

Cena 13,00 zł
(VAT 8%)

Indeks 381306
e-ISSN 2543-8476
PL ISSN 0043-518X

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

MAJ / MAY
ROK / VOLUME 66

2021 | 5

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

POLSKIE TOWARZYSTWO STATYSTYCZNE
POLISH STATISTICAL ASSOCIATION



WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

MAJ / MAY
ROK / VOLUME 66

2021 | 5 (720)

RADA NAUKOWA / SCIENTIFIC COUNCIL

dr Dominik Rozkrut (przewodniczący/chairman) – Uniwersytet Szczeciński; Prof. Anthony Arundel – Maastricht University; Eric Bartelsman, PhD, Assoc. Prof. – Vrije Universiteit Amsterdam; prof. dr hab. Czesław Domański – Uniwersytet Łódzki; prof. dr hab. Elżbieta Gołata – Uniwersytet Ekonomiczny w Poznaniu; Semen Matkovskyy, PhD, Assoc. Prof. – Ivan Franko National University of Lviv; prof. dr hab. Włodzimierz Okrasa – Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie; prof. dr hab. Józef Oleński – Polskie Towarzystwo Statystyczne; prof. dr hab. Tomasz Panek – Szkoła Główna Handlowa w Warszawie; Juan Manuel Rodriguez Poo, PhD, Assoc. Prof. – University of Cantabria; Iveta Stankovičová, BEng, PhD, Assoc. Prof. – Comenius University in Bratislava; prof. dr hab. Marek Walesiak – Uniwersytet Ekonomiczny we Wrocławiu; prof. dr hab. Józef Zegar – Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy

sekretarz/secretary: Paulina Kucharska-Singh

KOLEGIUM REDAKCYJNE / EDITORIAL BOARD

Tudorel Andrei, PhD, Assoc. Prof. – Bucharest Academy of Economic Studies; mgr Renata Bielak – Główny Urząd Statystyczny; dr Marek Cierpiął-Wolan – Uniwersytet Rzeszowski; dr hab. Grażyna Dehnel, prof. UEP – Uniwersytet Ekonomiczny w Poznaniu; dr Jacek Kowalewski – Uniwersytet Ekonomiczny w Poznaniu; dr Jan Kubacki – Urząd Statystyczny w Łodzi; dr Grażyna Marciniak; dr hab. Andrzej Młodak, prof. AK – Akademia Kaliska im. Prezydenta Stanisława Wojciechowskiego; dr hab. Mateusz Pipień, prof. UEK – Uniwersytet Ekonomiczny w Krakowie; Marek Rojček, BEng, PhD – University of Economics, Prague; Anna Shostya, PhD, Assoc. Prof. – Pace University in New York; dr hab. Małgorzata Tarczyńska-Łuniewska, prof. US – Uniwersytet Szczeciński; dr Wioletta Wrzaszcz – Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy; dr inż. Agnieszka Zgierska – Główny Urząd Statystyczny

ZESPÓŁ REDAKCYJNY / EDITORIAL STAFF

redaktor naczelny / editor-in-chief: Marek Cierpiął-Wolan

zastępca redaktora naczelnego / deputy editor-in-chief: Andrzej Młodak

redaktorzy tematyczni / thematic editors: Jan Kubacki, Małgorzata Tarczyńska-Łuniewska, Agnieszka Zgierska

redaktor merytoryczny / substantive editor: Wioletta Wrzaszcz

sekretarz/secretary: Małgorzata Zygmunt

ADRES REDAKCJI / EDITORIAL OFFICE ADDRESS

Główny Urząd Statystyczny / Statistics Poland, al. Niepodległości 208, 00-925 Warszawa
tel./phone +48 22 608 32 25, e-mail: redakcja.ws@stat.gov.pl

Redakcja językowa: Wydział Czasopism Naukowych, Główny Urząd Statystyczny
Language editing: Scientific Journals Division, Statistics Poland

Redakcja techniczna, skład i łamanie, wykresy, korekta: Zakład Wydawnictw Statystycznych – zespół pod kierunkiem Wojciecha Szuchty

Technical editing, typesetting, figures, proof-reading: Statistical Publishing Establishment – team supervised by Wojciech Szuchta



Zakład Wydawnictw
Statystycznych

Druk i oprawa / Printed and bound:

Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment
al. Niepodległości 208, 00-925 Warszawa, zws.stat.gov.pl

Wersja elektroniczna, stanowiąca wersję pierwotną czasopisma, jest dostępna na ws.stat.gov.pl
The original version of the journal is the electronic issue, available at ws.stat.gov.pl

© Copyright by Główny Urząd Statystyczny

Indeks 381306

Informacje w sprawie sprzedaży czasopisma / Sales of the journal:

Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment

tel./phone +48 22 608 32 10, +48 22 608 38 10

Prenumerata jest prowadzona przez / Subscription is available at RUCH S.A.

Zamówienia na prenumeratę można składać na stronie / Subscriptions can be ordered at
www.prenumerata.ruch.com.pl

SPIS TREŚCI

CONTENTS

Od redakcji	4
From the editorial team	
 Studia metodologiczne	
Methodological studies	
Mieczysław Kowerski, Jarosław Bielał	
Kilka uwag na temat pomiaru zależności pomiędzy zmiennymi o panelowej strukturze danych	7
A selection of remarks on the measurement of correlations between variables of a panel data structure	
 Statystyka w praktyce	
Statistics in practice	
Marcin Salamaga	
Assessment of the risk of foreign divestment in Poland during the COVID-19 pandemic	26
Ocena ryzyka dezinvestycji zagranicznych w Polsce w warunkach pandemii COVID-19	
Dorota Pekasiewicz	
Zastosowanie rozkładu Pareta do aproksymacji rozkładów dochodów zamożnych gospodarstw domowych	43
The application of the Pareto distribution to approximate income distributions of wealthy households in Poland	
 Dyskusje. Recenzje. Informacje	
Discussions. Reviews. Information	
Justyna Gustyn	
Wydawnictwa GUS. Kwiecień 2021	60
Publications of Statistics Poland. April 2021	
 Dla autorów	62
For the authors	
 Zakres tematyczny działów	73
Thematic scope of sections	

OD REDAKCJI

W majowym wydaniu „Wiadomości Statystycznych. The Polish Statistician” proponujemy Państwu lekturę artykułów z zakresu studiów metodologicznych oraz zastosowań statystyki w praktyce.

Dr hab. Mieczysław Kowerski i mgr Jarosław Bielak w artykule *Kilka uwag na temat pomiaru zależności pomiędzy zmiennymi o panelowej strukturze danych* przedstawiają argumenty świadczące o niezasadności stosowania współczynnika korelacji liniowej Pearsona w analizie zależności danych panelowych. Wskazują, że współczynnik korelacji liniowej Pearsona nie uwidacznia informacji o zależnościach badanych zjawisk jednocześnie w czasie i przestrzeni, których dostarczanie jest podstawową zaletą danych panelowych. Jako bardziej adekwatną (choć również niedoskonałą) miarę autorzy proponują współczynnik korelacji pomiędzy wartościami empirycznymi i teoretycznymi modelu panelowego z jedną zmienną objaśniającą. Artykuł stanowi przyczynek do dyskusji nad badanym zagadnieniem. Rozważania teoretyczne zostały uzupełnione przykładem empirycznym obrazującym sytuację spółek budowlanych notowanych na Giełdzie Papierów Wartościowych.

W artykule *Assessment of the risk of foreign divestment in Poland during the COVID-19 pandemic* dr hab. Marcin Salamaga, prof. UEK, bada znaczenie bezpośrednich inwestycji zagranicznych (BIZ) dla rozwoju gospodarczego Polski oraz oddziaływanie kryzysów na napływ tych inwestycji do Polski. W tym celu wykorzystuje wyniki sondażu przeprowadzonego w 2020 r. wśród prawie 500 przedsiębiorstw inwestujących w Polsce. Przy zastosowaniu modeli logitowych identyfikuje czynniki, które w największym stopniu wpływają na dez-inwestycje w gospodarce w warunkach pandemii COVID-19. Z badania wynika, że przy założeniu znacznych dez-inwestycji ryzykogenność lokowania BIZ w usługach jest wyższa niż ich lokowania w przemyśle przetwórczym, a w przypadku niewielkich dez-inwestycji największe ryzyko dotyczy branży IT.

Celem badania przedstawionego w artykule dr hab. Doroty Pekasiewicz, prof. UŁ, pt. *Zastosowanie rozkładu Pareta do aproksymacji rozkładów dochodów zamożnych gospodarstw domowych* jest aproksymacja rozkładów ekwiwalentnych dochodów zamożnych gospodarstw domowych w poszczególnych grupach społeczno-ekonomicznych. Do aproksymacji autorka stosuje rozkład Pareta z parametrami oszacowanymi metodą największej wiarygodności. Wykorzystuje dane za lata 2014–2017 pochodzące z badań budżetów gospodarstw domowych. Rozkłady teoretyczne dobrze dopasowane do danych empirycznych posłużyły autorce do oszacowania miar charakteryzujących rozkłady ekwiwalentnych dochodów gospodarstw domowych, w tym miar nierównomierności, i pozwoliły na porównywanie grup pod względem ich zamożności.

Wydanie kończy się omówieniem nowości wydawniczych GUS przygotowanym przez Justynę Gustyn.

Zapraszamy do lektury.

FROM THE EDITORIAL TEAM

The May issue of *Wiomości Statystyczne. The Polish Statistician* features articles relating to methodological studies and practical applications of statistics.

Mieczysław Kowerski, PhD, DSc, and Jarosław Bielak, MSc, in their article entitled *A selection of remarks on the measurement of correlations between variables of a panel data structure* present arguments proving that the application of the Pearson linear correlation coefficient is unsuitable for the analysis of panel data dependencies. The authors indicate that the Pearson linear correlation coefficient fails to provide information on the dependencies between examined phenomena simultaneously in time and space, while the basic advantage of panel data is the ability to yield such information. A correlation coefficient between the empirical and fitted values of a panel model with one explanatory variable is proposed as a more adequate measure (although still not a perfect one). The article is a contribution to the discussion on the examined issue. The theoretical considerations have been complemented by an empirical example illustrating the situation of construction companies listed on the Warsaw Stock Exchange.

Marcin Salamaga, PhD, DSc, Professor at the Cracow University of Economics, in his article *Assessment of the risk of foreign divestment in Poland during the COVID-19 pandemic* examines the importance of foreign direct investments for the economic development of Poland, and the impact crises have on the inflow of such investments to Poland. The research is based on a survey conducted in 2020 among nearly 500 firms investing in Poland. The author applies logit models in order to identify which factors cause divestment in the economy struck by the COVID-19 pandemic to the largest extent. The results of the research demonstrate that if divestment occurs on a large scale, then the services sector bears a greater risk of it than the processing industry, while the IT sector is likely to face the largest risk in case smaller-scale divestment occurs.

The aim of the study presented in the article entitled *The application of the Pareto distribution to approximate income distributions of wealthy households in Poland* by Dorota Pekasiewicz, PhD, DSc, Professor at the University of Lodz, is the approximation of the equivalent income distribution of wealthy households in particular socio-economic groups. The approximation is based on the Pareto distribution with parameters estimated by means of the maximum likelihood method. The author uses data for the years 2014–2017 collected in the Household Budget Survey. Theoretical distributions appropriately fitted to empirical data enabled the author to estimate the measures characterising the distribution of the equivalent household income, including measures of inequality, and allowed the comparison of the groups under study in terms of the level of their wealth.

The issue concludes with a discussion on Statistics Poland's new publications, prepared by Justyna Gustyń.

We wish you pleasant reading.

Kilka uwag na temat pomiaru zależności pomiędzy zmiennymi o panelowej strukturze danych

Mieczysław Kowerski^a, Jarosław Bielak^b

Streszczenie. W licznych artykułach modelowanie na podstawie danych panelowych rozpoczyna się od przedstawienia macierzy współczynników korelacji liniowej Pearsona pomiędzy zmiennymi przyjętymi do badania. Celem niniejszego artykułu jest pokazanie nieprzydatności takiego podejścia w analizie zależności w przypadku danych panelowych oraz próba zaproponowania bardziej adekwatnej miary – współczynnika korelacji pomiędzy wartościami empirycznymi i teoretycznymi zmiennej objaśnianej oszacowanego modelu panelowego (z efektami stałymi lub losowymi) względem zmiennej, której zależność w stosunku do zmiennej objaśnianej jest obliczana.

Współczynnik korelacji liniowej Pearsona nie uwidacznia podstawowej zalety danych panelowych, jaką jest dostarczanie informacji o zależnościach badanych zjawisk jednocześnie w czasie i przestrzeni. Dla obliczenia tego współczynnika nie jest bowiem istotne, że pewna obserwacja dotyczy obiektu i w okresie t , a inna – obiektu j w okresie $t + 1$. Można go natomiast wykorzystać w analizach danych panelowych do obliczeń cząstkowych.

Prowadzone rozważania zilustrowano obliczeniami zależności pomiędzy strukturą kapitału oraz rentowością i wielkością 17 spółek budowlanych notowanych na Giełdzie Papierów Wartościowych w Warszawie w latach 2009–2018 (170 obserwacji), tworzących panel zbilansowany. Obliczenia te pozwoliły na sformułowanie zalet i wad zaproponowanego rozwiązania.

Słowa kluczowe: dane o strukturze panelowej, współczynnik korelacji liniowej Pearsona, model panelowy

JEL: C18, C23

A selection of remarks on the measurement of correlations between variables of a panel data structure

Abstract. Many articles featuring panel data modelling tend to begin their considerations with an introduction of the Pearson linear correlation coefficients matrix between the analysed variables. The aim of the article is to prove such an approach unsuitable in the analysis of panel data dependencies. Instead, an attempt has been made to propose a more appropriate measure – a correlation coefficient between the empirical and fitted values of the dependent variable of the estimated panel model (with fixed or random effects) in relation to the variable whose dependency towards the dependent variable is being studied.

^a Uczelnia Państwowa im. Szymona Szymonowica w Zamościu, Instytut Społeczno-Ekonomiczny, Polska / State Higher School of Vocational Education in Zamość, Department of Social and Economic Sciences, Poland. ORCID: <https://orcid.org/0000-0002-2147-2037>. Autor korespondencyjny / Corresponding author, e-mail: mieczyslaw.kowerski@upz.edu.pl.

^b Uczelnia Państwowa im. Szymona Szymonowica w Zamościu, Instytut Społeczno-Ekonomiczny, Polska / State Higher School of Vocational Education in Zamość, Department of Social and Economic Sciences, Poland. ORCID: <https://orcid.org/0000-0001-8537-8624>. E-mail: jaroslaw.bielak@upz.edu.pl.

Pearson's linear correlation coefficient does not reflect the basic advantage of panel data, which is the ability to provide information about the dependencies of the studied phenomena simultaneously in time and space. The fact that one observation relates to object i during period t and another to object j during period $t + 1$ is irrelevant for the calculation of the coefficient. Pearson's coefficient, however, can be used when conducting sub-calculations in panel data analysis.

The presented considerations have been illustrated by the calculations of the relationships between the structure of capital and the profitability and size of 17 construction companies listed on the Warsaw Stock Exchange in the years 2009–2018 (170 observations) which created a balanced panel. A specification of the advantages and disadvantages of the proposed solution was formulated on the basis of the calculations.

Keywords: panel data structure, Pearson linear correlation coefficient, panel model

1. Wprowadzenie

Terminem *dane panelowe* określa się zbiór informacji o jednoznacznie identyfikowalnych obiektach (jednostkach) obserwowanych w czasie. Metody analizy i modelowania zjawisk przy użyciu takich danych cieszą się w ostatnich latach rosnącym zainteresowaniem ekonomistów. Wynika to z wielu zalet danych panelowych i przewagi modelowania zjawisk ekonomicznych na podstawie tego rodzaju danych nad oddzielnym modelowaniem zjawisk opisanych przez szeregi czasowe czy przekrojowe; wykorzystanie danych panelowych umożliwia pełniejszy opis analizowanego zjawiska, większą liczbę dostępnych informacji, a także eliminację lub redukcję obciążenia estymatorów (Witkowski, 2012, s. 268–269).

Jak wskazuje literatura przedmiotu (zob. przykłady omówione poniżej), pierwszym krokiem w modelowaniu na podstawie danych panelowych dość często jest badanie korelacji liniowej między zmiennymi za pomocą współczynnika Pearsona.

Na podstawie zgromadzonych danych panelowych opisujących sytuację finansową badanych spółek w latach 1998–2016 Herman (2019, s. 45–46) obliczył wartości współczynnika korelacji liniowej Pearsona, a następnie wykorzystał je do przeprowadzenia wnioskowania o sile i kierunku zależności pomiędzy potencjalnymi zmiennymi objaśniającymi a zmiennymi objaśnianymi. Stwierdził przy tym, że właśnie taka analiza korelacji powinna poprzedzać podejście wielowymiarowe.

Borsuk i Kostrzewa (2020) poprzedzili prezentację wyników estymacji modeli ryzyk systemowych oszacowanych na podstawie kwartalnych danych panelowych obejmujących działalność banków komercyjnych na polskim rynku w latach 2007–2017 przedstawieniem macierzy współczynników korelacji liniowej Pearsona analizowanych zmiennych, ale nie podjęli się ich interpretacji.

Driver i in. (2020) zaprezentowali macierz współczynników korelacji liniowej Pearsona zmiennych przyjętych do badania przed wykonaniem estymacji panelowych modeli Lintnera opisujących politykę dywidendową spółek notowanych na

giełdzie londyńskiej w latach 1997–2012. Za pomocą analizy poziomu istotności współczynników korelacji autorzy próbowali odpowiedzieć na pytanie, które zmienne powinny znaleźć się w modelach panelowych jako zmienne objaśniające. Dociekali także, czy może wystąpić zjawisko przybliżonej współliniowości pomiędzy potencjalnymi zmiennymi objaśniającymi.

Karkowska (2019) obliczyła macierz współczynników korelacji Pearsona przyjętych zmiennych i dokonała oceny ich istotności oraz kierunku zależności pomiędzy potencjalnymi zmiennymi objaśniającymi i zmiennymi objaśnianymi, aby później wykonać estymację – na podstawie danych panelowych z 4678 banków w 31 krajach europejskich w latach 1996–2011 – modeli źródeł ryzyka systematycznego sektora bankowego.

Pluskota (2020) obliczyła macierz korelacji przyjętych zmiennych i oceniła istotność współczynników korelacji, a następnie, na podstawie danych panelowych pochodzących z rozwiniętych krajów europejskich w latach 1996–2017, przeprowadziła estymację modeli opisujących wpływ korupcji na tempo wzrostu gospodarczego i innowacje. Nieistotnie skorelowana ze zmiennymi objaśnianymi zmienna *corruption* okazała się istotna w oszacowanych modelach panelowych (s. 82–84), co może być sygnałem, że współczynnik korelacji liniowej Pearsona nie mierzy zależności w przypadku danych panelowych.

Z kolei Urbanek (2017), zanim przystąpił do estymacji modeli panelowych zależności pomiędzy wynikami finansowymi i siłą marki spółek notowanych na Giełdzie Papierów Wartościowych (GPW) w Warszawie w latach 2008–2014, obliczył macierz korelacji przyjętych zmiennych i już na podstawie oceny istotności współczynników korelacji liniowej Pearsona doszedł do wniosku, że ich wartości w pełni potwierdzają postawione hipotezy¹.

Wymienione wyżej artykuły zawierają prawidłowo zastosowane metody estymacji modeli panelowych uwzględniające zróżnicowanie obserwacji ze względu na przestrzeń i czas. Zdaniem autorów niniejszej pracy obliczanie i prezentacja wartości współczynnika korelacji liniowej Pearsona pomiędzy zmiennymi o panelowej strukturze danych nie są więc potrzebne, a niekiedy wręcz wprowadzają badaczy w błąd. Nie oznacza to, że nie należy obliczać zależności pomiędzy zmiennymi opisanymi za pomocą danych panelowych, powinno się jednak robić to z uwzględnieniem ich panelowego charakteru.

Celem artykułu jest pokazanie nieprzydatności współczynnika korelacji liniowej Pearsona do analizy zależności pomiędzy zmiennymi o panelowej strukturze danych oraz próba zaproponowania bardziej adekwatnej miary. Prowadzony wywód został

¹ W związku z czym można zapytać o celowość szacowania modeli panelowych w dalszej części artykułu.

zilustrowany obliczeniami zależności pomiędzy strukturą kapitału oraz rentownością i wielkością 17 spółek budowlanych notowanych na GPW w Warszawie w latach 2009–2018 (170 obserwacji), tworzących panel zbilansowany.

2. Podstawowe pojęcia

Dane panelowe mają jednocześnie cechy danych przekrojowych (opisujących zbiorowość w jednym momencie) i szeregów czasowych (opisujących jednostkę w różnych okresach). Jeżeli we wszystkich okresach obserwowane są te same jednostki, to mamy do czynienia z panelem zbilansowanym, a jeżeli dla niektórych okresów brakuje danych o wszystkich jednostkach – z panelem niezbilansowanym.

Podstawową zaletą danych panelowych, w porównaniu np. z danymi przekrojowymi, jest większa liczba informacji o tych samych obiektach. Dane panelowe umożliwiają jednoczesną obserwację zróżnicowania analizowanych obiektów oraz ich ewolucję w czasie, a zatem lepsze rozpoznanie badanego zjawiska. Pozwalają także na kontrolowanie i/lub identyfikację nieobserwowalnych efektów indywidualnych w modelach regresji. W związku z tym dzięki wykorzystaniu panelu możliwe staje się usunięcie obciążenia estymatora z powodu pominięcia ważnego czynnika, jakim jest stały, indywidualny (specyficzny) efekt charakterystyczny dla każdej jednostki (Witkowski, 2012, s. 268–269). Dane panelowe zapewniają znacznie większą liczbę obserwacji, co poprawia precyzję wnioskowania oraz umożliwia oszacowanie dynamiki zjawisk, nawet gdy liczba okresów jest niewielka.

Współczynnik korelacji liniowej Pearsona² pozwala zmierzyć siłę i kierunek zależności pomiędzy dwiema jednowymiarowymi (w szeregu czasowym lub przekrojowym) zmiennymi, przede wszystkim gdy zakłada się, że jest to zależność liniowa. Stąd też, jeżeli mamy panel zbilansowany N jednostek w T okresach, to w formule współczynnika korelacji liniowej każdą z $N \cdot T$ obserwacji traktuje się niezależnie od tego, jakiego obiektu i jakiego okresu dotyczy. Współczynnik korelacji liniowej Pearsona jest więc interpretowany jako miara zależności pomiędzy dwiema zmiennymi w jednym obiekcie w $N \cdot T$ okresach (np. latach) lub jako miara zależności w $N \cdot T$ jednostkach w jednym okresie (np. roku). Przykładowo jeżeli mamy panel 100 obiektów w ciągu 10 lat (a więc 1000 obserwacji), to współczynnik korelacji możemy traktować jako miarę zależności dwóch zmiennych w 1000 obiektów w jednym roku (co w praktyce zachodzi dość często) lub w jednym obiekcie w ciągu 1000 lat (co jest w zasadzie niemożliwe).

Współczynnik korelacji liniowej Pearsona nie uwzględnia podstawowej zalety danych panelowych, jaką jest dostarczanie pełniejszej informacji o badanym zjawie-

² W 1895 r. Karl Pearson podał ostateczną matematyczną formułę współczynnika korelacji liniowej, który odtąd nazywany jest współczynnikiem Pearsona.

sku. Dla jego obliczenia nie jest istotne, że pewna obserwacja dotyczy obiektu i w okresie t , a inna – obiektu j w okresie $t + 1$, a to ma zasadnicze znaczenie w analizie danych panelowych. Tak więc obliczanie i wnioskowanie o zależnościach pomiędzy zmiennymi o panelowej strukturze danych na podstawie macierzy liniowych współczynników korelacji Pearsona jest niewłaściwe.

