, ażeby stać się bardziej „naukowe”. W ten zaś sposób badanie naukowe, lub jego brak nie tyle inspiruje, co przesądza o możliwości zastosowania danego rozwiązania w praktyce.

Wydaje się, że najlepszą ilustracją tego rozumowania są intelektualne „inspiracje” polityków gospodarczych z ostatnich dekad. Otóż w latach 80tych rozpowszechniło się traktowanie prywatyzacji, liberalizacji i deregulacji jako lekarstwa na wszystkie choroby gospodarcze. W latach 90tych popularne było myślenie o „końcu historii” czy o „teorii skapywania”, dekadę później karierę zrobiła koncepcja „turbokapitalizmu”, a współcześnie głośnym fenomenem jest tak zwana „pułapka średniego dochodu”. Podkreślenia wymaga fakt, że nie są to koncepcje o charakterze beletrystycznym, ale bardzo mocno osadzone w badaniach empirycznych wytwory nauk ekonomicznych. Historia pokazuje, że niemal żaden polityk gospodarczy nie mógł przejść obojętnie obok tego typu „skrzydlatych pojęć” a największą przychylność mediów oraz mainstreamu ekonomicznego gwarantowało uczynienie z tych idei myśli przewodniej, wokół której budowano strategię rozwoju kraju.

BIBLIOGRAFIA

- Aarts A. A., Anderson J., Anderson C. (2015) Estimating the Reproducibility of Psychological Science. *Science*, 349(6251), aac4716-1 – aac4716-8.
- Ahteensuu M. (2013) The Precautionary Principle and the Justifiability of Three Imperatives. *Homo Oeconomicus*, 30(1), 17-36.
- Begley C.G. (2013) Reproducibility: Six Red Flags for Suspect Work. *Nature*, 497(7450), 433-434.
- Bohannon J., Koch D., Himm P., Driehaus A. (2015) Chocolate with High Cocoa Content as a Weight – Loss Accelerator. *International Archives of Medicine*, 8, 55.
- Chang A., Li P. (2015) Is Economic Research Replicable? Sixty Published Papers from Thirteen Journals Say “Usually Not”. *Finance and Economics Discussion Series*, 2015-083. Washington: Board of Governors of the Federal Reserve System, <http://dx.doi.org/10.17016/FEDS.2015.083>.
- Gigerenzer G. (2014) *Risk Savvy*. Penguin Random House.
- Góralczyk B. (2018) *Wielki Renesans*. Wydawnictwo Akademickie Dialog, Warszawa.
- Ioannidis J. P. A. (2005) Why Most Published Research Findings Are False. *PLoS Med*, 2(8): e124, 696-701.

- Ioannidis J. P. A., Stanley T. D., Doucouliagos H. (2017) The Power of Bias in Economics Research. *The Economic Journal*, 127(10), F236-F265.
- Jablonowski M. (2006) *Precautionary Risk Management*. Palgrave Macmillan, New York.
- Koźmiński A. K. (2008) *Zarządzanie w warunkach niepewności*. PWN, Warszawa.
- Koźmiński A. K., Latusek - Jurczak D. (2011) *Rozwój Teorii Organizacji*. Wolters Kluwer Polska, Warszawa.
- Malinowski G. M. (2017) Renesans strategii – czyli o niemożliwości uprawiania Evidence based Policy. *Organizacja i Zarządzanie*, 113, 265-284.
- Malinowski G. M. (2018) Zasada Ostrożności, czyli heurystyka strachu oraz heurystyka odwagi w kontekście polityki gospodarczej. *Prakseologia*, 160, 291-332.
- Lemons J., Shrader – Frechette K., Cranor C. (1997) The Precautionary Principle: Scientific Uncertainty and Type I and Type II Errors. *Foundations of Science*, 2, 207-236.
- Lo Ch. (2009) Risks, Scientific Uncertainty and The Approach of Applying Precautionary Principle. *Medicine and Law*, 28, 283-300.
- Potocki T., Opolski K. (2015) Decyzje w obliczu „niepewnych ryzyk” – rola heurystyk i nurtu racjonalności adaptacyjnej. *Finanse*, 1(8), 43-70.
- Ravetz J. (2003) The Post – Normal Science of Precaution. *Futures*, 36, 347-357.
- Silberzahn R., Uhlmann E. L., Martin D. (2017) Many Analyst, One Dataset: Making Transparent how Variations in Analytical Choices Affect Results. *Advances in Methods and Practices in Psychological Science*, 1(3), 337-356.
- Taleb N. (2016) *The Logic of Risk Taking*.
<http://www.fooledbyrandomness.com/rationality.pdf>, z dn. 30.01. 2019.
- The Economist (2013) Trouble at the Lab, October 19th, 26-30.
- Weiss Ch. (2003) Expressing Scientific Uncertainty. *Law, Probability and Risk*, 2, 25-46.

UNCERTAINTY OF SCIENCE AND DECISION-MAKING – PROBLEMS WITH EVIDENCE-BASED POLICY

Abstract: The purpose of this article is primarily to introduce the topic of scientific uncertainty to the wider context of economics and management. Scientific uncertainty is one of the manifestations of irreducible uncertainty and reflection on it should enable better decision making. An entity that bases its operation on current scientific research, which depreciates over time and ultimately leads to erroneous decisions, is referred to as the “loser”. The text indicates estimation of potential scale of this problem supplemented by an outline of sociological difficulties identified in the analysis of the process of building scientific statements. The article ends with a sketch of the answer to the question “how to act in the context of scientific uncertainty?”.

Keywords: science uncertainty, irreducible uncertainty, loser problem, statistical errors, precautionary principle

JEL classification: D01, D81, H12