3. Zastosowanie współczynnika korelacji liniowej Pearsona do badania zależności w przypadku danych panelowych

Nieprzydatność współczynnika korelacji liniowej Pearsona do obliczania zależności pomiędzy zmiennymi o panelowej strukturze danych nie oznacza, że nie znajduje on zastosowania w analizach panelowych. Można go wykorzystać do pewnych obliczeń częściowych.

Do unaocznienia problemu niech posłuży następujący przykład. Analizie poddano zależności finansowe w zbilansowanym panelu 17 spółek budowlanych notowanych na GPW w Warszawie w latach 2009–2018 (170 obserwacji). Badano zależności struktury kapitału spółki mierzonej relacją zadłużenia ogółem do aktywów ogółem (CS) względem wielkości tej spółki mierzonej logarytmem naturalnym z wartości aktywów ogółem tego podmiotu (SIZE) oraz rentowności spółki mierzonej wskaźnikiem ROA.

Tabl. 1. Wartości współczynnika korelacji liniowej Pearsona (r) pomiędzy CS oraz SIZE i ROA (dla 170 obserwacji, bez uwzględnienia struktury panelowej danych)

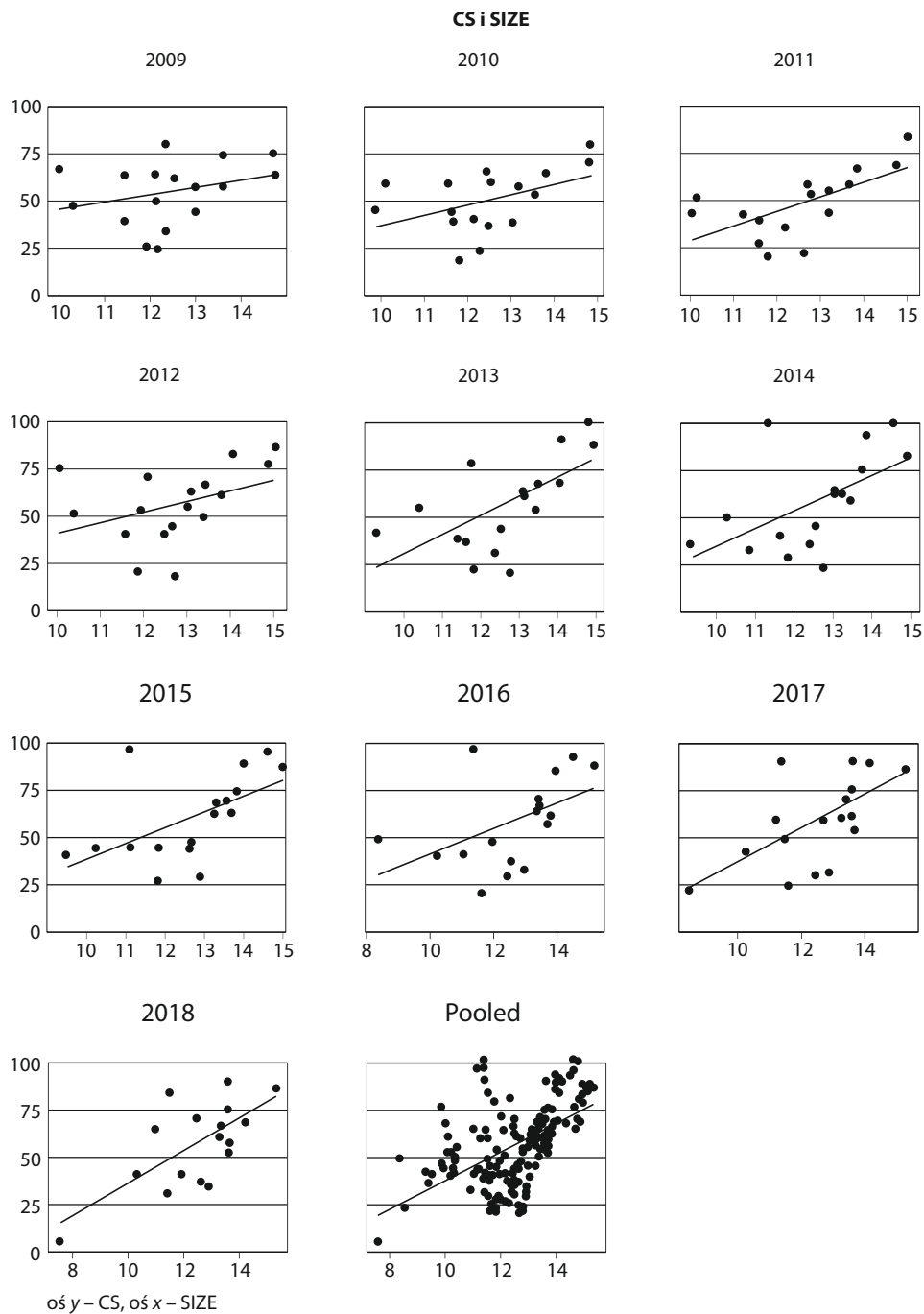
Zmienne	r	p
SIZE	0,5363	<0,0001
ROA	-0,0737	0,3384

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.).

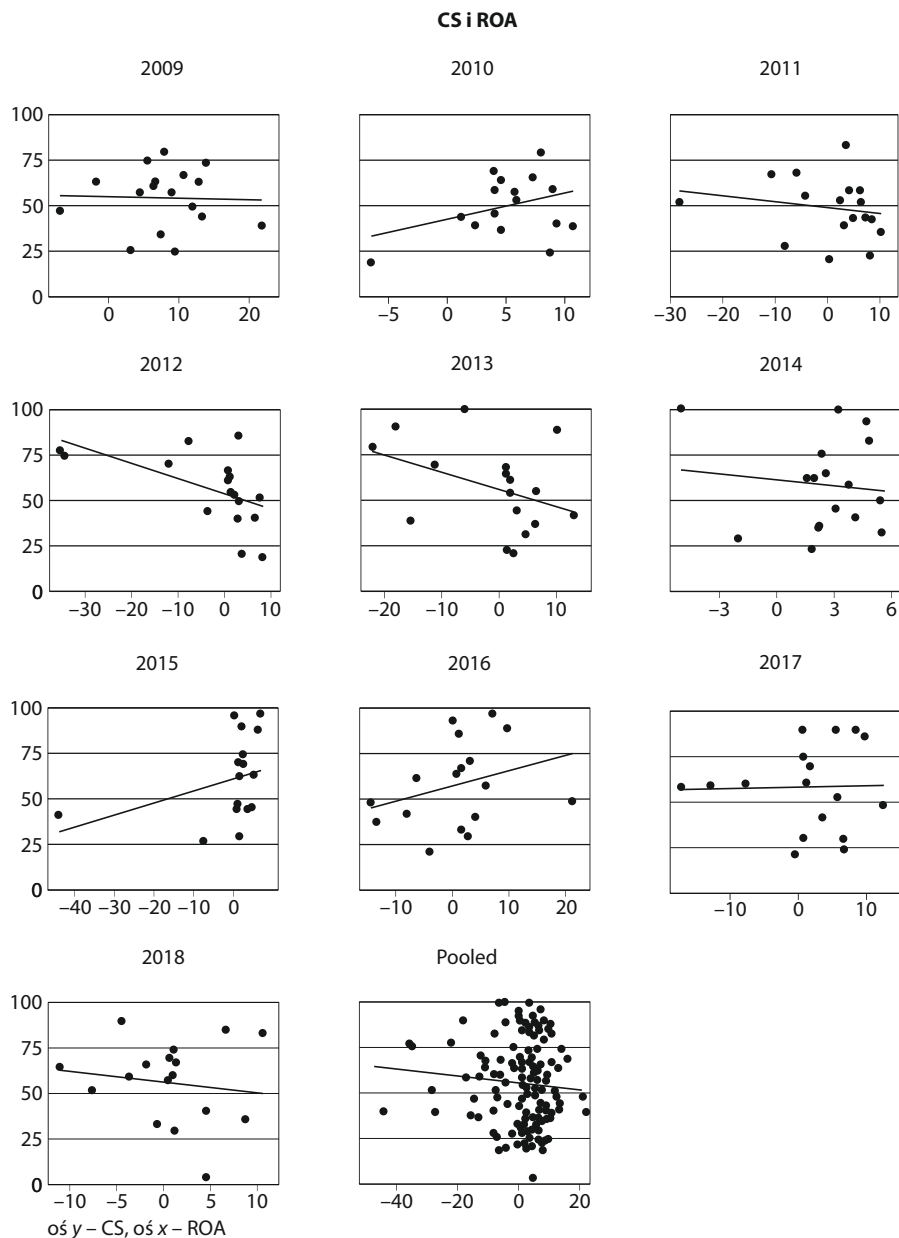
Obliczone wartości współczynnika korelacji (tabl. 1), zgodnie z wcześniejszym wywoдем, odzwierciedlają zależności występujące dla jednej hipotetycznej spółki w ciągu 170 lat lub 170 hipotetycznych spółek w jednym roku. Z formalnego punktu widzenia warto odnotować, że zależność pomiędzy CS i SIZE jest istotna na poziomie <0,0001, natomiast pomiędzy CS i ROA – nieistotna³. Zbiór danych to jednak panel zbilansowany 17 spółek w ciągu 10 lat, a zaprezentowane współczynniki nie uwzględniają tego faktu.

³ W artykule do oceny istotności przyjęto poziom $p = 0,05$.

Wykr. 1. Wykresy rozrzutu z uwzględnieniem struktury panelowej wraz z ilustracją liniowej korelacji pomiędzy zmiennymi CS i SIZE oraz CS i ROA dla spółek w poszczególnych latach



Wykr. 1. Wykresy rozrzutu z uwzględnieniem struktury panelowej wraz z ilustracją liniowej korelacji pomiędzy zmiennymi CS i SIZE oraz CS i ROA dla spółek w poszczególnych latach (dok.)



Uwaga. Wykresy zatytułowane „Pooled” dotyczą wszystkich obserwacji bez uwzględnienia struktury panelowej.

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.) z wykorzystaniem pakietu oprogramowania R.

Korzystając z danych panelowych, można natomiast obliczyć wartości współczynnika korelacji liniowej Pearsona pomiędzy wyspecyfikowanymi zmiennymi, które są:

- „przekrojowe” – w każdym roku ($T = 10$) dla wszystkich spółek ($N = 17$);
- „czasowe” (dynamiczne) – dla każdej spółki w ciągu 10 lat.

Tabl. 2. Wartości współczynnika korelacji pomiędzy CS oraz SIZE i ROA w poszczególnych latach

L a t a	CS i SIZE		CS i ROA	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
2009	0,2968	0,2474	-0,0250	0,9241
2010	0,4496	0,0702	0,3643	0,1506
2011	0,6379	0,0059	-0,1938	0,4560
2012	0,3868	0,1251	-0,5581	0,0199
2013	0,6176	0,0082	-0,3934	0,1182
2014	0,5518	0,0216	-0,1177	0,6528
2015	0,5606	0,0192	0,3401	0,1817
2016	0,4825	0,0498	0,3034	0,2365
2017	0,6091	0,0095	0,0327	0,9010
2018	0,6759	0,0029	-0,1362	0,6023

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.).

W poszczególnych latach wartości „przekrojowego” współczynnika korelacji są zróżnicowane, co oznacza, że okres analizy ma wpływ na wyniki. W przypadku korelacji pomiędzy CS i SIZE we wszystkich analizowanych okresach zależności były dodatnie i w siedmiu latach istotne na poziomie $<0,05$. Natomiast w przypadku korelacji pomiędzy CS i ROA w sześciu okresach wartości te były ujemne, w tym w jednym roku współczynnik okazał się istotny, podczas gdy w czterech latach zaobserwowano dodatnie, ale nieistotne statystycznie wartości tego współczynnika (tabl. 2, wyk. 1).

Tabl. 3. Wartości współczynnika korelacji pomiędzy CS oraz SIZE i ROA w latach 2009–2018 według spółek

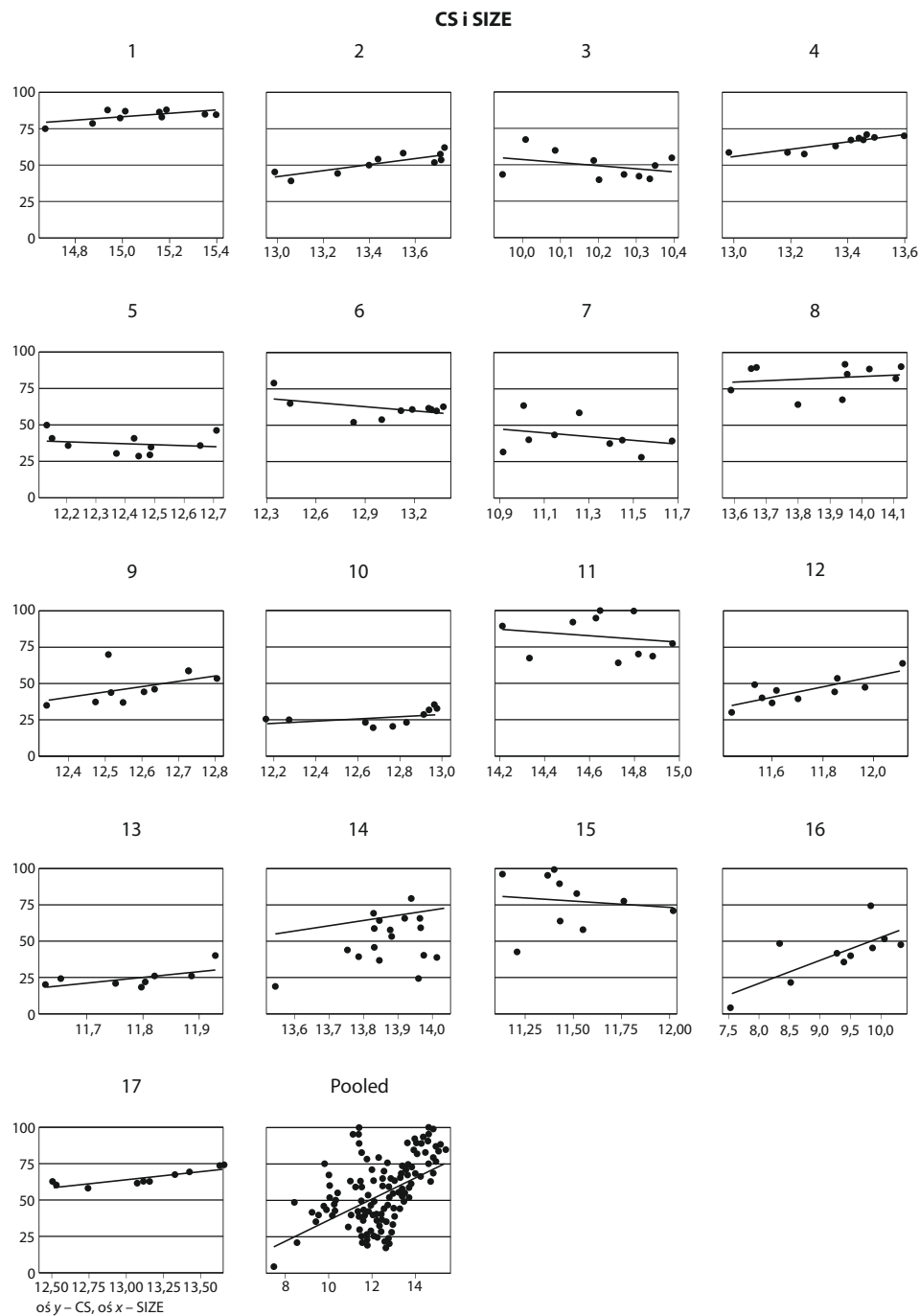
Lp.	Spółki	CS i SIZE		CS i ROA	
		<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
1	Budimex	0,6588	0,0383	0,3130	0,3785
2	Elbudowa	0,8459	0,0020	–0,4080	0,2419
3	Enap	–0,3610	0,3054	0,9552	<0,0001
4	Erbud	0,8757	0,0009	–0,2013	0,5771
5	Instalkrk	–0,1981	0,5832	0,3427	0,3323
6	Mirbud	–0,5376	0,1090	0,5117	0,1306
7	Mostalplc	–0,3129	0,3788	–0,3397	0,3368
8	Mostalwar	0,2084	0,5634	–0,1734	0,6319
9	Mostalzab	0,4370	0,2066	–0,2449	0,4952
10	Panova	0,4487	0,1934	–0,5959	0,0691
11	Polimexms	–0,1784	0,6220	–0,0585	0,8725
12	Prochem	0,8014	0,0053	–0,1913	0,5965
13	Projprzem	0,6758	0,0320	0,3332	0,3469
14	Rafako	0,7082	0,0219	–0,2073	0,5655
15	Remak	–0,1276	0,7255	0,0104	0,9774
16	Resbud	0,7331	0,0158	–0,4069	0,2432
17	Unibep	0,9046	0,0003	–0,5848	0,0758

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.).

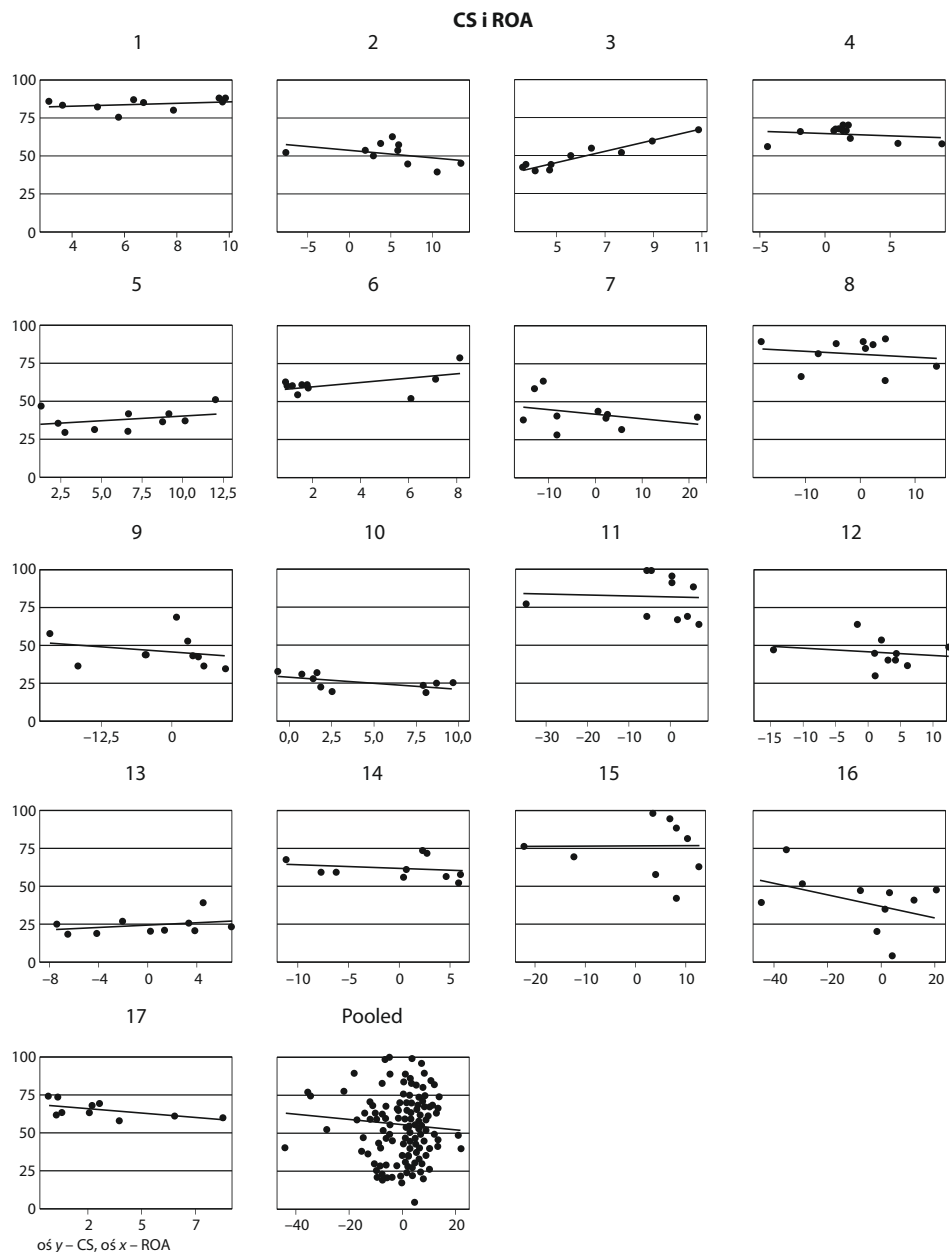
Zróznicowane są również wartości „czasowego” współczynnika korelacji. W przypadku korelacji pomiędzy CS i SIZE w latach 2009–2018 dla sześciu spółek wartości te były ujemne, ale nieistotne, a dla 11 spółek – dodatnie, w tym dla ośmiu istotne na poziomie 0,05. Natomiast zależności pomiędzy CS i ROA dla 11 spółek były ujemne, ale nieistotne, a dla sześciu spółek – dodatnie, w tym dla jednej współczynnik korelacji był istotny (tabl. 3, wykr. 2).

Powyższa analiza nie daje syntetycznej odpowiedzi na pytanie, jakie są kierunek i siła zależności pomiędzy analizowanymi zmiennymi w badanym panelu.

Wykr. 2. Wykresy rozrzutu z uwzględnieniem struktury panelowej wraz z ilustracją liniowej korelacji pomiędzy zmiennymi CS i SIZE oraz CS i ROA według spółek



Wykr. 2. Wykresy rozrzutu z uwzględnieniem struktury panelowej wraz z ilustracją liniowej korelacji pomiędzy zmiennymi CS i SIZE oraz CS i ROA według spółek (dok.)



Uwaga. Numery od 1 do 17 oznaczają kolejne spółki (zob. tabl. 3). Wykresy zatytułowane „Pooled” dotyczą wszystkich obserwacji bez uwzględnienia struktury panelowej.

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.) z wykorzystaniem pakietu oprogramowania R.

4. Propozycje pomiaru zależności pomiędzy zmiennymi o panelowej strukturze danych

Przedstawiona argumentacja nie oznacza, że nie należy badać zależności pomiędzy poszczególnymi zmiennymi, które są przewidziane do modelowania panelowego. Badanie zależności daje możliwość wyboru właściwego zestawu zmiennych objaśniających do modelu ekonometrycznego, jak również zbadania koincydencji parametrów (Hellwig, 1976). Jednak narzędzia oceny zależności muszą być adekwatne do panelowego charakteru danych. Nie ma panelowego współczynnika korelacji analogicznego do współczynnika Pearsona, ale podejmowane są próby rozwiązania tego problemu.

4.1. Metoda Famy-MacBetha

Jako pierwszą warto omówić metodę Famy-MacBetha (Fama i MacBeth, 1973), polegającą na szacowaniu w każdym roku parametrów (w tym wypadku są to wartości współczynnika korelacji liniowej Pearsona) na podstawie danych przekrojowych, a następnie – badaniu za pomocą testu *t*-Studenta istotności średnich wartości parametrów z całego analizowanego okresu. Fama i French (2002, s. 11–12) zaproponowali dodatkowo, aby w procesie wnioskowania wartość krytyczną statystyki *t* powiększyć 2,5 razy ze względu na autokorelację w czasie.

Już w latach 70. XX w. znane i stosowane były metody estymacji modeli panelowych z efektami ustalonymi (*least square with dummy variables* – LSDV) albo losowymi (*generalized least squares* – GLS) (Maddala, 2006, s. 644–648). Na początku lat 80. Anderson i Hsiao (1981) opracowali metodę estymacji dynamicznych modeli panelowych, a 10 lat później Arellano i Bond (1991) przedstawili uogólnioną metodę momentów (*general moments method*). Ale finansiści jeszcze w pierwszej dekadzie XXI w. stosowali metodę Famy-MacBetha do oceny istotności parametrów modeli panelowych⁴.

Tabl. 4. Test *t*-Studenta istotności średniego współczynnika korelacji liniowej Pearsona w latach 2009–2018

Wyszczególnienie	CS i SIZE	CS i ROA
Średni <i>r</i>	0,5269	–0,0384
Statystyka <i>t</i> (9)	13,7686	0,3918

Uwaga. Przy poziomie istotności 0,05 dla dziewięciu stopni swobody wartość statystyki *t* wynosi 2,2622. Ze względu na autokorelację wartość krytyczna statystyki *t* wynosi $2,2622 \cdot 2,5 = 5,6555$.

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.).

Na podstawie kryterium zaproponowanego przez Famę i Frencha można stwierdzić, że tylko zależność pomiędzy CS i SIZE okazała się istotna (w rzeczywistości współczynnik korelacji jest istotny na poziomie 0,001 – zob. tabl. 4).

⁴ Jako przykłady można podać prace: Booth i Zhou (2008), Chay i Suh (2009), Denis i Osobov (2008), Fama i French (2001).

Metoda obliczania współczynnika korelacji dla danych panelowych według Famy-MacBetha ma dość poważną wadę. W przypadku gdy w poszczególnych grupach panelu występują korelacje dodatnie i ujemne, podczas obliczania średniej ich wartości mogą się wzajemnie „znosić”, czego rezultatem jest ostateczny wynik bliski 0 (por. tabl. 4 dla korelacji CS i ROA), pomimo dość wysokich wartości dla poszczególnych grup. Rozwiązaniem tego problemu może być obliczanie średniej na podstawie wartości bezwzględnych współczynników korelacji uzyskanych na pierwszym etapie. Dzięki temu wyeliminowane zostaje „znoszenie się” wartości o przeciwnych znakach, a ostateczny współczynnik znacznie lepiej informuje o sile zależności między dwiema cechami z uwzględnieniem ich panelowej struktury. Odbywa się to kosztem utraty informacji o kierunku zależności, niemniej jednak autorzy niniejszej pracy proponują, żeby metodę Famy-MacBetha skorygować w opisany wyżej sposób i obliczać średnią na podstawie wartości bezwzględnych otrzymanych współczynników korelacji. Końcowy współczynnik będzie miarą unormowaną w przedziale $[0, 1]$, informującą o sile zależności liniowej pomiędzy dwiema zmiennymi, z uwzględnieniem panelowego charakteru danych, chociaż niepokazującą kierunku tej zależności.

Tabl. 5. Absolutna średnia wartość współczynnika korelacji według skorygowanej metody Famy-MacBetha w latach 2009–2018

Wyszczególnienie	CS i SIZE	CS i ROA
Średni r	0,5269	0,2465
Statystyka t (9)	13,7686	4,4928

Uwaga. Przy poziomie istotności 0,05 dla dziewięciu stopni swobody wartość statystyki t wynosi 2,2622. Ze względu na autokorelację wartość krytyczna statystyki t wynosi $2,2622 \cdot 2,5 = 5,6555$.

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.).

W przypadku zależności pomiędzy CS i SIZE – dla których wszystkie wartości współczynnika korelacji były dodatnie – wartość statystyki empirycznej się nie zmieniła. W przypadku zależności CS i ROA wartość statystyki t znacznie wzrosła, ale po zastosowaniu zaproponowanego przez Famę i Frencha skorygowanego kryterium istotności zależność dla wskazanych zmiennych była nadal nieistotna (tabl. 5).

4.2. Dobór zmiennych do mierników taksonomicznych konstruowanych na podstawie danych panelowych

Młodak i in. (2016, s. 8–9) zaproponowali metodę doboru zmiennych do mierników taksonomicznych, która uwzględnia panelowy charakter danych. W tym celu dla każdego roku policzyli macierz współczynników korelacji liniowej Pearsona pomiędzy potencjalnymi zmiennymi objaśniającymi. Następnie skonstruowali kompleksową macierz z tych współczynników korelacji, które miały największą wartość bez-

względna. Dalej przyjęto postępowanie analogiczne do postępowania przy metodzie odwroconej macierzy korelacji (Malina i Zeliaś, 1998; Młodak, 2006; Polozij, 1966).

Kompleksowa macierz zaproponowana w tej metodzie nie jest co prawda macierzą korelacji, ponieważ nie zostały spełnione kryteria przechodniości skorelowania (Hellwig, 1976) i z tego powodu wartości na głównej przekątnej macierzy odwrotnej były nawet ujemne, niemniej jednak przy „wniesionej dozie subiektywizmu” (Młodak i in., 2016, s. 9) umożliwiła wybranie właściwych zmiennych o strukturze panelowej potrzebnych do konstrukcji miernika.

4.3. Współczynnik korelacji pomiędzy wartościami empirycznymi i teoretycznymi zmiennej objaśnianej modelu panelowego

Autorzy niniejszego artykułu proponują rozwiązanie polegające na szacowaniu modeli panelowych z jedną zmienną objaśniającą, którą jest kolejno każda ze zmiennych przeznaczonych do tej roli (w przykładzie przedstawionym w artykule: SIZE i ROA), oraz ocenie siły zależności za pomocą współczynnika korelacji liniowej Pearsona pomiędzy wartościami empirycznymi i teoretycznymi zmiennej objaśnianej (w przedstawionym przykładzie: CS). Kwadrat tego współczynnika jest często stosowany jako miara dopasowania modelu do danych empirycznych. Warto podkreślić, że w przypadku nieuwzględniania panelowego charakteru danych, a więc zastosowania metody najmniejszych kwadratów, współczynnik ten jest równy współczynnikowi korelacji liniowej Pearsona pomiędzy obiema zmiennymi, a jego kwadrat – współczynnikowi determinacji. Kierunek zależności określa się na podstawie znaku parametru przy zmiennej objaśniającej.

Każdy z modeli należałoby oszacować dwukrotnie: jako model z ustalonymi oraz losowymi efektami indywidualnymi i czasowymi. W omawianym przykładzie szacowane więc będą modele zmiennej CS w zależności od SIZE i od ROA, każdy w dwóch alternatywnych wersjach (łącznie cztery modele). Oczywiście należy się spodziewać innych wyników w zależności od przyjętego założenia o charakterze efektów, które mogą być stałe lub losowe⁵.

Obliczenia rozpoczyna się od określenia zależności pomiędzy CS i SIZE. Oszacowany model, ze stałymi efektami indywidualnymi i czasowymi CS względem SIZE, pokazuje, że efekty indywidualne dla spółek w sposób istotny wpływały na zależność pomiędzy tymi zmiennymi (co oznacza odrzucenie hipotezy zerowej o wspólnym wyrazie wolnym), natomiast wpływ efektów czasowych był nieistotny na poziomie 0,05. Współczynnik korelacji wartości empirycznej i teoretycznej zmiennej objaśnianej CS wynosi 0,9077 ($p < 0,0001$). Jest on więc znacznie większy niż współczynnik

⁵ Stanowi to problem, jeżeli ta procedura miałaby zostać zastosowana do doboru zmiennych, ponieważ może się okazać, że niektóre potencjalne zmienne objaśniające są silnie powiązane ze zmienną objaśnianą w przypadku założenia stałych efektów, a inne – w przypadku efektów losowych.

korelacji obliczony dla obu zmiennych przy założeniu, że dane nie mają charakteru panelowego (np. pochodzą ze 170 spółek w jednym roku). Oszacowany model z losowymi efektami indywidualnymi i czasowymi CS względem SIZE pokazuje również, że tylko efekty indywidualne istotnie wpływały na zależność pomiędzy tymi zmiennymi (tzn. brakuje podstaw do odrzucenia hipotezy zerowej o zgodności estymatora, czyli do odrzucenia hipotezy o efektach losowych – test Hausmana). W tym przypadku współczynnik korelacji wartości empirycznej i teoretycznej zmiennej objaśnianej CS wynosi 0,8959 ($p < 0,0001$). Jest on tylko nieznacznie mniejszy niż współczynnik korelacji obliczony dla modelu z ustalonymi efektami.

Z oszacowanych modeli o ustalonych i losowych efektach wynika, że parametr przy zmiennej SIZE jest dodatni i istotny statystycznie, a więc zależność pomiędzy CS i SIZE jest dodatnia i istotna. Jednocześnie siła tej zależności znacznie wzrosła (w porównaniu z zależnością mierzoną współczynnikiem korelacji liniowej Pearsona) po uwzględnieniu panelowego charakteru danych (tabl. 6, wyk. 3).

Tabl. 6. Panelowe zależności pomiędzy CS (zmienna objaśniana) oraz SIZE i ROA (zmiennie objaśniające); modele z jedną zmienną

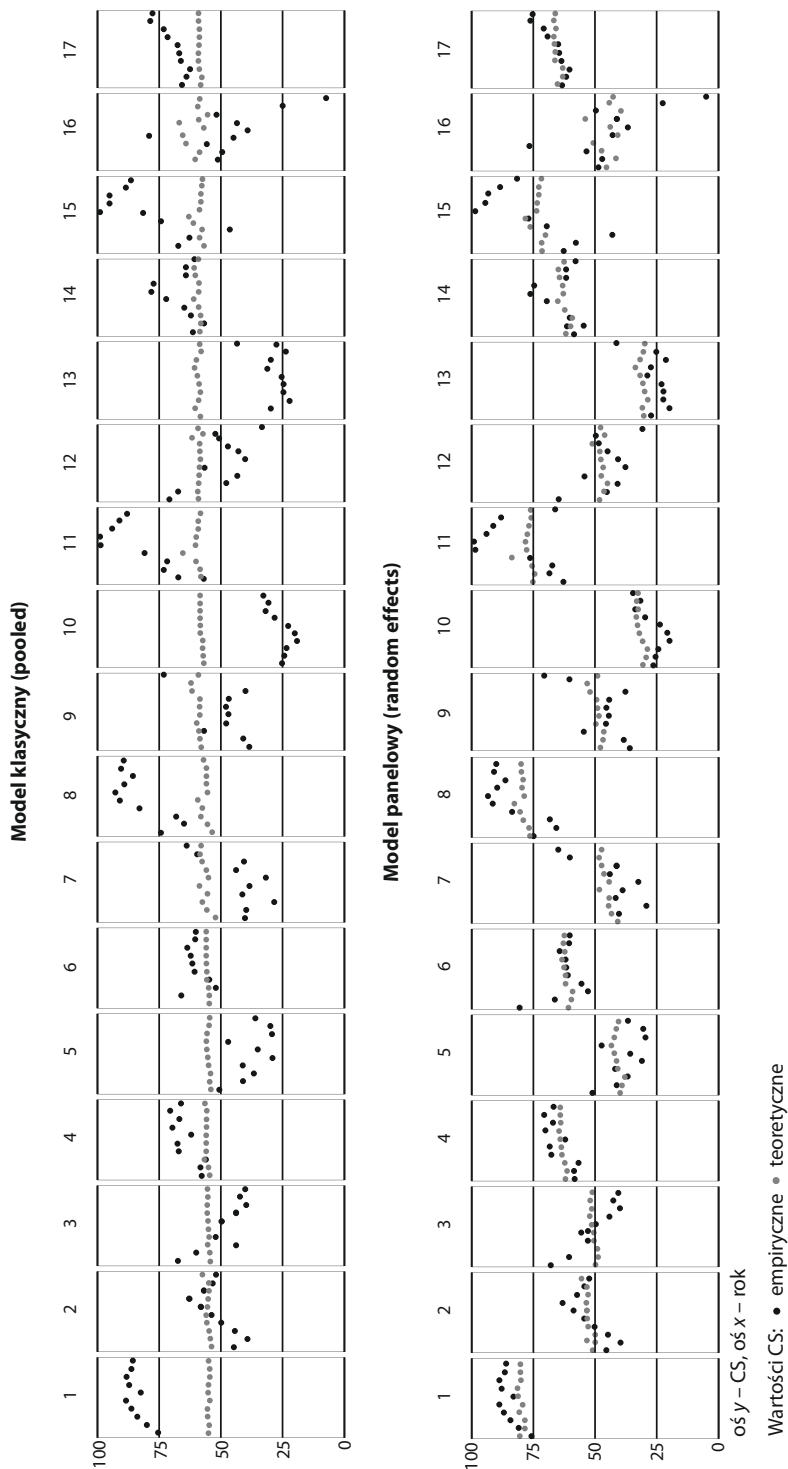
Modele	Wartość parametru przy zmiennej objaśniającej	Współczynnik korelacji Pearsona między wartościami teoretycznymi i empirycznymi CS ^a	Test Lagrange'a na istotność efektów. Hipoteza zerowa: efekty nieistotne. Statystyka testu LM ^b		Test Hausmana. Hipoteza zerowa: estymator GLS jest zgodny. Statystyka testu chi-kwadrat
			efekty indywidualne	efekty czasowe	
SIZE					
Pooled ^c	7,5195 ($<0,0001$)	0,5363 ($<0,0001$)	–	–	} 0,1173 (0,7320)
Panelowy: FE ^d	9,3440 (0,0391)	0,9077 ($<0,0001$)	19,2000 ($<0,0001$)	–1,2388 (0,8923)	
RE ^e	8,7319 ($<0,0001$)	0,8959 ($<0,0001$)	–	–	
ROA					
Pooled ^c	–0,1749 (0,3829)	0,0737 (0,3394)	–	–	} 0,0480 (0,8266)
Panelowy: FE ^d	–0,1989 (0,0297)	0,9013 ($<0,0001$)	21,1840 ($<0,0001$)	–1,4453 (0,9258)	
RE ^e	–0,2057 (0,0063)	0,8962 ($<0,0001$)	–	–	

a „Klasyczny” współczynnik korelacji liniowej Pearsona między dwiema zmiennymi. b Test Lagrange'a Multiplier z efektami two-way (Honda) dla paneli zbilansowanych. c Parametry szacowane klasyczną metodą najmniejszych kwadratów (KMNK), odporne błędy standardowe (HC1). d Model panelowy z ustalonymi efektami indywidualnymi i czasowymi, estymator KMNK na danych przekształconych, odporne błędy standardowe (HC1, Arellano, 1987). e Model panelowy z losowymi efektami two-way, estymator według metody Swamy-Arora, odporne błędy standardowe (HC1, Arellano, 1987).

Uwaga. W nawiasach wartości poziomu istotności p .

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.).

Wykr. 3. Zależność pomiędzy wartościami empirycznymi i teoretycznymi zmiennej CS w modelach ze zmienną objaśniającą ROA według szeregów czasowych dla poszczególnych spółek



Uwaga. Numery od 1 do 17 oznaczają kolejne spółki (zob. tabl. 3).

Źródło: opracowanie własne na podstawie danych z Bankier.pl (b.r.) z wykorzystaniem pakietu oprogramowania R.

Podobne wyniki otrzymano w przypadku zależności pomiędzy CS i ROA. Warto jednak zauważyć, że współczynnik korelacji nieuwzględniający struktury panelowej okazał się ujemny i nieistotny statystycznie. Po uwzględnieniu panelowej struktury zmiennych zależność nadal była ujemna (ze względu na ujemną wartość oszacowanego parametru przy ROA), ale istotna statystycznie. A zatem także w tym przypadku uwzględnienie panelowej struktury zmiennych znacznie poprawiło siłę zależności. Gdyby jako kryterium doboru zmiennych przyjąć istotność współczynnika korelacji Pearsona, to wskaźnik ROA nie znalazłby się wśród zmiennych objaśniających. Natomiast ujemna i istotna zależność pomiędzy strukturą kapitału mierzoną udziałem zadłużenia w aktywach i rentownością została potwierdzona bardzo wieloma badaniami przeprowadzonymi na danych panelowych (Nehrebecka i in., 2016, s. 22–49).

Niedoskonałością zaproponowanej miary jest (w przeciwieństwie do współczynnika korelacji Pearsona) jej niesymetryczność. Wyznaczona w ten sposób zależność pomiędzy SIZE i CS nie będzie taka sama, jak pomiędzy CS i SIZE.

5. Podsumowanie

W artykule autorzy wykazują, że w przypadku zmiennych o panelowej strukturze danych nie należy obliczać wartości liniowego współczynnika korelacji Pearsona, ponieważ nie odzwierciedlają one panelowego charakteru danych. Wniosek ten dla większości badaczy jest oczywisty, ale w związku z tym, że wciąż publikowane są artykuły uwzględniające takie podejście, warto zabrać głos w tej sprawie.

Obliczanie wartości współczynnika korelacji Pearsona dla zmiennych o panelowej strukturze danych jest prostym przeniesieniem metodologii modelowania na podstawie szeregów czasowych lub szeregów przekrojowych, które rozpoczyna się od prezentacji statystyk opisowych i macierzy korelacji; w takim przypadku szacowanie wartości współczynnika Pearsona jest w pełni uzasadnione i zazwyczaj pomaga w wyborze zmiennych objaśniających do modeli.

Oczywiście nie oznacza to, że w ogóle nie należy liczyć zależności pomiędzy parami zmiennych o panelowej strukturze danych. Takie miary można wykorzystać przy opisie analizowanych zjawisk, jak również podczas wstępnego doboru zmiennych do modeli panelowych. Jednak miary te muszą uwzględniać panelowy charakter danych. Niekiedy w takiej sytuacji wykorzystuje się współczynnik Pearsona do obliczeń cząstkowych (pośrednich).

Zaproponowana w artykule miara wykorzystująca współczynnik korelacji pomiędzy wartościami empirycznymi i teoretycznymi modelu panelowego z jedną zmienną objaśniającą została skonstruowana przede wszystkim w celu pokazania braku przydatności współczynnika korelacji Pearsona do pomiaru zależności pomiędzy zmiennymi o strukturze panelowej. Ta miara, jakkolwiek bardziej adekwatna, jest niedoskonała. Jej mankamentem jest to, że w przeciwieństwie do współczynnika korelacji Pearsona jest ona niesymetryczna. Problemy mogą się pojawiać również przy próbach zastosowania tej miary do procedury doboru zmiennych. Na przykład

gdy jedna zmienna zostanie wybrana na podstawie modelu z ustalonymi efektami (ponieważ zastosowane testy pokazały, że ten model będzie właściwy), a druga – na podstawie modelu z efektami losowymi, to pojawia się wątpliwość, który model należy zastosować do oszacowania modelu panelowego zmiennej objaśnianej względem obu zmiennych.

Do pomiaru zależności pomiędzy zmiennymi o strukturze panelowej potrzebna jest miara o cechach współczynnika korelacji liniowej Pearsona (unormowanie na przedział $[-1, 1]$, spełnienie zasady, zgodnie z którą im większa wartość bezwzględna, tym silniejsza zależność), a jednocześnie uwzględniająca panelowy charakter danych. Taką miarę można by nazwać „panelowym współczynnikiem korelacji”.

Artykuł stanowi przyczynek do dyskusji nad badanym zagadnieniem. Nie przedstawiono w nim rozwiązań najodpowiedniejszych dla badania zależności między zmiennymi o panelowej strukturze danych, niemniej jednak zasugerowano pewne rozwiązania. Zwrócono przy tym uwagę na problem metodologiczny, który wymaga dalszych prac.

Bibliografia

- Anderson, T. W., Hsiao, Ch. (1981). Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association*, 76(375), 598–606. <https://doi.org/10.2307/2287517>.
- Arellano, M. (1987). Computing Robust Standard Errors for Within-groups Estimators. *Oxford Bulletin of Economics and Statistics*, 49(4), 431–434. <https://doi.org/10.1111/j.1468-0084.1987.mp49004006.x>.
- Arellano, M., Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2), 277–297. <https://doi.org/10.2307/2297968>.
- Bankier.pl. (b.r.). *Notowania GPW – akcje*. Pobrane 20 lutego 2020 r. z <https://www.bankier.pl/gielda/notowania/akcje>.
- Booth, L., Zhou, J. (2008). *Market Power and Dividend Policy: A Risk-Based Perspective*. <https://dx.doi.org/10.2139/ssrn.1296940>.
- Borsuk, M., Kostrzewa, K. (2020). Miary ryzyka systemowego dla Polski. Jak ryzyko systemowe wpływa na akcję kredytową banków?. *Bank i Kredyt*, 51(3), 211–238. https://bankikredyt.nbp.pl/content/2020/03/BIK_03_2020_01.pdf.
- Chay, J. B., Suh, J. (2009). Payout Policy and Cash-Flow Uncertainty. *Journal of Financial Economics*, 93(1), 88–107. <https://doi.org/10.1016/j.jfineco.2008.12.001>.
- Denis, D., Osobov, I. (2008). Why do firms pay dividends? International evidence on the determinants of dividend policy. *Journal of Financial Economics*, 89(1), 62–82. <https://doi.org/10.1016/j.jfineco.2007.06.006>.
- Driver, C., Grosman, A., Scaramozzino, P. (2020). Dividend policy and investor pressure. *Economic Modelling*, 89, 559–576. <https://doi.org/10.1016/j.econmod.2019.11.016>.
- Fama, E. F., French, K. R. (2001). Disappearing dividends: changing firm characteristics or lower propensity to pay?. *Journal of Financial Economics*, 60(1), 3–43. <https://doi.org/10.1111/j.1745-6622.2001.tb00321.x>.

- Fama, E. F., French, K. R. (2002). Testing Trade-Off and Pecking Order Predictions About Dividends and Debt. *The Review of Financial Studies*, 15(1), 1–33. <https://doi.org/10.1093/rfs/15.1.1>.
- Fama, E. F., MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81(3), 607–636. <https://doi.org/10.1086/260061>.
- Hellwig, Z. (1976). Przechodniość relacji skorelowania zmiennych losowych i płynące stąd wnioski ekonometryczne. *Przegląd Statystyczny*, 23(1), 3–20.
- Herman, S. (2019). Impact of joint-stock companies' financial condition on real activities manipulation to manage earnings. *Wiadomości Statystyczne. The Polish Statistician*, 64(10), 36–52. <https://doi.org/10.5604/01.3001.0013.7588>.
- Karkowska, R. (2019). Systematic risk affected by country level development. The case of the European banking sector. *Argumenta Oeconomica*, (2), 255–282. <https://doi.org/10.15611/aoe.2019.2.11>.
- Maddala, G. S. (2006). *Ekonometria*. Warszawa: Wydawnictwo Naukowe PWN.
- Malina, A. (2002). Wielokryterialna taksonomia w analizie porównawczej struktur gospodarczych Polski. W: A. Zeliaś (red.), *Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych* (s. 305–312). Kraków: Wydawnictwo Akademii Ekonomicznej.
- Malina, A., Zeliaś, A. (1998). On Building Taxonomic Measures on Living Conditions. *Statistics in Transition*, 3(3), 523–544.
- Młodak, A. (2006). *Analiza taksonomiczna w statystyce regionalnej*. Warszawa: Difin.
- Młodak, A., Józefowski, T., Wawrowski, Ł. (2016). Zastosowanie metod taksonomicznych w estymacji wskaźników ubóstwa. *Wiadomości Statystyczne*, 61(2), 1–24. <https://doi.org/10.5604/01.3001.0014.0900>.
- Nehrebecka, N., Białek-Jaworska, A., Dzik-Walczak, A. (2016). *Źródła finansowania przedsiębiorstw. Stan badań i ich metaanaliza*. Warszawa: Difin.
- Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58(347–352), 240–242. <https://doi.org/10.1098/rsp1895.0041>.
- Pluskota, A. (2020). The Impact of Corruption on Economic Growth and Innovation in an Economy in Developed European Countries. *Annales Universitatis Mariae Curie-Skłodowska. Sectio H. Oeconomia*, 54(2), 77–87. <https://doi.org/10.17951/h.2020.54.2.77-87>.
- Polożij, G. N. (1966). *Metody przybliżonych obliczeń*. Warszawa: Wydawnictwa Naukowo-Techniczne.
- Urbanek, G. (2017). Analysing brand strength – corporate financial performance link for companies listed on the Warsaw Stock Exchange. *Ekonometria*, (2), 92–102. <https://doi.org/10.15611/ekt.2017.2.06>.
- Witkowski, B. (2012). Modele danych panelowych. W: M. Gruszczyński (red.), *Mikroekonometria. Modele i metody analizy danych indywidualnych* (s. 267–308). Warszawa: Oficyna Wolters Kluwer Business.

Assessment of the risk of foreign divestment in Poland during the COVID-19 pandemic

Marcin Salamaga^a

Abstract. The importance of foreign direct investment (FDI) to the economic development of Poland cannot be overestimated, both at the regional level and in relation to the economy as a whole. Since FDIs are powered by capital sensitive to various national and international crises, it seems natural to ask whether the situation connected with the global COVID-19 pandemic is reflected in the reduction of FDI inflows to Poland. The aim of the paper is to identify the determinants of the foreign divestment process in the Polish economy as a whole and in its main sectors during the COVID-19 pandemic. The article analyses scenarios of foreign divestment in Poland in an annual perspective, starting from the outbreak of the epidemic in March 2020 up to February 2021. The study used data from a survey conducted in April and May 2020 among nearly 500 enterprises realising FDI in Poland. The benchmark for the surveyed companies was the level of their involvement in FDIs covering a one-year period prior to the announcement of the epidemic. The application of logit models allowed the identification of the most important factors of foreign divestment during COVID-19, including the location of FDIs in the services sector, industry, the IT sector, increased market openness and interactions of variables taking into account the restrictions introduced to the economy due to the pandemic. The level of risk of divestment of these variables depends, however, on the volume of FDI reductions declared by investors and on the sector of the economy. If considerable divestment is assumed, FDIs in the services sector are then burdened with a higher risk of divestment than FDIs in the processing industry. Assuming small divestments, FDIs in the IT sector constitute a factor bearing the greatest risk of FDI reduction in the entire economy.

Keywords: logit model, interactions of variables, foreign direct investment, FDI, divestment, COVID-19, risk assessment

JEL: C25, F21

Ocena ryzyka dezinwestycji zagranicznych w Polsce w warunkach pandemii COVID-19

Streszczenie. Znaczenie bezpośrednich inwestycji zagranicznych (BIZ) dla rozwoju gospodarczego Polski jest nie do przecenienia zarówno na poziomie regionalnym, jak i gospodarki jako całości. Ze względu na zasilanie BIZ kapitałem, który jest wrażliwy na różne kryzysy krajowe i międzynarodowe, naturalne wydaje się pytanie, czy sytuacja pandemiczna na świecie znajduje odzwierciedlenie w ograniczeniu napływu tych inwestycji do Polski. Celem badania przedstawionego w artykule jest identyfikacja czynników, które w największym stopniu wpływają na dezinwestycje w całej gospodarce i jej najważniejszych działach w warunkach pandemii COVID-19. Przeanalizowano scenariusze dezinwestycji zagranicznych w Polsce w perspektywie

^a Cracow University of Economics, College of Economics, Finance and Law, Poland / Uniwersytet Ekonomiczny w Krakowie, Kolegium Ekonomii, Finansów i Prawa, Polska.
ORCID: <https://orcid.org/0000-0003-0225-6651>. E-mail: salamaga@uek.krakow.pl.

rocznej od momentu ogłoszenia epidemii, czyli od marca 2020 r., do lutego 2021 r. Wykorzystano dane z badania sondażowego przeprowadzonego w kwietniu i maju 2020 r. wśród prawie 500 przedsiębiorstw, które lokowały BIZ w Polsce. Punktem odniesienia był poziom zaangażowania badanych firm w inwestycje bezpośrednie w ciągu roku poprzedzającego ogłoszenie epidemii. Dzięki zastosowaniu modeli logitowych udało się wskazać najistotniejsze czynniki dezinvestycji zagranicznych w dobie pandemii COVID-19. Należą do nich m.in.: lokowanie BIZ w sektorze usług, przemyśle i branży IT, wzrost otwartości rynku oraz interakcje zmiennych uwzględniające obostrzenia wprowadzane w gospodarce w związku ze stanem pandemii. Poziom ryzykogenność dezinvestycji tych zmiennych zależy jednak od rozmiarów redukcji BIZ deklarowanej przez inwestorów oraz od działu gospodarki. Wykazano m.in., że przy założeniu znacznych dezinvestycji lokowanie BIZ w usługach jest obciążone przeważnie wyższym ryzykiem dezinvestycji niż w przypadku przemysłu przetwórczego. Przy założeniu niewielkich dezinvestycji najbardziej ryzykownym czynnikiem redukcji BIZ w całej gospodarce jest lokowanie ich w branży IT.

Słowa kluczowe: model logitowy, interakcje zmiennych, bezpośrednie inwestycje zagraniczne, BIZ, dezinvestycje, COVID-19, ocena ryzyka

1. Introduction

The COVID-19 pandemic has had significant social and economic effects on the global economy and it will continue to be so in the future. The freezing of many economic sectors, the imposing of strict sanitary measures, and finally, the closing of borders have resulted in the reduced production capacity in various economic sectors, as well as the disruption of domestic and international supply chains. At the macroeconomic level, these repercussions have caused a decrease in gross domestic product (GDP), an increase in government debt due to the triggering of public aid measures (as part of anti-crisis shields and lowered tax income), an increased unemployment rate, more frequent enterprise bankruptcies, and further serious social and economic repercussions (Aidukaite et al., 2021; Bergsen, 2020; Davahli et al., 2020).

Regression has also affected FDI, which, due to the epidemic, has been limited, postponed or withdrawn in host countries and in countries providing capital. The scale of the reduced inflow of FDIs during the pandemic depends on various factors, e.g. on the structure of the economy, as it is evident that not all economic sectors have suffered from the crisis in the same way. Sanitary restrictions have hit certain sectors more, having led to a full ban on production or services, while others suffered less (where the risk of virus transmission is low or where the sector is of strategic importance to the functioning of the economy). The economic sectors most affected by the restrictions related to the COVID-19 pandemic include gastronomy, hospitality, health and wellness, while construction and industry were less affected (Bergsen, 2020; Davahli et al., 2020).

Thus, the situation caused by the pandemic prompts investors of many industries affected by COVID-19 to further restrict or suspend investment in sectors more

susceptible to the effects of the phenomenon, e.g. where potential losses related to a halt on production or services provision may be considerable. The high sensitivity of foreign capital to the COVID-19 pandemic also results from deep globalisation processes connected with the emergence of global value chains (GVC) taking advantage of international labour division. The scattering of the production process and services among countries in the framework of international labour division, supervised by transnational enterprises (offshoring), acts as a closed system: interference with a GVC link in one country results in immediate consequences in the other countries involved. Investments in enterprises which are elements of such international systems seem more risky during the pandemic and may incline investors to divest. Divestment involves restricting the previous scope and scale of operations of a business which is the target of direct investment as a result of abandoning part of its operations or a complete transfer of the enterprise by its investor (Borga et al., 2019; Doctor & Bagwell, 2020; Martins & Esteves, 2008; Shin, 2000; Trencher et al., 2020). In practical terms, divestment means a change of ownership (co-ownership), rather than the closing of a business. For this reason, it is necessary to devise and assess the factors that cause an increased investment risk during a pandemic.

Such research intensifies globally during local or global economic (financial) crises and, likewise, publications on divestment usually become more frequent. Analyses of this kind mostly cover microeconomic and macroeconomic factors. Berry (2010), Harrigan (1981), Markides (1992) and Pashley & Philippatos (1990) argue that poor results of the mother company (investor) in host countries of FDIs may also be conducive to divestment. However, no unambiguous relation between the financial results of affiliates and divestment has been confirmed (Berry, 2013; Hamilton & Chow, 1993; Markides, 1992). Bergh (1997) and Berry (2010, 2013) have shown that the diversification of business operations also encourages the withdrawal of FDIs. Furthermore, the level of the internationalisation of transnational enterprises which are chief providers of capital is an important diversification factor (Berry, 2010, 2013; Norbäck et al., 2015). Borga et al. (2019) prove that the risk of foreign divestment increases along with the intensification of enterprise internationalisation. Jovanovic & MacDonald (1994) prove that technological changes increase the odds of divestment, while Chatterjee et al. (2003) and Norbäck et al. (2015) argue that institutional changes at the sector level increase chances of divestment.

Research shows that the GDP, the level of economic openness, the level of salaries and wages, currency exchange rates, inflation, political stability, membership of a country in economic associations, free trade zones and others belong to important macro-economic factors which may have an influence on divestment (Berry, 2010; Blake & Moschieri, 2017; Norbäck et al., 2015). Berry (2010), Blake & Moschieri

(2017) and Norbäck et al. (2015) proved a negative relation between economic growth and divestment, while Borga et al. (2019) and Norbäck et al. (2015) demonstrated that a greater openness of an economy encourages divestment. The papers on divestment were written before COVID-19 was declared a pandemic, so researchers did not take into account the risk factors of divestment typical of this phenomenon and the consequences of the measures introduced to counteract it. In order to correctly diagnose divestment risk, one needs to consider the factors related to the direction of the development of the COVID-19 pandemic and the potential economic restrictions imposed by governments of various countries.

The aim of the paper is to identify the most important determinants of the foreign divestment process in the entire economy of Poland and its main sectors during the COVID-19 pandemic. The process involves the estimation of the divestment risk generated by individual macroeconomic variables. The article presents the application of a logit model with effects of interaction to study the inclination for foreign divestment in Poland. The analysis is based on data from a survey carried out among foreign enterprises conducting direct investment in Poland. Such research seems necessary to monitor tendencies in the investment market on an ongoing basis. It can also be used as potential support for decision-makers in the management of foreign investment.

2. Research methodology

The subject of the study concerned foreign companies which located FDIs in Poland in the years 2015–2019. The study used data from a survey conducted in April and May 2020 on a random sample of 708 randomly selected foreign subjects (simple random sampling without replacement). A questionnaire was sent to the selected entities in order to obtain information on how investors assess various risk factors leading to divestment in the context of different scenarios of the development of the COVID-19 pandemic. The return of the questionnaires reached the level of 70% and ultimately the study covered 496 correctly completed questionnaires, which constituted approximately 2% of the enterprises locating FDI in Poland. The article analyses scenarios of foreign divestment in Poland covering a year from the moment the epidemic was announced (from March 2020 to February 2021). The benchmark for the surveyed companies during the COVID-19 pandemic was the level of their involvement in FDIs during a one-year period preceding the announcement of the epidemic. The randomly selected companies represented main industries, including manufacturing, construction, IT, finance and more. The surveyed enterprises came from countries which have been in the recent years the main suppliers of FDI to

Poland: the Netherlands, Germany, France, Luxembourg,¹ and others. When preparing the research probe, the databases of such companies as Orbis (Bureau van Dijk, n.d. a) and Zephyr (Bureau van Dijk, n.d. b) were used. As the aim of the study was to estimate the level of divestment risk generated by different variables measured on the nominal and rank scale, a logit model with possible interactions between the variables was selected for the analysis. The ease of interpreting the results in the form of an odds ratio is also an advantage of using the logit model. The model was estimated for 3 divestment scenarios: small divestment (up to 20%), medium divestment (from 20% to 40%) and significant divestment (above 40%).

For each of the three divestment scenarios, the following coding was adopted for binary dependent variables:

- $Y_1 = 1$, when a decision on small foreign divestment was made and $Y_1 = 0$, when the decision on divestment was not made,
- $Y_2 = 1$, when the decision on medium divestment was made and $Y_2 = 0$, when the decision on divestment was not made,
- $Y_3 = 1$, when a decision on significant divestment was made and $Y_3 = 0$, when no decision on divestment was made.

The Chart below presents a list of all the variables along with the coding of the answers given by the respondents.

Chart. Specification of variables with questions and sets of possible answers included in the survey

Designation of the variable	Question	A set of possible answers
ΔGDP_c	expected decrease in GDP in the investor's country	1 – up to 5% 2 – over 5% to 10% 3 – over 10% to 15% 4 – over 15% to 20% 5 – over 20%
ΔGDP	expected decrease in GDP in Poland	
ΔEx	expected depreciation of the Polish zloty (PLN) in relation to the euro	
ΔOMI	expected increase in the market openness index	
ΔLC	expected increase in unit labour costs	
M	investing in the industrial sector and the construction industry	1 – yes 0 – no
S	investing in the services sector	
IT	investing in the IT sector	
O	investing in other sectors	

¹ Luxembourg's presence here may be due to the fact that it is considered to be a 'tax haven' (which has also been criticised by the European Commission). As a result, many of the actual sources of FDI may be hidden.

Chart. Specification of variables with questions and sets of possible answers included in the survey (cont.)

Designation of the variable	Question	A set of possible answers
Δi	expected increase in the inflation rate	1 – up to 5%
$\Delta R\&D$	expected decrease in domestic expenditures on research and development in relation to GDP	2 – over 5% to 10% 3 – over 10% to 15% 4 – over 15% to 20% 5 – over 20%
DC	distance from Warsaw to the capital of the investor's country	1 – up to 700 km 2 – over 700 km to 1400 km 3 – over 1400 km to 2100 km 4 – over 2800 km to 3500 km 5 – over 3500 km
$L1$	introduction of minor restrictions on the economy in relation to the epidemic risk	1 – yes 0 – no
$L2$	introduction of moderately burdensome restrictions on the Polish economy and a partial freeze of the economy in relation to the development of COVID-19	
$L3$	introduction of very rigorous restrictions on the Polish economy and a considerable economic freeze	

Source: own study based on the survey results.

Almost all of the specified variables relate to investors' expectations regarding changes in individual factors.

The basic logit model used in the paper considers second-order interaction and takes the following form (Harrell, 2001; Jaccard, 2001):

$$L(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \beta_{k+1} X_1 X_2 + \beta_{k+2} X_1 X_3 + \dots + \beta_{k+(k)_2} X_{k-1} X_k,$$

where:

$X_1, X_2, X_3, \dots, X_k$ – independent variables,

$\beta_1, \beta_2, \dots, \beta_{k+(k)_2}$ – model parameters,

$L(p)$ – logit, where $L(p) = \frac{p}{1-p}$ and $p = P(Y = 1 | X_1, X_2, \dots, X_k)$.

The parameters of the model have been estimated using the maximum likelihood method. The results of McFadden's pseudo R^2 index, the Likelihood Ratio Test (LRT), the Schwarz information criterion, and the statistical significance of parameters have been used to select the final model.

3. Determinants of the divestment process in various sectors of the Polish economy

The results of the estimation of the logit models for three divestment variants in Poland are provided in Table 1. The majority of the parameters in the estimated models have a significance level of 0.01 or 0.05, whereas the results of the pseudo R^2 , Likelihood Ratio (LR), LRT and the Bayesian Information Criterion (BIC) indicate a satisfactory adjustment of models to the empirical data. Selected values of the odds ratio, calculated based on the assessment of parameters of individual models have been interpreted.

Table 1. Estimation results of the logit model describing the inclination of foreign divestment according to three divestment scenarios in comparison to the one year before the outbreak of the COVID-19 pandemic

Variables	Divestment					
	up to 20%		from 20% to 40%		above 40%	
	coefficient	odds ratio	coefficient	odds ratio	coefficient	odds ratio
<i>Constant</i>	-12.15***	0.00	-9.12***	0.00	-7.02***	0.00
ΔGDP_c	0.12**	1.13	0.15*	1.16	-0.10***	0.91
ΔGDP	0.30**	1.35	0.33**	1.39	0.06***	1.07
ΔEx	0.21**	1.24	0.10***	1.11	-0.06*	0.94
ΔOMI	0.43**	1.53	0.62***	1.86	0.90***	2.45
ΔLC	0.37***	1.45	0.60***	1.83	0.46***	1.58
<i>M</i>	-0.19**	0.83	0.38***	1.46	0.85***	2.35
<i>S</i>	0.43**	1.53	0.22***	1.25	0.42**	1.52
<i>IT</i>	0.85***	2.33	0.42***	1.52	0.22*	1.24
<i>O</i>	-0.25***	0.78	-0.19*	0.83	-0.06***	0.94
Δi	0.02*	1.02	-0.22*	0.81	-0.17**	0.84
$\Delta R\&D$	0.08*	1.08	-0.28**	0.76	-0.68**	0.57
ΔDC	0.37***	1.45	0.22***	1.24	0.32*	1.37
<i>L1</i>	0.75***	2.13	-0.05*	0.96	-0.23*	0.79
<i>L2</i>	0.35*	1.43	0.64***	1.90	0.11**	1.12
<i>L3</i>	-0.27*	0.77	0.26**	1.29	0.87***	2.39
<i>IT</i> · <i>L1</i>	0.90**	2.45	.	.	.	.
<i>IT</i> · <i>L2</i>	.	.	0.76***	2.15	.	.
<i>IT</i> · <i>L3</i>	.	.	.	.	0.76***	2.14
<i>S</i> · <i>L1</i>	0.75**	2.11	.	.	.	.
ΔOMI · <i>L1</i>	.	.	0.63***	1.88	.	.
ΔOMI · <i>L3</i>	.	.	.	.	0.75***	2.12
ΔLC · <i>L2</i>	0.64**	1.90	.	.	.	.
ΔLC · <i>L3</i>	.	.	0.58***	1.79	0.58***	1.79
ΔGDP · <i>M</i>	0.17***	1.18	0.50***	1.65	0.47**	1.60
ΔOMI · <i>M</i>	0.55**	1.73	.	.	.	.
ΔGDP_c · <i>IT</i>	.	.	0.37***	1.45	.	.
ΔGDP_c · <i>M</i>	.	.	.	.	0.87***	2.39
<i>M</i> · <i>L2</i>	0.24**	1.27	.	.	.	.
<i>S</i> · <i>L3</i>	.	.	0.31**	1.37	.	.
$\Delta R\&D$ · <i>IT</i>	.	.	-0.35*	0.70	.	.
ΔO · Δi	.	.	.	.	-0.32*	0.73

Table 1. Estimation results of the logit model describing the inclination of foreign divestment according to three divestment scenarios in comparison to the one year before the outbreak of the COVID-19 pandemic (cont.)

Variables	Divestment					
	up to 20%		from 20% to 40%		above 40%	
	coefficient	odds ratio	coefficient	odds ratio	coefficient	odds ratio
McFadden's pseudo R ²	0.22		0.32		0.19	
LR	-166.83		-176.33		-160.29	
LRT	91.93***		162.17***		75.69***	
BIC	433.09		452.09		420.02	

Note. Significance at the level of: *** – 0.01, ** – 0.05, * – 0.1.

Source: own study based on the survey results.

Table 1 shows that if individual factors are taken into consideration, the greatest increase in divestment risk at a level of no more than 20% (in comparison to the pre-pandemic period) is generated by FDIs in IT and an optimistic course of the pandemic. The odds of divestment increase for the above reasons by approximately 133% and 113%, respectively. It should also be noted that investment in branches of the processing industry may reduce the risk of divestment by approximately 17%. The results incline the author to assume that in conditions of inconsiderable divestment the industrial sector is perceived to be safer than the services sector. Interactions in most cases increase considerably the risk of divestment in comparison to the combination with at least one factor creating an interaction. Thus, the combination of FDIs in the IT sector with the occurrence of weaker restrictions imposed on the economy due to the pandemic increases the risk of minor divestment by approximately 145%, whereas the interaction of FDIs in the services sector and an optimistic scenario of the development of the pandemic increases the risk of divestment by about 111%.

As regards medium divestment (from 20% to 40%) in comparison to the pre-pandemic period, an increase in labour costs and in the value of the market openness index (OMI) bears the greatest risk. An increase in labour costs by every subsequent 5 p.p. increases the risk of moderate divestment by about 83%, while an increase in the value of the OMI by every subsequent 5 p.p. increases the risk of divestment by approximately 86%. Moderate restrictions introduced to counteract the pandemic pose an even higher risk of moderate divestment (risk increases by around 90%); the same applies to the interaction of this factor with FDIs in the sector of IT services (risk increases by about 115%). Individual factors which increase the risk level of

FDI reduction when the prospect of divestment changes from optimistic (up to 20%) to moderately pessimistic (from 20% to 40%) include: a decrease in the GDP of the investor's country, a decrease in the GDP of Poland, a reduced OMI, increased labour costs, FDIs in the processing industry, FDIs in sectors of the economy other than the services and processing industries, the optimistic variant of pandemic development, and the moderate or pessimistic variant for pandemic development.

Table 1 suggests that the greatest risk of divestment at a level above 40% in comparison to the pre-pandemic period is generated by considerable economic restrictions imposed in response to the development of a pessimistic pandemic scenario, an increase in the OMI in Poland and FDIs in the processing industry. The pessimistic scenario of the development of the pandemic increases the risk of considerable divestment by approximately 139%. An increase in the value of the OMI by another 5 p.p. increases this risk by about 145% on average, while FDIs in the processing industry increase the risk of these FDIs by around 135% on average. With regard to factor interactions, the combination of FDIs in the processing industry with a decreased GDP in the investor's country by another 5 p.p. is characterised by the greatest risk of divestment at a level exceeding 40% (risk increases by approximately 140%). More than a twofold increase in the risk of considerable divestment is also generated by a combination of FDIs in the IT sector with a pessimistic scenario of pandemic development and a combination of an increase in the OMI in Poland with severe restrictions on the economy related to the development of the pandemic. Individual factors that have increased the risk of FDI reduction in the case of a change of the divestment perspective from moderately pessimistic (from 20% to 40%) to pessimistic (above 40%) include: an increased OMI, FDIs in the processing industry, FDIs in the services sector, FDIs in economic sectors other than the services and processing industries, an increase in inflation, the geographic distance between Poland and the investor's country, and the occurrence of a pessimistic scenario of pandemic development.

FDI in Poland mainly occurs in the processing industry and the services sector (Narodowy Bank Polski, n.d.). These sectors are characterised by specific investments and a distinct level of investment risk of different factors, so an examination of the inclination for divestment in these separate economic sectors seems definitely justified. Table 2 presents the results of the logit model estimations used to describe the inclination for FDI reduction in the processing industry.

Table 2. Estimation results of the logit model describing the inclination of foreign divestment in the processing industry according to three divestment scenarios in comparison to the one year before the outbreak of the COVID-19 pandemic

Variables	Divestment					
	up to 20%		from 20% to 40%		above 40%	
	coefficient	odds ratio	coefficient	odds ratio	coefficient	odds ratio
<i>Constant</i>	-6.54***	0.00	-8.01***	0.00	-13.10***	0.00
ΔGDP_c	0.21**	1.23	0.46**	1.58	0.32***	1.38
ΔGDP	0.32***	1.38	0.65***	1.92	0.64***	1.90
ΔEx	0.06**	1.06	0.20***	1.22	0.14**	1.15
ΔOMI	0.35***	1.43	0.62**	1.87	0.51***	1.67
ΔLC	0.16**	1.18	0.78***	2.19	0.67***	1.96
Δi	-0.03*	0.97	-0.14**	0.87	-0.21**	0.81
$\Delta R\&D$	0.08***	1.09	0.18***	1.20	0.04*	1.04
ΔDC	0.13***	1.14	0.22***	1.24	0.36***	1.43
$L1$	0.51***	1.66	0.32*	1.37	-0.11*	0.90
$L2$	0.19**	1.21	1.14***	3.11	0.72***	2.06
$L3$	0.12***	1.13	0.80***	2.23	1.30***	3.68
$\Delta GDP_c \cdot L1$	0.63*	1.87	.	.	.	.
$\Delta GDP_c \cdot L2$	.	.	0.80***	2.23	.	.
$\Delta GDP \cdot L3$	.	.	.	.	1.64***	5.17
$\Delta GDP \cdot L1$	0.81***	2.24	.	.	.	.
$\Delta GDP \cdot L2$	.	.	0.94***	2.55	.	.
$\Delta OMI \cdot L2$	.	.	.	.	1.72***	5.56
$\Delta OMI \cdot L1$	0.91***	2.49	.	.	.	.
$\Delta LC \cdot L3$	.	.	1.10***	3.01	.	.
$\Delta LC \cdot L2$	.	.	.	.	1.11***	3.03
$\Delta GDP \cdot \Delta OMI$	0.72**	2.06	1.00**	2.71	.	.
$\Delta OMI \cdot \Delta LC$	.	.	.	.	0.91***	2.49
$\Delta GDP \cdot \Delta Ex$	.	.	.	.	1.10***	3.01
McFadden's pseudo R^2	0.42		0.36		0.31	
LR	-181.43		-162.94		-199.07	
LRT	259.54***		186.51***		177.21***	
BIC	462.29		425.31		497.58	

Note. Significance at the level of: *** – 0.01, ** – 0.05, * – 0.1.

Source: own study based on the survey results.

The analysis of the results provided in Table 2 lead to the conclusion that, with reference to individual macroeconomic factors, the greatest increase in divestment risk in the processing industry at a level up to 20% in comparison to the pre-pandemic period is generated by a decrease in Poland's GDP, an increase in the OMI and the occurrence of an optimistic variant of the pandemic. With a decrease in GDP and an increase in OMI by every subsequent 5 p.p., the odds of divestment increase by approximately 38% and 43%, respectively, whereas minor restrictions on the economy imposed to counteract the pandemic increase the risk of divestment by about 66%. In general, interactions of variables increase the risk of divestment in the processing industry by twofold in comparison to the individual factors. For example, the

combination of a decrease in the GDP of Poland (by every 5 p.p.) with minor economic restrictions related to the pandemic increases the risk of minor divestment in the industry by 124%, whereas an interaction of OMI increases (by every 5 p.p.) and the same economic restrictions lead to an increase in the risk of divestment by approximately 149% on average.

In relation to moderate divestment in the industry (from 20% to 40%) in comparison to the period before the pandemic outbreak, economic recession in Poland, an increase in labour costs and a medium variant of the pandemic development scenario create the greatest risk of divestment. A decrease in Poland's GDP by every subsequent 5 p.p. increases the risk of divestment by approximately 92%. An increase in labour costs by every subsequent 5 p.p. increases the risk of moderate divestment by about 119%. A medium variant of the pandemic development increases the odds of divestment by around 211%. As regards the statistically significant interaction of the observed factors, the following increase the risk of moderate divestment to the greatest extent: an increase in the labour costs (by every 5 p.p.) combined with the pessimistic variant of pandemic development and a decrease in Poland's GDP combined with an increase in the value of the OMI (by every subsequent 5 p.p.). In the former case, an increase in the odds of moderate divestment amounts to 201%, while in the latter case, it equals 171%. It should be stressed that all individual factors increase the level of the FDI reduction risk along with a change in the divestment prospect from optimistic (up to 20%) to moderately pessimistic (from 20% to 40%), except for an increase in inflation and the occurrence of an optimistic variant of pandemic development.

Table 2 indicates that in comparison to the pre-pandemic period, the greatest risk of considerable divestment (at a level above 40%) in the processing industry is generated by the imposing of considerable and moderate economic restrictions in response to the pessimistic development of the pandemic scenario, a greater economic recession in Poland, and increased labour costs. A decrease in Poland's GDP by every subsequent 5 p.p. generates an increase in moderate divestment by approximately 90%, and an increase in labour costs by every subsequent 5 p.p. increases the risk of divestment by about 96% on average. Moderate economic restrictions introduced to counteract the pandemic increase the risk of divestment by about 106%, whereas the pessimistic variant of the development of the pandemic increases the risk of divestment by an average of 268%.

With reference to factor interactions, the combination of a decrease in Poland's GDP by every subsequent 5 p.p. with the pessimistic variant of pandemic development in Poland, and the interaction of an increase in the OMI by every subsequent

5 p.p. with moderate economic restrictions to counteract the pandemic are characterised by the greatest risk of foreign divestment at a level above 40% in the processing industry. In the former case, the risk of moderate divestment increases by approximately 417%, and in the latter case – by about 456%.

It should also be stressed that all individual factors have increased the risk of FDI reduction along with a change of the divestment prospect from optimistic (up to 20%) to pessimistic (above 40%), with the exception of an increase in inflation, decrease in R&D expenditures and the occurrence of an optimistic variant of pandemic development.

Table 3 presents the results of the estimation of logit models used to describe the inclination for FDI reduction in the services sector.

Table 3. Estimation results of the logit model describing the inclination of foreign divestment in the services sector according to three divestment scenarios in comparison to the one year before the COVID-19 pandemic outbreak

Variables	Divestment					
	up to 20%		from 20% to 40%		above 40%	
	coefficient	odds ratio	coefficient	odds ratio	coefficient	odds ratio
<i>Constant</i>	-5.14***	0.00	-13.14***	0.00	-6.53***	0.00
ΔGDP_c	0.18***	1.19	0.20***	1.22	0.95***	2.59
ΔGDP	0.33***	1.39	0.49***	1.63	0.34*	1.41
ΔEx	0.08*	1.08	-0.15*	0.86	0.91***	2.47
ΔOMI	0.49***	1.63	0.82***	2.28	0.77***	2.15
ΔLC	0.38***	1.47	0.69***	2.00	0.52***	1.68
Δi	0.17*	1.19	-0.12**	0.89	0.43***	1.54
$\Delta R\&D$	-0.65***	0.52	0.40***	1.49	0.49***	1.63
ΔDC	0.01*	1.01	0.62***	1.85	-0.01*	0.99
$L1$	0.56***	1.75	0.26***	1.30	0.48**	1.61
$L2$	0.01	1.01	0.82***	2.26	1.48***	4.40
$L3$	-0.20*	0.82	0.41***	1.51	1.84***	6.27
$\Delta GDP_c \cdot L1$	0.42***	1.53	.	.	.	.
$\Delta GDP_c \cdot L2$	.	.	1.41***	4.10	.	.
$\Delta GDP_c \cdot L3$	.	.	.	.	1.94***	6.97
$\Delta GDP \cdot \Delta OMI$	1.38***	3.98	.	.	0.81***	2.24
$\Delta OMI \cdot \Delta DC$	0.78***	2.18	1.52***	4.58	1.09***	2.98
$\Delta OMI \cdot L1$	1.49***	4.44	.	.	.	.
$\Delta OMI \cdot L3$	.	.	1.63***	5.09	.	.
$\Delta OMI \cdot \Delta LC$	.	.	.	.	1.13***	3.10
$\Delta GDP \cdot \Delta R\&D$	.	.	1.38***	3.98	0.51***	1.67
McFadden's pseudo R^2	0.25		0.34		0.21	
LR	-159.08		-210.53		-184.26	
LRT	108.33***		218.85***		96.78***	
BIC	417.59		520.49		467.95	

Note. Significance at the level of: *** – 0.01, ** – 0.5, * – 0.1.

Source: own study based on the survey results.

The greatest risk of minor divestment (up to 20%) in the services sector is generated by the occurrence of an optimistic scenario of pandemic development, an increase in the OMI in Poland and increased labour costs (see Table 3).

The occurrence of an optimistic variant of pandemic development increases the risk of minor divestment in the services sector by approximately 75%. An increase in the OMI by every subsequent 5 p.p. increases the odds of divestment by an average of 63%, whereas an increase in the labour costs of the same scale increases the odds of foreign divestment by about 47% on average.

As regards the interaction of factors considerably increasing the risk of minor divestment in the services sector, they include the combination of an economic recession in Poland with an increased OMI (increase in divestment risk by approximately 298%), and the combination of an increased OMI with an optimistic scenario of pandemic development (increase in the odds of divestment in the services sector by about 344%).

Factors which increase the odds of moderately pessimistic divestment the most in the services sector (from 20% to 40%) in comparison to the pre-pandemic period include an economic recession in Poland, an increased OMI and moderate restrictions imposed on the economy to counteract the pandemic. These increase the risk of divestment in the services sector by 63% (when the GDP falls by every subsequent 5 p.p.), 128% (with an increase in OMI by every subsequent 5 p.p.) and 126%, respectively. The interaction of variables taken into account in the model provided in Table 3 has increased the risk level of moderately pessimistic foreign divestment by at least two and a half times. For example, the combination of an increase in OMI in Poland (by every subsequent 5 p.p.) with an increase in the geographic distance from the investor's country increases the risk of divestment in the services sector by 358% on average, whereas the interaction of an increase in OMI (by every subsequent 5 p.p.) and the occurrence of severe restrictions on the economy leads to a risk of divestment increased by an average of about 409%.

It should be emphasised that all individual factors have increased the risk level of the reduction of FDI in the services sector along with a change of the divestment perspective from an optimistic one (up to 20%) to a moderately pessimistic one (from 20% to 40%), with the exception of a change in the exchange rate of the PLN, an increase in inflation and the occurrence of the optimistic variant of pandemic development.

Based on Table 3, it can be concluded that the greatest risk of considerable divestment (at a level above 40%) in the services sector in relation to the pre-pandemic

period is generated by the introduction of considerable and moderate restrictions on the economy in response to the pessimistic development of the pandemic scenario, increased economic recession in the investor's country, and the depreciation of the Polish currency. A decrease in the GDP of the investor's country by every subsequent 5 p.p. generates an increase in the risk of considerable divestment by approximately 159% on average, whereas a decrease in the exchange rate of PLN in relation to the euro by every subsequent 5 p.p. increases the risk of divestment by an average of about 147%. Moderate restrictions on the economy imposed to counteract the pandemic increase the risk of divestment by around 340%, while the pessimistic variant of pandemic development increases the risk of divestment by approximately 527% on average.

Regarding the interactions of factors, the combination of a decrease in the GDP of an investor's country by every subsequent 5 p.p. with a pessimistic scenario of pandemic development in Poland, and the interaction between an increase in the OMI by every subsequent 5 p.p. and a relative increase in labour costs by a similar value bear the greatest risk of considerable foreign divestment in the services sector at a level above 40%. In the former case, the risk of moderate divestment increases by approximately 597% and in the latter case – by about 210%. It should be emphasised that all individual factors have increased the FDI reduction risk level in the services sector along with a change in the divestment prospect from optimistic (up to 20%) to pessimistic (above 40%), with the exception of geographic distance from the investing country to Poland and the occurrence of the optimistic variant of pandemic development.

4. Conclusions

The COVID-19 pandemic is a new phenomenon in the modern world which threatens the stability of many sectors of the world economy and the economies of individual countries. The full economic consequences of the pandemic have not yet been entirely investigated due to the short time horizon of this phenomenon occurring. The paper attempts to fill this research gap with regard to the impact of the pandemic on FDI. In the present study, it was possible to identify the factors which increased the risk of foreign divestment in Poland, so the aim of the paper has been achieved. The application of the logit model allowed the identification of the most important factors of foreign divestment in the time of the COVID-19 pandemic. This tool proved successful in estimating the risk of divestment due to the specificity of the variables (measured on a nominal or ordinal scale) and the ease of estimating the

model and interpreting the results presented in the form of odds ratios. The obtained research results lead to the conclusion that the most important determinants of divestment in the economy as a whole are: the location of FDI in the services sector, in industry and in the IT sector, increased market openness and interactions of variables taking into account the restrictions introduced in the economy in connection with the state of the pandemic. The level of risk of divestment for these variables depends, however, on the size of the FDI reduction declared by investors and on the sector of the economy.

This is particularly visible if considerable divestment is assumed: FDIs in the services sector are then burdened with a higher risk of divestment than FDIs in the processing industry. If average divestment at a low level is assumed, FDIs in the IT sector constitute a factor bearing the greatest risk of FDI reduction in the entire economy. For the expected medium-level divestment and for the divestment level above 40%, the OMI is the FDI reduction factor bearing the greatest risk.

Economic restrictions aimed at limiting the transmission of COVID-19 and their interactions, especially interactions with the economic recession in Poland and in the investor's country and with the OMI are the factors of foreign divestment bearing the greatest risk in the processing industry and in the services sector. With the introduction of factor interactions to the model, it was possible to detect the feedback between independent variables. The interaction of variables, especially economic restrictions introduced to counteract the pandemic taken into account in logit models, in many cases, increases the risk of foreign divestment by multiple times. During the COVID-19 pandemic, the estimated logit models may make it easier for investors to make decisions to divest or facilitate the choice of new directions of investment. The changing intensity and development of the COVID-19 pandemic, which is difficult to predict, justify the need to constantly update the results of such research and to repeat it in the future, which will allow for an ongoing monitoring of the degree of investment risk of various macroeconomic and microeconomic factors. It cannot be excluded that, considering the further development of the COVID-19 pandemic, the priorities of investors regarding the scale and direction of investments will further change and divestment will turn out to be unavoidable in certain industry sectors.

Acknowledgements

The publication was financed from a subvention allocated to Cracow University of Economics.

References

- Aidukaite, J., Saxonberg, S., Szelewa, D., & Szikra, D. (2021). Social policy in the face of a global pandemic: Policy responses to the COVID-19 crisis in Central and Eastern Europe. *Social Policy & Administration*, 55(2), 358–373. <https://doi.org/10.1111/spol.12704>.
- Bergh, D. D. (1997). Predicting Divestiture of Unrelated Acquisitions: An Integrative Model of Ex Ante Conditions. *Strategic Management Journal*, 18(9), 715–731. [https://doi.org/10.1002/\(SICI\)1097-0266\(199710\)18:9<715::AID-SMJ912>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0266(199710)18:9<715::AID-SMJ912>3.0.CO;2-6).
- Bergsen, P. (2020). A new political economy for Europe post-COVID-19. *European View*, 19(2), 131–137. <https://doi.org/10.1177/1781685820968301>.
- Berry, H. (2010). Why Do Firms Divest?. *Organization Science*, 21(2), 380–396. <https://doi.org/10.1287/orsc.1090.0444>.
- Berry, H. (2013). When Do Firms Divest Foreign Operations?. *Organization Science*, 24(1), 246–261. <https://doi.org/10.1287/orsc.1110.0724>.
- Blake, D. J., & Moschieri, C. (2017). Policy risk, strategic decisions and contagion effects: Firm-specific considerations. *Strategic Management Journal*, 38(3), 732–750. <https://doi.org/10.1002/smj.2509>.
- Borga, M., Flores, P. I., & Sztajerowska, M. (2019). *Drivers of divestment decisions of multinational enterprises – A cross-country firm-level perspective* (OECD Working Papers on International Investment No. 2019/3). <https://doi.org/10.1787/5a376df4-en>.
- Bureau van Dijk. (n.d. a). *Orbis*. <https://www.bvdinfo.com/en-gb/our-products/data/international/orbis>.
- Bureau van Dijk. (n.d. b). *Zephyr*. <https://www.bvdinfo.com/en-gb/our-products/data/economic-and-ma/zephyr>.
- Chatterjee, S., Harrison, J. S., & Bergh, D. D. (2003). Failed takeover attempts, corporate governance and refocusing. *Strategic Management Journal*, 24(1), 87–96. <https://doi.org/10.1002/smj.279>.
- Davahli, M. R., Karwowski, W., Sonmez, S., & Apostolopoulos, Y. (2020). The Hospitality Industry in the Face of the COVID-19 Pandemic: Current Topics and Research Methods. *International Journal of Environmental Research and Public Health*, 17(20), 1–20. <https://doi.org/10.3390/ijerph17207366>.
- Doctor, A. C., & Bagwell, S. (2020). Risky Business: Foreign Direct Investment and the Economic Consequences of Electoral Violence. *Journal of Global Security Studies*, 5(2), 339–360. <https://doi.org/10.1093/jogss/ogz070>.
- Hamilton, R. T., & Chow, Y. K. (1993). Why managers divest—Evidence from New Zealand's largest companies. *Strategic Management Journal*, 14(6), 479–484. <https://doi.org/10.1002/smj.4250140606>.
- Harrell, F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer-Verlag. <https://doi.org/10.1007/978-1-4757-3462-1>.
- Harrigan, K. R. (1981). Deterrents to Divestiture. *The Academy of Management Journal*, 24(2), 306–323. <https://doi.org/10.5465/255843>.

- Jaccard, J. (2001). *Interaction Effects in Logistic Regression*. Thousand Oaks: Sage Publications. <https://dx.doi.org/10.4135/9781412984515>.
- Jovanovic, B., & MacDonald, G. M. (1994). The Life Cycle of a Competitive Industry. *Journal of Political Economy*, 102(2), 322–347. <https://doi.org/10.1086/261934>.
- Markides, C. C. (1992). Consequences of corporate refocusing: ex ante evidence. *Academy of Management Journal*, 35(2), 398–412. <https://doi.org/10.5465/256379>.
- Martins, P. S., & Esteves, L. A. (2008). *Foreign Ownership, Employment and Wages in Brazil: Evidence from Acquisitions, Divestments and Job Movers* (IZA Discussion Papers No. 3542). <http://ftp.iza.org/dp3542.pdf>.
- Narodowy Bank Polski. (n.d.). *Cykliczne materiały analityczne NBP*. <https://www.nbp.pl/home.aspx?f=/publikacje/zib/zib.html>.
- Norbäck, P.-J., Tekin-Koru, A., & Waldkirch, A. (2015). Multinational Firms and Plant Divestiture. *Review of International Economics*, 23(5), 811–845. <https://doi.org/10.1111/roie.12199>.
- Pashley, M. M., & Philippatos, G. C. (1990). Voluntary divestitures and corporate life-cycle: some empirical evidence. *Applied Economics*, 22(9), 1181–1196. <https://doi.org/10.1080/00036849000000038>.
- Shin, S. (2000). The Foreign Divestment Factors in South Korea: An Analysis of the Trading Sector. *The Multinational Business Review*, 8(2), 98.
- Trencher, G., Downie, Ch., Hasegawa, K., & Asuka, J. (2020). Divestment trends in Japan's international coal businesses. *Renewable and Sustainable Energy Reviews*, 124. <https://doi.org/10.1016/j.rser.2020.109779>.

Zastosowanie rozkładu Pareta do aproksymacji rozkładów dochodów zamożnych gospodarstw domowych

Dorota Pekasiewicz^a

Streszczenie. Celem badania omawianego w artykule jest aproksymacja rozkładów ekwiwalentnych dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych za pomocą rozkładu Pareta z parametrami oszacowanymi metodą największej wiarygodności. Za zamożne gospodarstwa domowe uznano te, których dochody przekroczyły ustalony próg bogactwa (zamożności). Rozkłady dochodów zamożnych gospodarstw domowych są zazwyczaj zeromodalne i cechują się grubymi ogonami, dlatego do ich aproksymacji wykorzystano rozkład Pareta. Analizowano ekwiwalentne dochody zamożnych gospodarstw w Polsce ogółem i w poszczególnych grupach społeczno-ekonomicznych w latach 2014–2017. Dane uzyskano z badania budżetów gospodarstw domowych. Stopień dopasowania rozkładów teoretycznych do empirycznych zbadano za pomocą wybranych miar podobieństwa.

Wyniki potwierdziły dużą zgodność empirycznych rozkładów dochodów z modelem Pareta. Bardzo dobre aproksymacje uzyskano w szczególności dla zamożnych gospodarstw domowych pracowników, pracujących na własny rachunek oraz emerytów i rencistów. Nieznacznie gorsze wyniki otrzymano dla gospodarstw rolników. Rozkłady teoretyczne dobrze dopasowane do danych empirycznych zostały wykorzystane do szacowania wybranych charakterystyk rozkładu, m.in. miar położenia, zróżnicowania i nierównomierności, oraz do porównywania grup pod względem ich zamożności.

Słowa kluczowe: dochód, gospodarstwo domowe, rozkład Pareta, zamożność

JEL: C10, J30

The application of the Pareto distribution to approximate income distributions of wealthy households in Poland

Abstract. The aim of the paper is to approximate the equivalent income distributions of wealthy households in particular socio-economic groups using the Pareto distribution, with parameters estimated by means of the maximum likelihood estimation method. Households whose income exceeded the established wealth threshold were classified as wealthy households. Income distributions of wealthy households are usually non-modal and heavy-tailed, thus, the Pareto distribution was applied as their theoretical model. The equivalent income of wealthy households in Poland was analysed in total and in particular socio-economic groups. The research was based on data from the 2014–2017 Household Budget Survey. Selected similarity measures were used to examine the degree to which the theoretical distributions proved consistent with the empirical ones.

The obtained results confirmed the high level of consistency of empirical income distributions with the Pareto model. Moreover, very good approximations were obtained especially for

^a Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Polska / University of Lodz, Faculty of Economics and Sociology, Poland. ORCID: <https://orcid.org/0000-0001-8275-3345>.
E-mail: dorota.pekasiewicz@uni.lodz.pl.

wealthy households of employees and self-employed, as well as pensioners. Slightly worse results were obtained for the farmers group. Theoretical distributions well fitted to empirical data were used to estimate selected distribution characteristics, including measures of location, dispersion and inequality, and to compare the different groups in terms of their wealth.

Keywords: income, household, Pareto distribution, wealth

1. Wprowadzenie

Rozkłady dochodów, w tym ich nierównomierność, są przedmiotem badań zarówno statystyków, jak i ekonomistów. Badania te obejmują takie nurty, jak poszukiwanie modelu teoretycznego opisującego rozkłady empiryczne, tworzenie teorii wyjaśniających proces kształtowania się rozkładów dochodów i ich nierównomierności oraz doskonalenie metod badania nierównomierności rozkładów.

Dotychczasowe badania zaowocowały dużą liczbą rozkładów teoretycznych, które znajdują zastosowanie w analizie dochodów. W pracach Jędrzejczak (2006), Jędrzejczak i Pekasiewicz (2020), Kleibera (1996), Kleibera i Kotza (2003), Ostasiewicz (2013) i Salamagi (2016) można znaleźć wyniki analiz wykorzystujących wiele rozkładów, w tym Daguma, Singha-Maddali czy Fiska, do aproksymacji rozkładów dochodów, ze wskazaniem tych, które najlepiej odwzorowują rozkłady empiryczne dla różnych grup gospodarstw domowych. Rozkłady dochodów gospodarstw najczęściej są jednomodalne i charakteryzują się prawostronną asymetrią oraz dodatnią kurtozą (Kot, 1999).

Inny typ rozkładu cechuje gospodarstwa domowe o wysokich dochodach. Są one zazwyczaj zeromodalne i mają nieskończone wariancje. Już wyniki badania przedstawione w pracy Mandelbrota (1963) pokazały, że w przypadku grup o wysokich dochodach dobre dopasowanie do danych empirycznych wykazuje rozkład Pareta, dlatego jest on wykorzystywany do aproksymacji miesięcznych dochodów ekwiwalentnych gospodarstw domowych w grupach społeczno-ekonomicznych. Problematykę zamożności i bogactwa oraz modelowania rozkładów indywidualnych i ekwiwalentnych dochodów zamożnych gospodarstw domowych za pomocą rozkładu Pareta rozważają m.in. Brzeziński (2014), Brzeziński i in. (2020), Charpentier i Flachaire (2019), Jagielski i Kutner (2011), Peichl i in. (2008), Sączewska-Piotrowska (2018) oraz Włodarczyk (2013).

W analizach dotyczących dochodów przypadających na osobę i porównywania dochodów w gospodarstwach domowych istotnym zagadnieniem jest uwzględnienie struktury demograficznej gospodarstwa. Koszty utrzymania nie są proporcjonalne do liczby członków gospodarstwa, dlatego wielkość dochodów ekwiwalentnych wyznacza się z zastosowaniem wybranej skali ekwiwalentności. W analizach przeprowadzonych w ramach badania omawianego w niniejszym artykule wykorzystano skalę potęgową z elastycznością ekwiwalentności równą 0,5.

Zagadnienia poruszane w artykule dotyczą aproksymacji empirycznych rozkładów dochodów ekwiwalentnych zamożnych gospodarstw domowych za pomocą rozkładu Pareta. Wiadomo, że o zamożności gospodarstw decydują nie tylko dochody. Duże znaczenie ma również wielkość dochodów niewydatkowanych na konsumpcję i kumulowanych w postaci oszczędności rzeczowych oraz finansowych, generujących posiadanie aktywów. Niemniej wielkość dochodu jest jednym z najważniejszych czynników wpływających na sytuację finansową gospodarstw. Za gospodarstwo domowe o wysokich dochodach uznaje się to, którego dochody przekraczają ustalony próg (granice) bogactwa, zwany też granicą zamożności.

Wartość granicy zamożności jest ustalana różnie, m.in. na poziomie dwu- lub trzykrotności mediany rozkładu dochodów ekwiwalentnych albo wybranego kwantyla rozkładu dochodów (zob. Brzeziński, 2014; Peichl i in., 2008). W literaturze czasem odróżnia się zamożnych od bogatych i bardzo bogatych. Według Franziniego i in. (2016) gospodarstwa domowe osiągające dochody powyżej trzykrotności mediany traktuje się jako zamożne, osiągające dochody pięciokrotnie przewyższające medianę – jako bogate, a osiągające dochody przynajmniej 10 razy wyższe od mediany – jako bardzo bogate. W omawianym badaniu za zamożne uznano gospodarstwa o miesięcznych dochodach ekwiwalentnych przekraczających wartość dwukrotności mediany i aproksymowano rozkłady ich dochodów za pomocą rozkładu Pareta o parametrach szacowanych metodą największej wiarygodności.

Celem badania było dopasowanie rozkładu teoretycznego do rozkładów ekwiwalentnych dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych. Rozkłady Pareta aproksymujące empiryczne rozkłady dochodów zamożnych gospodarstw wyznaczono dla grup społeczno-ekonomicznych w latach 2014–2017. Rozkłady teoretyczne umożliwiają dokładniejszą analizę oraz opracowanie prognoz dotyczących rozkładów dochodów. W omawianym badaniu zostały wykorzystane do oszacowania miar położenia, zróżnicowania oraz nierównomierności ekwiwalentnych dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych w 2017 r.

2. Rozkład Pareta i jego wybrane charakterystyki

Wśród rozkładów, które opracowano w wyniku obserwacji rozkładów dochodów ludności uzyskanych ze statystyk podatkowych różnych krajów europejskich, znajduje się rozkład Pareta (zob. Pareto, 1897).

Funkcja gęstości rozkładu Pareta jest określona wzorem

$$f(y) = \begin{cases} \frac{a\theta^a}{y^{a+1}} & \text{dla } y \geq \theta, \\ 0 & \text{dla } y < \theta, \end{cases} \quad (1)$$

a dystrybuanta – wzorem

$$F(y) = \begin{cases} 1 - \left(\frac{y}{\theta}\right)^{-a} & \text{dla } y \geq \theta, \\ 0 & \text{dla } y < \theta, \end{cases} \quad (2)$$

gdzie:

θ – minimalny dochód,

a – parametr zwany stałą Pareta.

Momenty zwykłe rzędu r istnieją, gdy parametr kształtu a spełnia nierówność $a > r$. Wówczas przyjmują one postać (zob. np. Krishnamoorthy, 2006)

$$E(Y^r) = \theta^r \frac{a}{a - r} \quad \text{dla } a > r. \quad (3)$$

Wartość oczekiwana wyraża się wzorem

$$E(Y) = \theta \frac{a}{a - 1} \quad \text{dla } a > 1. \quad (4)$$

Rozkład dochodów określony krzywą Pareta ma skończoną wariancję tylko dla $a > 2$ i przyjmuje postać

$$D(Y) = \frac{\theta^2}{(a - 1)^2} \cdot \frac{a}{a - 2} \quad \text{dla } a > 2. \quad (5)$$

Rozkład Pareta może charakteryzować się niewielką liczbą skończonych momentów, co oznacza, że może być rozkładem z ciężkim ogonem (*heavy tail*). Im mniejsza liczba skończonych momentów rozkładu, tym grubszy jest jego prawy koniec.

Dobre dopasowanie rozkładu Pareta do rozkładu dochodów może być wykorzystane do oszacowania funkcji Lorenza oraz miar określających stopień nierównomierności rozkładu, takich jak współczynnik Giniego czy współczynnik Schutza-Pietry.

Funkcja Lorenza dla rozkładu Pareta wyraża się wzorem

$$L(p) = 1 - (1 - p)^{1 - \frac{1}{a}} \quad \text{dla } a > 1, \quad (6)$$

gdzie $p \in [0, 1]$.

Współczynnik Giniego i współczynnik Schutza-Pietry przyjmują postać odpowiednio:

$$G = \frac{1}{2a-1} \quad \text{dla } a > 1, \quad (7)$$

$$SP = \frac{1}{a-1} \left(\frac{a-1}{a} \right)^a \quad \text{dla } a > 1. \quad (8)$$

Oszacowanie $\hat{a}$ parametru a pozwala zatem wyznaczyć zarówno funkcję Lorenza $\hat{L}(p) = 1 - (1-p)^{1-\frac{1}{\hat{a}}}$ dla $p \in [0,1]$, jak i współczynniki nierówności: $\hat{G} = \frac{1}{2\hat{a}-1}$ oraz $\hat{SP} = \frac{1}{\hat{a}-1} \left(\frac{\hat{a}-1}{\hat{a}} \right)^{\hat{a}}$ dla $a > 1$.

Im większa jest wartość parametru a , tym wartości współczynników Giniego i Pietry są mniejsze, a tym samym mniejsze są nierównomierności rozkładu.

3. Metoda badania

Podstawowym źródłem informacji o dochodach, wydatkach, spożyciu oraz innych aspektach warunków życia określonych grup ludności są badania budżetów gospodarstw domowych (BBGD), prowadzone przez Główny Urząd Statystyczny (Główny Urząd Statystyczny [GUS], 2018b). Ich wyniki pozwalają na analizę wpływu różnych czynników na kształtowanie się poziomu i zróżnicowania dochodów gospodarstw pogrupowanych według określonych kryteriów, m.in. źródła utrzymania, klasy miejscowości zamieszkania albo typu biologicznego gospodarstwa, województwa czy regionu. Ze względu na źródła utrzymania wyróżnia się następujące grupy społeczno-ekonomiczne gospodarstw:

- pracowników;
- rolników;
- pracujących na własny rachunek;
- emerytów i rencistów;
- utrzymujących się ze źródeł niezarobkowych.

W omawianym badaniu wymienione grupy gospodarstw analizowano pod kątem ich zamożności.

Badanie budżetów gospodarstw domowych prowadzone jest metodą reprezentacyjną, która daje możliwość uogólnienia (z określonym błędem) uzyskanych wyników na wszystkie gospodarstwa domowe w Polsce lub na wybrane grupy gospodarstw. Jednostką obserwacji jest gospodarstwo domowe jedno- lub wieloosobowe. Za gospodarstwo zamożne uznaje się gospodarstwo o dochodach przekraczających próg bogactwa (zamożności). Na potrzeby pracy badawczej opisanej w artykule próg ten ustalono na poziomie dwukrotności mediany.

Do analiz wykorzystano dane za lata 2014–2017 pochodzące z BBDG (GUS, 2015, 2016, 2017, 2018a)¹. Dane te charakteryzują się jednak nadreprezentatywnością grup o niskich dochodach. Występuje tu także problem niedostatecznej reprezentacji najbogatszych gospodarstw domowych. Do zaniżania dochodów dochodzi nawet w przypadku danych uzyskiwanych z legalnych źródeł, co jest argumentem przemawiającym za analizowaniem wielkości wydatków, a nie dochodów w badaniach dotyczących dobrobytu (zob. Kot, 2003).

Przystępując do badania rozkładów dochodów ekwiwalentnych, należy podjąć decyzję o wyborze skali ekwiwalentności, która odzwierciedla wpływ struktury demograficznej gospodarstwa domowego na koszty utrzymania gospodarstw. W niniejszej pracy wykorzystano skalę potęgową – często stosowaną oprócz zmodyfikowanej i oryginalnej skali OECD – z elastycznością ekwiwalentności równą 0,5, czyli pierwiastek z liczby osób w gospodarstwie domowym.

Parametry rozkładu Pareta wykorzystywanego do aproksymacji rozkładów dochodów zamożnych gospodarstw domowych szacowano metodą największej wiarygodności (MNW). Metoda ta została wybrana na podstawie badań symulacyjnych, których rezultaty przedstawiono w pracy Jędrzejczak i Pekasiewicz (2020). Świadczą one o dobrych własnościach estymatorów otrzymanych za pomocą MNW w porównaniu z własnościami (obciążonością, średnim kwadratem błędu) estymatorów użytych innymi metodami, np. metodą momentów ważonych prawdopodobieństwami czy metodą najmniejszych kwadratów lub ważoną metodą najmniejszych kwadratów.

Dobroć, czyli trafność dopasowania rozkładu teoretycznego do empirycznego, oceniona została za pomocą współczynników zgodności: wskaźnika podobieństwa struktur (Wps), współczynnika Mortary² (IM) oraz współczynnika uwzględniającego różnicę między wartością średnią rozkładu empirycznego i wartością oczekiwaną rozkładu teoretycznego (I_r).

Do obliczenia wskaźnika podobieństwa struktur posłużono się wzorem (zob. np. Chomętowski i Sokołowski, 1978; Kruszka, 1989)

$$Wps = \sum_{j=1}^k \min(w_j, \hat{w}_j), \quad (9)$$

gdzie:

k – liczba przedziałów, na które zostały pogrupowane uporządkowane dane,

w_j – empiryczne wskaźniki struktury, $j = 1, \dots, k$,

$\hat{w}_j$ – teoretyczne wskaźniki struktury, $j = 1, \dots, k$.

¹ W trakcie pisania artykułu dane za 2017 r. były najnowszymi dostępnymi danymi.

² Giorgio Mortara (1885–1967) – włoski ekonomista, demograf i statystyk.

Współczynnik Mortary ma postać (zob. np. Zenga i in., 2010)

$$IM = \frac{1}{n} \sum_{j=1}^k |n_j - \hat{n}_j|, \quad (10)$$

gdzie:

k – liczba przedziałów,

n_j – częstości empiryczne,

$\hat{n}_j$ – częstości teoretyczne.

Im wskaźnik podobieństwa struktur jest bliższy 1, a współczynnik Mortary – bliższy 0, tym dopasowanie jest lepsze.

W omawianym badaniu liczba przedziałów, na które zostały podzielone dane z poszczególnych grup, służąca do obliczenia wskaźnika podobieństwa struktur i współczynnika Mortary, wynosiła od 15 do 20, w zależności od liczebności grup.

Zgodność rozkładu empirycznego z wybranym modelem teoretycznym bada się również poprzez porównanie najważniejszych charakterystyk liczbowych rozkładu teoretycznego z odpowiednimi charakterystykami empirycznymi – mogą to być średnie, odchylenia standardowe lub współczynniki nierównomierności. W przeprowadzonych analizach uwzględniono współczynnik będący względną różnicą między średnią wartością rozkładu empirycznego i oczekiwaną wartością rozkładu teoretycznego:

$$I_r = \frac{|\bar{Y} - EY|}{EY} \cdot 100\%, \quad (11)$$

gdzie:

$\bar{Y}$ – średnia arytmetyczna wartości dochodów $Y_1, Y_2, \dots, Y_n$,

EY – wartość oczekiwana teoretycznego rozkładu Pareta.

Przyjmuje się, że w przypadku bardzo dobrego dopasowania różnica między empirycznymi i teoretycznymi charakterystykami rozkładu powinna być mniejsza niż 5%.

4. Wyniki analiz rozkładów dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych

W badaniu analizowano rozkłady miesięcznych dochodów ekwiwalentnych zamożnych gospodarstw domowych w latach 2014–2017.

Na podstawie prób pochodzących z BBGD (GUS, 2015, 2016, 2017, 2018a) dla każdego analizowanego roku wyznaczono granicę zamożności (y_b^*) jako wartość dwukrotności mediany empirycznego rozkładu miesięcznych dochodów ekwiwalentnych. Otrzymano następujące wartości:

- 2014 r. – 4041,46 zł;
- 2015 r. – 4195,36 zł;
- 2016 r. – 4477,12 zł;
- 2017 r. – 4695,74 zł.

Zaobserwowano wzrost granicy zamożności z roku na rok, a średnie tempo zmian w analizowanym okresie wynosiło 5,13%. Udziały gospodarstw zamożnych w Polsce i w poszczególnych grupach społeczno-ekonomicznych, obliczone na podstawie wyznaczonych wartości y_b^* , przedstawiono w tabl. 1.

Tabl. 1. Odsetek gospodarstw o wysokich dochodach w grupach społeczno-ekonomicznych

L a t a	Gospodarstwa					
	ogółem	pracowników	rolników	pracujących na własny rachunek	emerytów i rencistów	utrzymujących się ze źródeł niezarobkowych
2014	8,67	11,05	16,32	19,65	2,71	1,82
2015	8,12	10,16	14,51	20,98	2,82	1,16
2016	7,34	9,33	15,57	19,49	2,07	1,76
2017	7,42	9,32	17,85	19,74	1,90	1,87

Źródło: obliczenia własne na podstawie: GUS (2015, 2016, 2017, 2018a).

W latach 2014–2017 gospodarstwa domowe osiągające miesięczne dochody ekwiwalentne powyżej dwukrotności mediany stanowiły niewiele ponad 7% wszystkich gospodarstw w Polsce, przy czym w analizowanym okresie zaobserwowano zmniejszenie się udziału gospodarstw o wysokich miesięcznych dochodach ekwiwalentnych. W grupie pracujących na własny rachunek i w grupie rolników odsetek gospodarstw domowych o dochodach ekwiwalentnych przekraczających granicę zamożności był najwyższy, a znacznie mniejszy – w grupach emerytów i rencistów oraz utrzymujących się ze źródeł niezarobkowych.

W tabl. 2 i na wyk. 1 przedstawiono charakterystyki empirycznych rozkładów miesięcznych dochodów ekwiwalentnych zamożnych gospodarstw domowych w 2017 r.

Miesięczne dochody ekwiwalentne zamożnych gospodarstw domowych w grupie rolników okazały się bardzo zróżnicowane. W wylosowanej próbie pojawiły się wartości odstające. Można je zidentyfikować na podstawie wykresu ramkowego, który

według Safariego i in. (2019) jest dobrą metodą wyznaczania nietypowych (odstających) wielkości dochodów, lub za pomocą metody kwartyłowej (zob. Chromiński i Tkacz, 2010). Za ekstremalnie nietypowe uznaje się wartości większe od $Q_3 + 3RQ$ lub mniejsze od $Q_1 - 3RQ$, gdzie Q_1 , Q_3 są odpowiednio kwartylem pierwszym i trzecim, a RQ oznacza rozstęp ćwiartkowy.

Tabl. 2. Charakterystyki empirycznych rozkładów dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych w 2017 r.

Gospodarstwa	Liczba gospodarstw	Minimum	Maksimum	Średnia	Mediana	Odchylenie standardowe	Współczynnik zmienności w %
		w zł					
Ogółem	2721	4696,10	264261,36	6965,53 (6870,94)	5773,50	6404,61 (4083,82)	91,95 (59,43)
Pracowników	1650	4696,10	66178,50	6444,46	5673,62	3031,41	47,12
Rolników	296	4715,67	264261,36	10692,10 (9832,57)	7479,66	16812,30 (7997,43)	157,24 (81,34)
Pracujących na własny rachunek	495	4698,94	57000,00	7048,34	6010,41	3854,53	54,69
Emerytów i rencistów	252	4702,26	22208,00	5857,17	5391,90	1584,69	27,06
Utrzymujących się ze źródeł niezarobkowych	28	4843,97	12625,50	6787,66	5859,72	2052,50	30,24

Uwaga. W nawiasach podano charakterystyki liczbowe po odrzuceniu wartości ekstremalnie odstającej, tzn. 264261,36 zł.

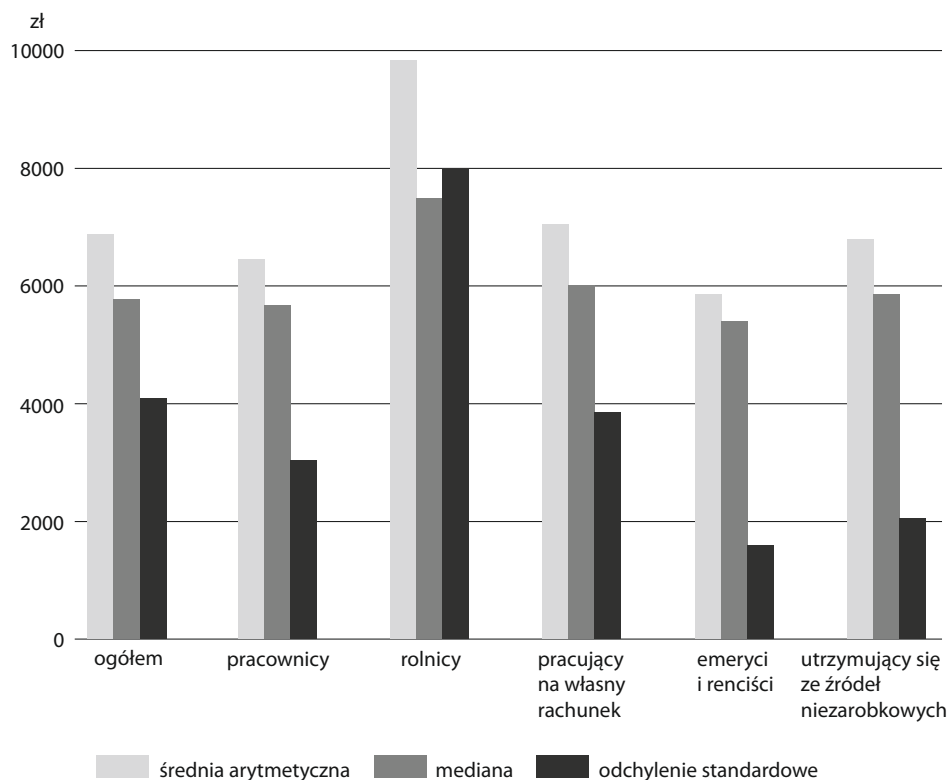
Źródło: obliczenia własne na podstawie: GUS (2018a).

W przypadku zamożnych gospodarstw rolników i wszystkich gospodarstw o wysokich dochodach ekstremalnie odstającą wartością jest miesięczny dochód ekwiwalentny w wysokości 264261,36 zł (dla grupy rolników otrzymano $Q_3 + 3RQ = 26056,977$ zł, a dla ogółu – $Q_3 + 3RQ = 22435,088$ zł), który został pominięty przy aproksymacji rozkładów miesięcznych dochodów zamożnych gospodarstw rolników i ogółu zamożnych gospodarstw domowych w Polsce. W przypadku rolników tak duża wartość miesięcznego dochodu może wynikać z sezonowości ich dochodów. W dalszych analizach ograniczono się do usunięcia jednej wartości odstającej (ekstremalnie odstającej), ponieważ pominięcie zbyt wielu informacji o wysokich dochodach może zaburzyć obraz zamożności.

Porównując wielkości dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych w 2017 r., można zauważyć, że średnie dochody (wyrażone zarówno średnią arytmetyczną, jak i medianą dochodów) są najwyższe w grupie rolników, a najniższe w grupie emerytów i rencistów. Stopień zróżnicowania

dochodów gospodarstw domowych osiągających wysokie dochody ekwiwalentne w grupach społeczno-ekonomicznych jest niejednakowy. Najbardziej niejednorodna jest grupa gospodarstw rolników, czyli takich, których wyłącznym lub głównym źródłem utrzymania jest dochód z użytkowanego gospodarstwa indywidualnego w rolnictwie, oraz grupa gospodarstw, w której źródłem utrzymania jest praca na własny rachunek poza gospodarstwem indywidualnym w rolnictwie lub wykonywanie wolnego zawodu (współczynniki zmienności odchylenia standardowego odpowiednio ponad 81% i 54%). Zamożne gospodarstwa emerytów i rencistów stanowią niewielki odsetek i są dość jednorodne (współczynnik zmienności – 27%).

Wykr. 1. Podstawowe miary położenia i zróżnicowania dochodów gospodarstw domowych w grupach społeczno-ekonomicznych w 2017 r.



Źródło: opracowanie własne na podstawie: GUS (2018a).

Podobnie kształtowała się sytuacja w latach 2014–2016.

Do oszacowania parametrów rozkładu Pareta aproksymującego rozkłady miesięcznych dochodów ekwiwalentnych wykorzystano MNW. Otrzymane wartości

parametrów rozkładu oraz miary określające dobroć dopasowania rozkładu teoretycznego do danych empirycznych przedstawiono w tabl. 3. W przypadku gospodarstw domowych utrzymujących się ze źródeł niezarobkowych nie można było dopasować rozkładu teoretycznego ze względu na zbyt małą liczebność próby.

Tabl. 3. Oszacowania parametrów rozkładu Pareta aproksymującego rozkład dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych oraz miary dopasowania

Gospodarstwa	Lata	Parametry rozkładu		Wps	IM	I_r w %
		θ	a			
O g ó ł e m	2014	4016,200	3,151	0,991	0,018	0,757
	2015	4195,410	3,225	0,994	0,012	1,292
	2016	4477,200	3,318	0,998	0,003	1,712
	2017	4696,101	3,235	0,995	0,009	1,084
Pracowników	2014	4016,197	3,361	0,996	0,008	0,237
	2015	4195,406	3,539	0,993	0,014	0,486
	2016	4477,195	3,666	0,991	0,017	0,158
	2017	4696,101	3,728	0,995	0,010	0,419
Rolników	2014	4029,094	1,856	0,942	0,112	7,111
	2015	4207,700	1,832	0,953	0,091	5,378
	2016	4482,770	1,928	0,988	0,023	4,867
	2017	4715,672	1,736	0,954	0,084	11,600
Pracujących na własny rachunek	2014	4017,000	2,619	0,976	0,042	1,586
	2015	4200,000	2,714	0,983	0,032	0,707
	2016	4483,057	2,940	0,958	0,074	2,070
	2017	4698,938	2,949	0,983	0,034	0,866
Emerytów i rencistów	2014	4017,020	5,030	0,925	0,143	0,457
	2015	4197,188	4,782	0,960	0,079	0,081
	2016	4486,593	4,641	0,971	0,057	0,025
	2017	4702,260	5,097	0,968	0,063	0,123

Uwaga. Wps – wskaźnik podobieństwa struktur, IM – współczynnik Mortary, I_r – współczynnik będący względną różnicą średniej wartości rozkładu empirycznego i oczekiwanej wartości rozkładu teoretycznego.

Źródło: obliczenia własne na podstawie: GUS (2015, 2016, 2017, 2018a).

Rozkład Pareta aproksymujący rozkład dochodów zamożnych gospodarstw rolników ma parametr $a < 2$, co oznacza, że momenty rzędów wyższych niż 1 nie istnieją, a tym samym nie istnieje odchylenie standardowe. Rozkład dochodów w tej grupie charakteryzuje się grubszy prawym ogonem niż rozkłady dochodów w pozostałych grupach społeczno-ekonomicznych.

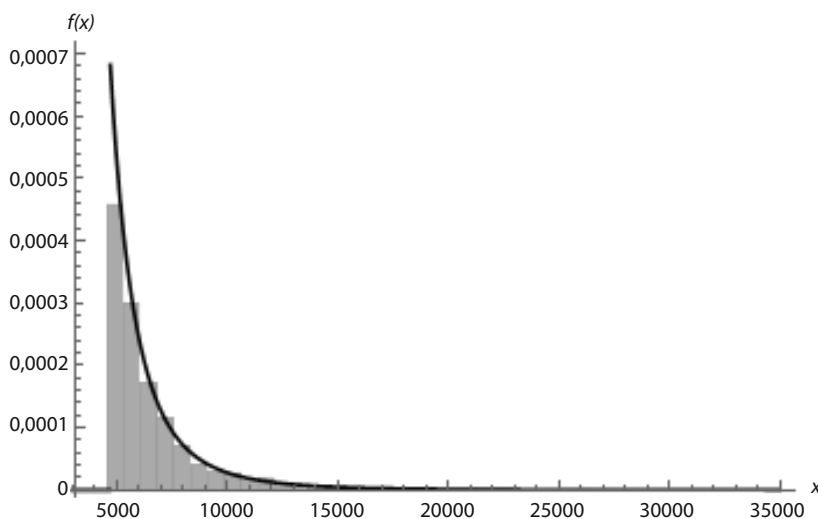
Obliczone miary (Wps, IM, I_r) świadczą o tym, że rozkład Pareta w większości przypadków bardzo dobrze aproksymuje rozkłady empiryczne. Problematiczne może być jedynie dopasowanie rozkładu teoretycznego w grupie gospodarstw rolników w całym analizowanym okresie oraz w grupie gospodarstw emerytów i rencistów w 2014 r. W przypadku grupy rolników o wysokich dochodach ekwiwalentnych współczynnik I_r , określony wzorem (11), jest bliski 5% w 2016 r. (4,87%) lub prze-

kracza 5% w latach 2014, 2015 i 2017, co świadczy o gorszym dopasowaniu rozkładu teoretycznego do danych empirycznych. Pozostałe miary, czyli wskaźnik podobieństwa struktur (większy niż 0,95) i współczynnik Mortara (bliski 0), również pozwalają uznać, że rozkład Pareta z oszacowanymi parametrami dobrze aproksymuje empiryczny rozkład dochodów zamożnych rolników w latach 2015–2017. Tylko w przypadku roku 2014 wszystkie trzy miary przyjmują niesatysfakcjonujące wartości: Wps wynosi nieco poniżej 0,95, I_r jest większy niż 5%, a IM wynosi 0,112. Rozkład Pareta aproksymujący empiryczny rozkład ekwiwalentnych dochodów zamożnych gospodarstw emerytów i rencistów można uznać za gorzej dopasowany, aczkolwiek miara $I_r < 5\%$.

Na wykresach 2 i 3 przedstawiono rozkłady empiryczne wraz z aproksymującymi je teoretycznymi funkcjami gęstości (f) rozkładu Pareta dla dochodów ogółu zamożnych gospodarstw domowych i dla poszczególnych grup społeczno-ekonomicznych w 2017 r.

Dobre dopasowanie rozkładu teoretycznego do danych empirycznych prowadzi do alternatywnego sposobu szacowania parametrów charakteryzujących rozważaną populację, polegającego najpierw na aproksymacji rozkładu empirycznego rozkładem teoretycznym, a następnie na wyznaczeniu odpowiednich wielkości na podstawie oszacowanych parametrów funkcji gęstości rozkładu teoretycznego.

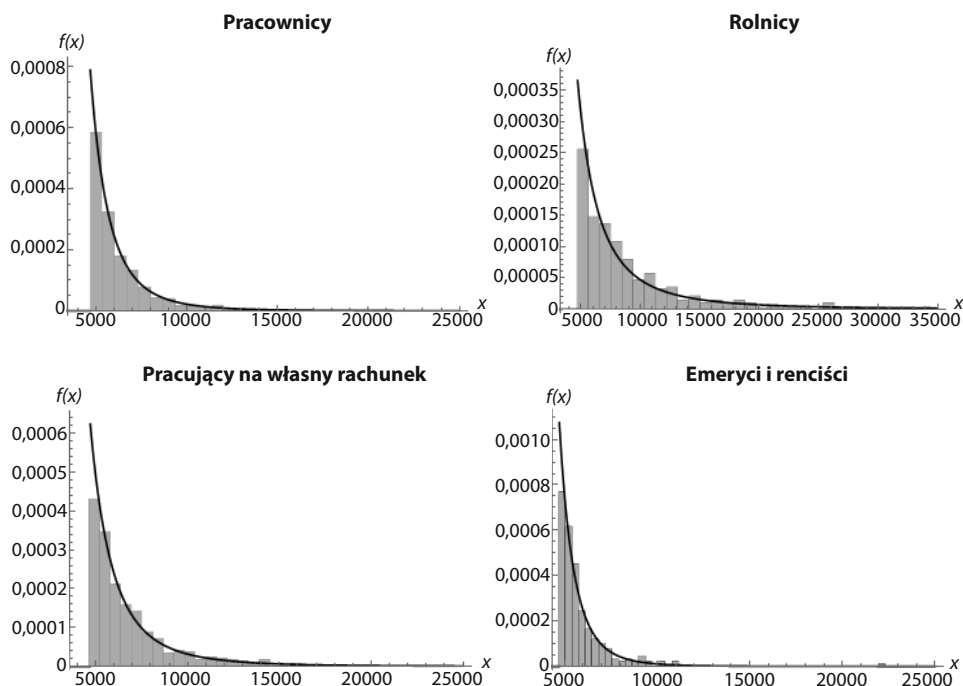
Wykr. 2. Aproksymacja rozkładu dochodów zamożnych gospodarstw domowych ogółem w 2017 r.



Uwaga. x – wielkość dochodu ekwiwalentnego gospodarstwa domowego w zł, $f(x)$ – funkcja gęstości rozkładu Pareta dochodu ekwiwalentnego.

Źródło: opracowanie własne na podstawie: GUS (2018a).

Wykr. 3. Aproksymacja rozkładów dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych w 2017 r.



Uwaga. Jak przy wykr. 2.

Źródło: opracowanie własne na podstawie: GUS (2018a).

Korzystając z otrzymanych rozkładów Pareta, wyznaczono następujące charakterystyki liczbowe rozkładów miesięcznych dochodów ekwiwalentnych w 2017 r.: wartość oczekiwana, mediana, odchylenie standardowe (z wyjątkiem gospodarstw rolników) i miary nierówności (tabl. 4). Różnice między wartościami z próby a teoretycznymi charakterystykami liczbowymi rozkładów są tym mniejsze, im dopasowanie rozkładu teoretycznego do danych empirycznych jest lepsze.

Tabl. 4. Charakterystyki rozkładów Pareta aproksymujących rozkłady dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych w 2017 r.

Gospodarstwa	Wartość oczekiwana	Mediana	Odchylenie standardowe	Współczynnik Giniego	Współczynnik Schutza-Pietry
	w zł				
O g ó ł e m	6797,27	5818,24	3400,66	0,18	0,14
Pracowników	6417,55	5655,69	2528,48	0,15	0,11
Rolników	11122,84	7029,88	.	0,41	0,31
Pracujących na własny rachunek	7109,89	5943,99	4250,03	0,20	0,15
Emerytów i rencistów	5849,99	5387,25	1472,41	0,11	0,08

Źródło: obliczenia własne na podstawie: GUS (2018a).

Porównując rozkłady dochodów zamożnych gospodarstw domowych w grupach społeczno-ekonomicznych, należy stwierdzić, że rozkład dochodów zamożnych gospodarstw rolników charakteryzuje największa wartość oczekiwana i mediana w porównaniu z rozkładami dochodów zamożnych gospodarstw w pozostałych grupach społeczno-ekonomicznych. W grupie rolników obserwuje się również największe nierówności dochodowe, o czym świadczy najwyższa wartość oszacowanego współczynnika Giniego i potwierdza to najwyższa wartość współczynnika Schutza-Pietry. Zamożne gospodarstwa emerytów i rencistów są grupą najbardziej jednorodną (małe odchylenie standardowe, współczynnik zmienności odchylenia standardowego równy 25,17%, nieduża wartość współczynników Giniego i Schutza-Pietry).

Współczynnik Schutza-Pietry określa, jaki odsetek sumy dochodów ekwiwalentnych wszystkich badanych gospodarstw powinien być transferowany z grupy gospodarstw o dochodach powyżej wartości oczekiwanej do grupy gospodarstw o dochodach równych wartości oczekiwanej lub niższych od niej, aby zniknęły nierówności. W przypadku zamożnych gospodarstw rolników transfer powodujący zniwelowanie nierówności dochodowych musiałby wynieść 31% dochodów, a w grupie emerytów – tylko 8%.

W pracy Włodarczyk (2013) przedstawiono analizy dotyczące rozkładów indywidualnych miesięcznych dochodów netto w Polsce w latach 2003–2011, w których najwyższe dochody modelowano rozkładem Pareta. Nierówności dochodowe mierzone współczynnikiem Giniego wzrastały od 0,156 do 0,216 w poszczególnych latach. W przypadku miesięcznych dochodów ekwiwalentnych gospodarstw domowych w latach 2014–2017 nie obserwuje się takiej tendencji. Różnice pomiędzy nimi w poszczególnych latach są znacznie mniejsze: $\hat{G}_{2014} = 0,189$, $\hat{G}_{2015} = 0,183$, $\hat{G}_{2016} = 0,177$ oraz $\hat{G}_{2017} = 0,183$.

5. Podsumowanie

W badaniu omówionym w artykule analizowano rozkłady miesięcznych dochodów ekwiwalentnych zamożnych gospodarstw domowych w poszczególnych grupach społeczno-ekonomicznych w latach 2014–2017. Badano zgodność rozkładów empirycznych z rozkładem Pareta o parametrach oszacowanych na podstawie próby MNW. Dobre dopasowanie rozkładów teoretycznych (w tym przypadku rozkładów Pareta) do rozkładów empirycznych i wykorzystanie parametrów tych rozkładów pozwoliło na oszacowanie istotnych miar charakteryzujących rozkłady ekwiwalentnych dochodów gospodarstw domowych, w tym miar nierównomierności.

W niniejszej pracy za zamożne gospodarstwa domowe uznano te, których miesięczne dochody ekwiwalentne przekroczyły wartość dwukrotności mediany. Innych

granic zamożności nie rozważano (np. trzykrotności mediany), gdyż dane uzyskane z BBGD nie pozwoliłyby na aproksymowanie dochodów w grupie gospodarstw utrzymujących się ze źródeł niezarobkowych, a także w innych grupach, m.in. gospodarstwach emerytów i rencistów, ze względu na zbyt małe próby, co prowadziło do błędów związanych z estymacją parametrów i dopasowaniem rozkładu Pareta do rozkładów empirycznych.

W przypadku zamożnych gospodarstw domowych ogółem, a także gospodarstw: pracowników, pracujących na własny rachunek oraz emerytów i rencistów rozkłady ekwiwalentnych dochodów zostały bardzo dobrze aproksymowane rozkładem Pareta z parametrami oszacowanymi za pomocą MNW, o czym świadczą wartości wszystkich zastosowanych miar dopasowania. Nieznacznie gorsze wyniki otrzymano dla gospodarstw rolników, szczególnie w 2014 r. i 2017 r., oraz gospodarstw emerytów i rencistów w 2014 r., ale i w tych grupach co najmniej jedna spośród trzech miar określających dobroć dopasowania pozwala uznać rozkład Pareta za dobrze aproksymujący dochody zamożnych gospodarstw domowych.

Rozkład dochodów gospodarstw domowych rolników osiągających wysokie dochody różni się od innych grup, gdyż charakteryzuje się grubszy prawym ogonem, a poza tym w tej grupie występuje większe zróżnicowanie wielkości dochodów ekwiwalentnych, o czym świadczą zarówno odchylenie standardowe i współczynnik zmienności odchylenia standardowego, jak i współczynniki Giniego oraz Schutza-Pietry oszacowane na podstawie dopasowanego rozkładu Pareta.

Parametry kształtu rozkładów Pareta były większe od 3 lub bardzo bliskie tej wartości, z wyjątkiem rozkładu dla zamożnych gospodarstw rolników, co świadczy o stosunkowo niewielkim odsetku gospodarstw domowych o skrajnie wysokich dochodach w badanych grupach. Współczynniki Giniego i Schutza-Pietry wskazują, że nierównomierności dochodów zamożnych gospodarstw domowych w badanych grupach społeczno-ekonomicznych są niezbyt duże.

Bibliografia

- Brzeziński, M. (2014). Statistical Inference for Richness Measures. *Applied Economics*, 46(14), 1599–1608. <https://doi.org/10.1080/00036846.2014.880106>.
- Brzeziński, M., Salach, K., Wroński, M. (2020). Wealth inequality in Central and Eastern Europe: Evidence from household survey and rich lists' data combined. *Economics of Transition and Institutional Change*, 28(4), 637–660. <https://doi.org/10.1111/ecot.12257>.
- Charpentier, A., Flachaire, E. (2019). *Pareto Models for Top Incomes* (Working Papers). <https://hal.archives-ouvertes.fr/hal-02145024/document>.
- Chomątowski, S., Sokołowski, A. (1978). Taksonomia struktur. *Przegląd Statystyczny*, 25(2), 217–226.

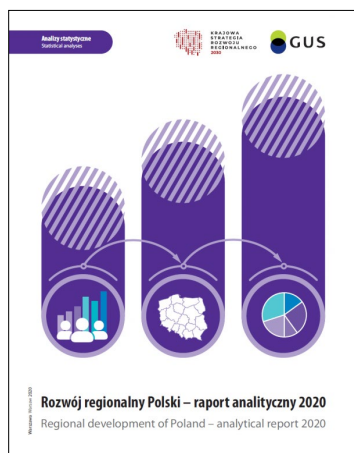
- Chromiński, K., Tkacz, M. (2010). Comparison of outlier detection methods in biometrical data. *Journal of Medical Informatics and Technologies*, 16, 89–94. <http://jmit.us.edu.pl/cms/jmitjrn/16/11-18-Chrominski.pdf>.
- Franzini, M., Granaglia, E., Raitano, M. (2016). *Extreme Inequalities in Contemporary Capitalism: Should We Be Concerned About the Rich?*. Switzerland: Springer International Publishing. <https://doi.org/10.1007/978-3-319-28811-6>.
- Główny Urząd Statystyczny. (2015). *Budżety gospodarstw domowych w 2014 r.* Warszawa. <https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/budzety-gospodarstw-domowych-w-2014-r-,9,9.html>.
- Główny Urząd Statystyczny. (2016). *Budżety gospodarstw domowych w 2015 r.* Warszawa. <https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/budzety-gospodarstw-domowych-w-2015-r-,9,10.html>.
- Główny Urząd Statystyczny. (2017). *Budżety gospodarstw domowych w 2016 r.* Warszawa. <https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/budzety-gospodarstw-domowych-w-2016-r-,9,11.html>.
- Główny Urząd Statystyczny. (2018a). *Budżety gospodarstw domowych w 2017 r.* Warszawa. <https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/budzety-gospodarstw-domowych-w-2017-r-,9,12.html>.
- Główny Urząd Statystyczny. (2018b). *Zeszyt metodologiczny. Badanie budżetów gospodarstw domowych.* Warszawa. <https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/zeszyt-metodologiczny-badanie-budzetow-gospodarstw-domowych,10,2.html>
- Jagielski, M., Kutner, R. (2011). Modelowanie zamożności polskich gospodarstw domowych metodami statystycznymi. *Ekonomia*, (25), 154–162. <http://ekonomia.wne.uw.edu.pl/ekonomia/getFile/716>.
- Jędrzejczak, A. (2006). The Characteristic of Theoretical Income Distributions and their Application to the Analysis of Wage Distributions in Poland by Regions. *Acta Universitatis Lodzensis. Folia Oeconomica*, 196, 183–198.
- Jędrzejczak, A., Pekasiewicz, D. (2020). *Teoretyczne rozkłady dochodów gospodarstw domowych i ich estymacja*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Kleiber, C. (1996). Dagum vs. Singh-Maddala Income Distributions. *Economics Letter*, 53(3), 265–268. [https://doi.org/10.1016/S0165-1765\(96\)00937-8](https://doi.org/10.1016/S0165-1765(96)00937-8).
- Kleiber, C., Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*. Hoboken: John Wiley & Sons. <https://doi.org/10.1002/0471457175>.
- Kot, S. M. (1999). Empiryczna weryfikacja rozkładów płac w Polsce. W: S. M. Kot (red.), *Analiza ekonometryczna kształtowania się płac w Polsce w okresie transformacji* (s. 64–98). Warszawa, Kraków: Wydawnictwo Naukowe PWN.
- Kot, S. M. (2003). Metodologiczne dylematy pomiaru nierówności i dobrobytu. *Nierówności Społeczne a Wzrost Gospodarczy*, (1), 161–179.
- Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Boca Raton: Chapman & Hall/CRC, Taylor & Francis Group.

- Kruszka, K. (1989). Miary podobieństwa struktury obiektów społeczno-ekonomicznych (studium porównawcze). *Akademia Ekonomiczna w Poznaniu, Zeszyty Naukowe, Seria I*, (159), 48–65.
- Mandelbrot, B. (1963). New methods in statistical economics. *Journal of Political Economy*, 71(5), 421–440.
- Ostasiewicz, K. (2013). Adekwatność wybranych rozkładów teoretycznych dochodów w zależności od metody estymacji. *Przegląd Statystyczny*, 60(4), 499–521. <https://ps.stat.gov.pl/Article/2013/4/499-522>.
- Pareto, V. (1897). *Cours d'economie politique*. Lausanne: Rouge.
- Peichl, A., Schaefer, T., Scheicher, C. (2008). *Measuring Richness and Poverty: A Micro Data Application to Europe and Germany* (IZA Discussion Paper No. 3790). <http://ftp.iza.org/dp3790.pdf>. 26.11.2012.
- Safari, M. A. M., Masseran, N., Ibrahim, K. (2019). On the identification of extreme outliers and dragon-kings mechanisms in the upper tail of income distribution. *Journal of Applied Statistics*, 46(10), 1886–1902. <https://doi.org/10.1080/02664763.2019.1566447>.
- Salamaga, M. (2016). Modelowanie rozkładów dochodów kobiet i mężczyzn w województwie małopolskim. *Wiadomości Statystyczne*, 61(8), 32–44. <https://doi.org/10.5604/01.3001.0014.1065>.
- Sączewska-Piotrowska, A. (2018). *Zamożność dochodowa gospodarstw domowych: determinanty, dynamika, zróżnicowanie przestrzenne*. Katowice: Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.
- Włodarczyk, J. (2013). Nierówności dochodowe w Polsce według rozkładów Pareta i Boltzmann-Gibbsa. *Studia Ekonomiczne*, 130, 76–87.
- Zenga, M. M., Pasquazzi, L., Zenga, M. (2010). *Rapporto n. 188: First applications of a new three-parameter distribution for non-negative variables*. Mediolan: Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali, Università degli Studi di Milano Bicocca.

WYDAWNICTWA GUS. KWIECIEŃ 2021 PUBLICATIONS OF STATISTICS POLAND. APRIL 2021

W ofercie wydawniczej Głównego Urzędu Statystycznego z ubiegłego miesiąca warto zwrócić uwagę na następującą publikację:

Among Statistics Poland's last month's publications, we would like to recommend:



Tytuł: *Rozwój regionalny Polski – raport analityczny 2020*

Title: *Regional development of Poland – analytical report 2020*

Język: polski (przedmowa, spis treści oraz synteza dodatkowo w języku angielskim)

Language: Polish (preface, contents and executive summary additionally in English)

Dodatkowe informacje: opracowanie dostępne w wersji elektronicznej

Additional information: publication available in the electronic version

Raport analityczny na temat rozwoju regionalnego kraju, opracowany na potrzeby monitorowania postępów w realizacji Krajowej Strategii Rozwoju Regionalnego 2030. Opracowanie obejmuje dia-

gnozę zjawisk związanych z obserwowanymi zagrożeniami środowiskowymi, procesami demograficznymi, uwarunkowaniami rozwoju kapitału ludzkiego i społecznego, gospodarczego oraz infrastrukturalnego, a także potencjałem administracyjnym samorządów. Podsumowano 2019 r., a dla zilustrowania zmian zachodzących w województwach przywołano dane o poszczególnych zagadnieniach obejmujące okres od 2010 r. W publikacji zamieszczono analizę kwestii rozwojowych dotyczących województw (lub regionów NUTS 2), a tam, gdzie to było możliwe, również obszarów strategicznej interwencji zdefiniowanych w KSRR 2030, czyli Polski Wschodniej, gmin zagrożonych trwałą marginalizacją, miast średniej wielkości tracących funkcje społeczno-gospodarcze, oraz delimitacji przestrzennych związanych z wymiarem miejskim.

Publikacja, wraz z częścią implementacyjną pt. *Sprawozdanie z realizacji Krajowej Strategii Rozwoju Regionalnego do 2030 r.* (przygotowaną przez Ministerstwo Funduszy i Polityki Regionalnej), stanowi kompleksowy raport z monitoringu Krajowej Strategii Rozwoju Regionalnego do 2030 r.

W kwietniu br. ukazały się ponadto:

- *Aktywność ekonomiczna ludności Polski – IV kwartał 2020 r.*;
- „Biuletyn statystyczny” nr 3/2020;

- *Ceny robót budowlano-montażowych i obiektów budowlanych (luty 2021 r.);*
- *Działalność przedsiębiorstw posiadających jednostki zagraniczne w 2019 r.;*
- *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach 2000–2021 (kwiecień 2021);*
- *Ludność. Stan i struktura ludności oraz ruch naturalny w przekroju terytorialnym (stan w dniu 31.12.2020);*
- *Nakłady i wyniki przemysłu w 2020 r.;*
- *Produkcja ważniejszych wyrobów przemysłowych w marcu 2021 r.;*
- *Sytuacja społeczno-gospodarcza kraju w I kwartale 2021 r.;*
- *Sytuacja społeczno-gospodarcza województw Nr 4/2020;*
- *„Wiadomości Statystyczne. The Polish Statistician” nr 4/2021;*
- *Zatrudnienie i wynagrodzenia w gospodarce narodowej w 2020 r.*

Wersje elektroniczne wszystkich publikacji GUS są dostępne na stronie stat.gov.pl/publikacje/publikacje-a-z.

Electronic versions of all the publications by Statistics Poland are available at stat.gov.pl/en/publications.

Justyna Gustyn

Główny Urząd Statystyczny, Departament Opracowań Statystycznych
Statistics Poland, Statistical Products Department

DLA AUTORÓW FOR THE AUTHORS

(for the English translation of the information given below, please visit ws.stat.gov.pl/ForAuthors)

W „Wiadomościach Statystycznych. The Polish Statistician” („WS”) zamieszczane są artykuły o charakterze naukowym poświęcone teorii i praktyce statystycznej, które prezentują wyniki oryginalnych badań teoretycznych lub analitycznych wykorzystujących metody statystyki matematycznej, opisowej bądź ekonometrii. Ukazują się również artykuły przeglądowe, recenzje publikacji naukowych oraz inne opracowania informacyjne. W czasopiśmie publikowane są prace w języku polskim i angielskim.

Od 2007 r. „WS” znajdują się na liście polskich punktowanych czasopism naukowych MEiN. Zgodnie z komunikatem Ministra Edukacji i Nauki z dnia 9 lutego 2021 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych „WS” otrzymały 40 punktów.

„Wiadomości Statystyczne. The Polish Statistician” są udostępniane w następujących bazach indeksacyjnych i repozytoriach: Agro, BazEkon, Central and Eastern European Academic Source (CEEAS), Central and Eastern European Online Library (CEEOL), Central European Journal of Social Sciences and Humanities (CEJSH), EBSCO Discovery Service, European Reference Index for the Humanities and Social Sciences (ERIH Plus), ICI Journals Master List, ICI World of Journals, Norwegian Register for Scientific Journals and Publishers (The Nordic List) oraz POL-index.

Za publikację artykułów na łamach „WS” autorzy nie otrzymują honorariów ani nie wnoszą opłat.

1. Zgłaszanie artykułów

Prace należy przysyłać na adres: redakcja.ws@stat.gov.pl.

Artykuł powinien być utrzymany w formie bezosobowej i zawierać streszczenie, słowa kluczowe, kod/kody JEL oraz ORCID, adres e-mail i afiliację autora. Tytuł, streszczenie i słowa kluczowe powinny być podane w językach polskim i angielskim.

Jeżeli w pracy występują tablice, wykresy lub mapy, powinny być umieszczone w treści artykułu. W osobnym pliku (najlepiej w formacie Excel) należy podać dane do wykresów.

Autor jest zobowiązany do podania w artykule wszelkich źródeł finansowania badań będących podstawą pracy. Jeżeli doszło do zaprezentowania podobnych materiałów podczas konferencji lub sympozjum naukowego, to podczas składania tekstu do publikacji w „WS” należy poinformować o tym redakcję.

Prosimy o niestosowanie stylów i ograniczenie formatowania do wymogów redakcyjnych. Więcej informacji w rozdziale *Wymogi redakcyjne*.

Razem z artykułem należy przesłać skan oświadczenia (do pobrania ze strony internetowej czasopisma) o oryginalności pracy i niezłożeniu jej w innym wydawnictwie, zawierającego zgodę na przeniesienie autorskich praw majątkowych, numer ORCID, afiliację lub afiliacje oraz dane kontaktowe autora, wraz ze wskazaniem proponowanego działu czasopisma. Oryginał oświadczenia należy wysłać na adres: Redakcja „Wiadomości Statystycznych. The Polish Statistician”, Główny Urząd Statystyczny, al. Niepodległości 208, 00-925 Warszawa.

Załączenie skanu oświadczenia jest warunkiem poddania pracy ocenie wstępnej i skierowania do recenzji.

2. Przebieg prac redakcyjnych

Zgłoszony artykuł jest oceniany i opracowywany w czteroetapowym procesie:

1. **Ocena wstępna**, dokonywana przez redakcję. Polega na weryfikacji naukowego charakteru artykułu oraz jego struktury i zawartości pod kątem wymogów redakcyjnych, a także zgodności tematyki z profilem czasopisma. Autor uzupełnia i poprawia artykuł stosownie do uwag redakcji, a w przypadku nieuwzględnienia danej uwagi uzasadnia swoje stanowisko. **Razem z poprawionym artykułem autor przesyła w osobnym pliku zanonimizowaną wersję pracy, przeznaczoną do recenzji.** Anonimizacja polega na utajnieniu nazwiska autora (także we właściwościach pliku), usunięciu podziękowań i informacji o źródłach finansowania, a także innych informacji wskazujących na afiliację lub umożliwiających zidentyfikowanie autora. Warunkiem skierowania pracy do recenzji jest potwierdzenie oryginalności tekstu uzyskane za pomocą systemu antyplagiatowego Similarity Check. W przypadku wykrycia znacznego podobieństwa do innych prac artykuł zostanie odrzucony.
2. **Ocena recenzentów**, dokonywana przez specjalistów w danej dziedzinie. Artykuł oceniają dwaj recenzenci spoza jednostki naukowej, przy której afiliowany jest autor; w przypadku pracy w języku angielskim co najmniej jeden recenzent jest afiliowany przy jednostce zagranicznej. W razie sprzecznych opinii dwóch recenzentów powoływany jest trzeci recenzent. Recenzenci kierują się kryteriami oryginalności i jakości opracowania zarówno w odniesieniu do treści, jak i formy.

Autorzy artykułów, które otrzymały pozytywne oceny, wprowadzają poprawki zalecane przez recenzentów i przesyłają do redakcji zmodyfikowaną wersję pracy. Jeśli pojawi się różnica zdań dotycząca zasadności proponowanych zmian, autorzy są zobligowani do uzasadnienia swojego stanowiska.

3. **Ocena Kolegium Redakcyjnego (KR)**, decydująca o przyjęciu pracy do publikacji. Jest dokonywana na podstawie recenzji, z uwzględnieniem opinii redaktorów tematycznego i merytorycznego. Polega m.in. na weryfikacji dokonania przez autora zmian w artykule stosownie do uwag recenzentów. KR ocenia artykuł pod względem poprawności i spójności merytorycznej oraz zaleca autorowi wprowadzenie poprawek, jeśli są one konieczne, aby praca spełniała wymogi czasopisma. Autorowi przysługuje prawo do odwołania od decyzji o niepublikowaniu artykułu. W takim przypadku powinien on skontaktować się z redakcją „WS” i przedstawić uzasadnienie. Ostateczna decyzja w tej sprawie należy do redaktora naczelnego.

W „WS” publikowane są wyłącznie te artykuły, które otrzymają pozytywną ocenę na każdym z wymienionych etapów i zostaną poprawione przez autora zgodnie z otrzymanymi uwagami (chyba że autor przedstawi argumenty uzasadniające nieuwzględnienie danej uwagi).

Artykuły przyjęte przez KR do publikacji są zamieszczane na stronie internetowej czasopisma w zakładce Early View. Znajdują się tam do czasu opublikowania w konkretnym wydaniu „WS”.

4. **Opracowanie redakcyjne, autoryzacja i korekta.** Artykuł zakwalifikowany do druku jest poddawany opracowaniu merytorycznemu i językowemu. Redakcja zastrzega sobie prawo

do zmiany tytułu i śródtytułów, modyfikowania tablic, wykresów i innych elementów graficznych oraz przeredagowania treści bez naruszenia zasadniczej myśli autora.

Po opracowaniu redakcyjnym artykuł jest przesyłany do autoryzacji. Tekst zatwierdzony przez autora, po składzie i łamaniu, jest poddawany korekcie i rewizji (II korekcie). Autor dokonuje korekty autorskiej tekstu na etapie rewizji. Wykresy i inne materiały graficzne są opracowywane na podstawie danych przekazanych przez autora i poddawane korekcie i rewizji. Autor dokonuje ich akceptacji na etapie rewizji.

W przypadku odkrycia błędów w opublikowanym artykule zamieszcza się na łamach „WS” sprostowanie, a artykuł w wersji elektronicznej jest poprawiany i umieszczany na stronie internetowej „WS” ze stosownym wyjaśnieniem.

3. Zasady etyki publikacyjnej COPE

Redakcja „WS” podejmuje wszelkie starania w celu utrzymania najwyższych standardów etycznych zgodnie z wytycznymi Komitetu ds. Etyki Publikacyjnej (COPE), dostępnymi na stronie internetowej www.publicationethics.org, oraz wykorzystuje wszystkie możliwe środki mające na celu zapobieżenie nadużyciom i nierzetelności autorskiej. Przyjęte zasady postępowania obowiązują autorów, Radę Naukową, Kolegium Redakcyjne, redakcję, pracowników Wydziału Czasopism Naukowych, recenzentów i wydawcę.

3.1. Odpowiedzialność autorów

1. Artykuły naukowe kierowane do opublikowania w „WS” powinny zawierać precyzyjny opis badanych zjawisk i stosowanych metod oraz autorskie wnioski i sugestie dotyczące rozwoju badań i analiz statystycznych. Autorzy powinni wyraźnie określić cel artykułu oraz jasno przedstawić wyniki przeprowadzonej analizy. Prezentacja efektów badań statystycznych zaprojektowanych i przeprowadzonych przez autorów wymaga opisanego zastosowania w nich metodologii. W przypadku nowatorskich metod analizy pożądanym jest podanie przykładu ilustrującego ich zastosowanie w praktyce statystycznej. Autorzy ponoszą odpowiedzialność za treści prezentowane w artykułach. W razie zgłaszania przez czytelników zastrzeżeń odnoszących się do tych treści autorzy są zobligowani do udzielenia odpowiedzi za pośrednictwem redakcji.
2. Na autorach spoczywa obowiązek zapewnienia pełnej oryginalności przedłożonych prac. Redakcja nie toleruje przejawów nierzetelności naukowej autorów, takich jak:
 - duplikowanie publikacji – ponowne publikowanie własnego utworu lub jego części;
 - plagiat – przywłaszczenie cudzego utworu lub jego fragmentu bez podania informacji o źródle;
 - fabrykowanie danych – oparcie pracy naukowej na nieprawdziwych wynikach badań;
 - autorstwo widmo (*ghost authorship*) – nieujawnianie współautorów, mimo że wnieśli oni istotny wkład w powstanie artykułu;
 - autorstwo gościnne (*guest authorship*) – podawanie jako współautorów osób o znikomym udziale lub niebiorących udziału w opracowywaniu artykułu;
 - autorstwo grzecznościowe (*gift authorship*) – podawanie jako współautorów osób, których wkład jest oparty jedynie na słabym powiązaniu z badaniem.

Autorzy deklarują w stosownym oświadczeniu, że zgłaszany artykuł nie narusza praw autorskich osób trzecich, nie był dotychczas publikowany i nie został złożony w innym wydawnictwie oraz że jest ich oryginalnym dziełem, i określają swój wkład w opracowanie artykułu. Jeżeli doszło do zaprezentowania podobnych materiałów podczas konferencji lub sympozjum naukowego, to podczas składania tekstu do publikacji w „WS” autorzy są zobowiązani poinformować o tym redakcję.

3. Autorzy są zobowiązani do podania w treści artykułu wszelkich źródeł finansowania badań będących podstawą pracy.
4. Główną odpowiedzialność za rzetelność przekazanych informacji, łącznie z informacją na temat wkładu poszczególnych współautorów w powstanie artykułu, ponosi zgłaszający artykuł.
5. Autorzy zgłaszający artykuły do publikacji w „WS” biorą udział w procesie recenzji double-blind peer review, dokonywanej przez co najmniej dwóch niezależnych ekspertów z danej dziedziny. Po otrzymaniu pozytywnych recenzji autorzy wprowadzają zalecane przez recenzentów poprawki i dostarczają redakcji zaktualizowaną wersję opracowania wraz z pisemnym poświadczeniem uwzględnienia poprawek. Jeśli pojawi się różnica zdań co do zasadności proponowanych zmian, należy wyjaśnić, które poprawki zostały uwzględnione, a w przypadku ich nieuwzględnienia – uzasadnić swoje stanowisko.
6. Jeżeli autorzy odkryją w swoim maszynopisie lub tekście już opublikowanym błędy, nieścisłości bądź niewłaściwe dane, powinni niezwłocznie poinformować o tym redakcję w celu dokonania korekty, wycofania tekstu lub zamieszczenia sprostowania. W przypadku korekty artykułu już opublikowanego jego nowa wersja jest zamieszczana na stronie internetowej „WS” wraz ze stosownym wyjaśnieniem.

3.2. Odpowiedzialność Rady Naukowej, Kolegium Redakcyjnego i Wydziału Czasopism Naukowych GUS

1. Rada Naukowa (RN) kształtuje profil programowy czasopisma, określa kierunki jego rozwoju i konsultuje jego zakres merytoryczny.
2. Kolegium Redakcyjne (KR) podejmuje decyzję o publikacji danego artykułu z uwzględnieniem ocen recenzentów oraz opinii zespołu redakcyjnego. W swoich rozstrzygnięciach członkowie KR kierują się kryteriami merytorycznej oceny wartości artykułu, jego oryginalności i jasności przekazu, a także ścisłego związku z celem i zakresem tematycznym „WS”. Oceniają artykuły niezależnie od płci, rasy, pochodzenia etnicznego, narodowości, religii, wyznania, światopoglądu, niepełnosprawności, wieku lub orientacji seksualnej ich autorów.
3. Zespół redakcyjny, wyodrębniony z KR, tworzą redaktor naczelny i jego zastępcę, redaktorzy tematyczni i redaktor merytoryczny. Członkowie zespołu redakcyjnego weryfikują nadane artykuły pod względem merytorycznym, oceniają ich zgodność z celem i zakresem tematycznym „WS” oraz sprawdzają spełnienie wymogów redakcyjnych i przestrzeganie zasad rzetelności naukowej. Ponadto wybierają recenzentów w taki sposób, aby nie wystąpił konflikt interesów, i dbają o zapewnienie uczciwego, bezstronnego i terminowego procesu recenzowania.
4. Za sprawny przebieg procesu wydawniczego, informowanie wszystkich jego uczestników o konieczności przestrzegania obowiązujących zasad i przygotowanie artykułów do

publikacji odpowiadają pracownicy Wydziału Czasopism Naukowych (WCN) GUS. W celu uzyskania obiektywnej oceny oryginalności nadsyłanych artykułów przed skierowaniem ich do recenzji WCN wykorzystuje system antyplagiatowy. Informacje dotyczące artykułu mogą być przekazywane przez WCN wyłącznie autorom, recenzentom, członkom RN i KR oraz wydawcy.

5. Zmiany dokonane w tekście na etapie przygotowania artykułu do publikacji nie mogą naruszać zasadniczej myśli autorów. Wszelkie modyfikacje o charakterze merytorycznym są z nimi konsultowane.
6. W przypadku podjęcia decyzji o niepublikowaniu artykułu nie może on zostać w żaden sposób wykorzystany przez wydawcę lub uczestników procesu wydawniczego bez pisemnej zgody autorów. Autorzy mogą się odwołać od decyzji o niepublikowaniu artykułu. W tym celu powinni się skontaktować z redaktorem naczelnym lub sekretarzem redakcji „WS” i przedstawić stosowną argumentację. Odwołania autorów są rozpatrywane przez redaktora naczelnego.
7. Członkowie RN i KR ani pracownicy WCN nie mogą pozostawać w jakimkolwiek konflikcie interesów w odniesieniu do artykułów zgłaszanych do publikacji. Przez konflikt interesów należy rozumieć sytuację, w której jakiekolwiek interesy lub zależności (służbowe, finansowe lub inne) mogą mieć wpływ na ocenę artykułu lub decyzję o jego publikacji.
8. W celu przeciwdziałania nierzetelności naukowej wymagane jest złożenie przez autorów oświadczenia, w którym deklarują, że zgłaszany artykuł nie narusza praw autorskich osób trzecich, nie był dotychczas publikowany i jest ich oryginalnym dziełem, a także określają swój wkład w opracowanie artykułu.
9. W celu zapewnienia wysokiej jakości recenzji wymagane jest złożenie przez recenzentów oświadczenia o przestrzeganiu zasad etyki recenzowania COPE i niewystępowaniu konfliktu interesów.
10. W przypadku uzasadnionego podejrzenia na jakimkolwiek etapie procesu wydawniczego, że autorzy dopuścili się nierzetelności naukowej (zob. pkt 3.1. Odpowiedzialność autorów), zespół redakcyjny skrupulatnie zbada sprawę ewentualnego nadużycia. Jeśli nierzetelność autorów zostanie udowodniona, to zgłoszony przez nich artykuł zostanie odrzucony przez KR, a autorzy otrzymają informację o podjętej decyzji wraz z jej uzasadnieniem.
11. Czytelnicy, którzy mają wobec autorów opublikowanego artykułu uzasadnione podejrzenia o nierzetelność naukową, powinni powiadomić o tym redaktora naczelnego lub sekretarza redakcji. Po zbadaniu sprawy ewentualnego nadużycia czytelnicy zostaną poinformowani o rezultacie przeprowadzonego postępowania. W przypadku potwierdzenia nadużycia, na łamach czasopisma zostanie zamieszczona stosowna informacja.

3.3. Odpowiedzialność recenzentów

1. Recenzenci przyjmują artykuł do recenzji tylko wtedy, gdy uznają, że:
 - posiadają odpowiednią wiedzę w określonej dziedzinie, aby rzetelnie ocenić pracę;
 - zgodnie z ich stanem wiedzy nie istnieje konflikt interesów w odniesieniu do autorów, przedstawionych w artykule badań i instytucji je finansujących, co potwierdzają w oświadczeniu;
 - mogą wywiązać się z terminu ustalonego przez redakcję, aby nie opóźnić publikacji.

2. Recenzenci są zobligowani do zachowania obiektywności i poufności oraz powstrzymania się od osobistej krytyki. Zawsze powinni uzasadnić swoją ocenę, przedstawiając stosowną argumentację.
3. Recenzenci powinni wskazać ważne dla wyników badań opublikowane prace, które w ich ocenie powinny zostać przywołane w ocenianym artykule.
4. W razie stwierdzenia wysokiego poziomu zbieżności treści recenzowanej pracy z innymi opublikowanymi materiałami lub podejrzenia innych przejawów nierzetelności naukowej recenzenci są zobowiązani poinformować o tym redakcję.
5. Po ukończeniu recenzji przechowywanie przesłanych przez redakcję materiałów (w jakiegokolwiek formie) oraz posługiwanie się nimi przez recenzentów jest niedozwolone.

3.4. Odpowiedzialność wydawcy

1. Materiały opublikowane w „WS” są chronione prawem autorskim.
2. Wydawca udostępnia pełną treść wszystkich artykułów w Internecie w trybie otwartego dostępu, tj. bezpłatnie i bez technicznych ograniczeń. Użytkownicy mogą czytać, pobierać, kopiować, drukować i wykorzystywać do innych celów artykuły zamieszczone online, zgodnie z właściwymi przepisami o dozwolonym użytku, pod warunkiem wskazania źródła pochodzenia artykułu. Inne sposoby wykorzystania treści artykułów „WS” wymagają zgody wydawcy.
3. Wydawca deklaruje gotowość do opublikowania poprawek, wyjaśnień oraz przeprosin.

4. Wymogi redakcyjne

Zgodnie z wymogami czasopisma omawiany w artykule problem badawczy powinien być jednoznacznie zdefiniowany oraz istotny dla oceny zjawisk społecznych lub gospodarczych. Artykuł powinien zawierać wyraźnie określony cel badania, precyzyjny opis badanych zjawisk i stosowanych metod, uzyskane wyniki przeprowadzonej analizy oraz autorskie wnioski.

4.1. Struktura i zawartość artykułu

Wymagane elementy artykułu:

1. Tytuł.
2. Dane autora: imię i nazwisko, ORCID, adres e-mail, afiliacja. Wśród autorów artykułu wieloautorskiego należy wskazać autora korespondencyjnego.
3. Streszczenie (zalecana objętość – do 1200 znaków ze spacjami, forma bezosobowa). W przypadku artykułu opisującego badanie empiryczne powinno zawierać: cel, przedmiot, okres i metodę badania, źródła danych i najważniejsze wnioski z badania. W przypadku artykułów o innym charakterze należy podać co najmniej cel pracy, jej przedmiot i najważniejsze wnioski.

Streszczenie to podstawowe źródło informacji o artykule, warunkujące też decyzję czytelnika o zapoznaniu się z całą pracą. Dlatego powinno być przygotowane ze szczególną starannością i dbałością o umieszczenie w nim wszystkich wymaganych elementów.

4. Słowa kluczowe – najistotniejsze pojęcia lub wyrażenia użyte w pracy (nie mniej niż trzy). Powinny być zawarte w streszczeniu i/lub tytule.
5. Kod/kody z klasyfikacji Journal of Economic Literature (JEL).
6. Tłumaczenie tytułu, streszczenia i słów kluczowych (na język angielski w przypadku artykułu napisanego w języku polskim, a na język polski w przypadku artykułu napisanego w języku angielskim).
7. W artykule opisującym badanie empiryczne wymagane są następujące części:
 - wprowadzenie, zawierające: syntetyczne przedstawienie zagadnień teoretycznych, uzasadnienie podjęcia danego problemu badawczego, cel badania i krytyczne odniesienie do literatury przedmiotu. W wyjątkowych przypadkach, kiedy istotne dla podjętego tematu jest obszerniejsze przedstawienie dyskusji toczącej się w literaturze, przegląd literatury może stanowić odrębną część artykułu;
 - metoda badania, zawierająca: przedmiot i okres badania, źródła danych i zastosowane metody badawcze, w tym uzasadnienie ich wyboru;
 - wyniki badania wraz z wnioskami;
 - podsumowanie, które powinno być zwięzłe i odzwierciedlać istotę problemu badawczego przedstawionego w artykule, bez podawania danych liczbowych; końcowe wnioski powinny odnosić się do treści artykułu, a w szczególności do celu badania.Wszystkie części powinny być opatrzone numerami.
8. Bibliografia, zawierająca pełny wykaz prac i materiałów przywołanych w artykule, przygotowana zgodnie z wymogami czasopisma.

4.2. Przygotowanie artykułu

1. Artykuł powinien być utrzymany w formie bezosobowej.
2. Tekst należy zapisać alfabetem łacińskim. Nazwy własne, tytuły itp. oryginalnie zapisane innym alfabetem powinny być poddane transliteracji.
3. Nie należy stosować stylów; formatowanie należy ograniczyć do wymogów redakcyjnych.
4. Objętość artykułu łącznie ze streszczeniem, słowami kluczowymi, bibliografią, tablicami, wykresami i innymi materiałami graficznymi nie powinna być mniejsza niż 10 stron maszynopisu ani przekraczać 20 stron.
5. Edytor tekstu: Microsoft Word, format *.doc lub *.docx.
6. Krój czcionki:
 - Arial – tytuł, autor, streszczenia, słowa kluczowe, kody JEL, śródtytuły, elementy graficzne (tablice, zestawienia, wykresy, schematy), przypisy;
 - Times New Roman – tekst główny, bibliografia.
7. Wielkość czcionki:
 - 14 pkt – tytuł, autor, tytuły rozdziałów;
 - 12 pkt – tekst główny, tytuły podrozdziałów;
 - 10 pkt – pozostałe elementy.
8. Marginesy – 2,5 cm z każdej strony.
9. Interlinia – 1,5 wiersza; tablice i przypisy – 1 wiersz; przed tytułami rozdziałów i podrozdziałów oraz po nich – pusty wiersz.
10. Wcięcie akapitowe – 0,4 cm; bibliografia – bez wcięcia, wysunięcie 0,4 cm.

11. Przy wyliczeniach należy posłużyć się listą punktowaną z punktorami w postaci kropek (wysunięcie 0,4 cm, wcięcie 0 cm); wiersze (oprócz ostatniego) zakończone średnikiem.
12. Strony ponumerowane automatycznie.
13. Tablice i elementy graficzne (wykresy, mapy, schematy) muszą być przywołane w tekście.
14. Wykresy, mapy i schematy należy zamieścić w tekście głównym. Wykresy powinny być edytowalne (optymalnie wykonane w programie Excel; w przypadku wykonania w programie graficznym powinny mieć postać wektorową). Dane, na podstawie których opracowano wykresy, należy przekazać osobno w pliku programu Excel (lub innym edytowalnym w pakiecie Microsoft Office), ewentualnie wykresy powinny dawać możliwość odczytania z nich danych.
15. Tablice muszą być edytowalne. Nie należy stosować rastrów, cieniowania, pogrubiania czy też podwójnych linii itp.
16. Wskazówki dotyczące opracowywania map znajdują się w publikacji *Mapy statystyczne. Opracowanie i prezentacja danych*, dostępnej na stronie internetowej GUS.
17. Pod tablicami i każdym elementem graficznym należy podać źródło opracowania, a także objaśnić użyte w nich skróty i symbole.
18. Literowe symbole liczb i innych wielkości niezłożonych należy zapisywać małą lub dużą literą i pismem pochyłym (np. a , A , $y(x)$, a_i); wektorów – małą literą, pismem pochyłym i pogrubionym (np. $\mathbf{a}$, $\mathbf{w}$, $\mathbf{y}(x)$, $\mathbf{w}_i$); macierzy – dużą literą i pismem pogrubionym (np. $\mathbf{A}$, $\mathbf{M}$, $\mathbf{Y}(x)$, $\mathbf{M}_i$).
19. Objasnienia znaków umownych w tablicach: kreska (–) – zjawisko nie wystąpiło; zero (0) – zjawisko istniało w wielkości mniejszej od 0,5; (0,0) – zjawisko istniało w wielkości mniejszej od 0,05; kropka (.) – brak informacji, konieczność zachowania tajemnicy statystycznej, wypełnienie pozycji jest niemożliwe lub niecelowe; „w tym” – oznacza, że nie podaje się wszystkich składników sumy.
20. Stosowane są następujące skróty: tablica – tabl., wykres – wyk.
21. Wszystkie zawarte w artykule informacje, dane i stwierdzenia wykraczające poza wiedzę powszechną – np. wyniki badań innych autorów, zarówno o charakterze empirycznym, jak i koncepcyjnym – muszą być opatrzone przypisem bibliograficznym. Przez wiedzę powszechną należy rozumieć informacje ogólnie znane i niebudzące wątpliwości ani kontrowersji w danej grupie społecznej, np. utworzenie GUS w 1918 r. lub powstanie UE w 1993 r. na podstawie traktatu z Maastricht. Natomiast dane statystyczne udostępniane lub publikowane np. przez GUS lub Eurostat nie należą do takich informacji. Charakteru wiedzy powszechnej nie mają również stwierdzenia odnoszące się do idei, zjawisk i procesów społecznych, politycznych czy gospodarczych. Nawet pozornie zdroworozsądkowe idee zmieniają bowiem swój sens w zależności od kultury, języka lub dyscypliny naukowej, a także bywają w rozmaity sposób konceptualizowane, jak np. pojęcie poznania w naukach społecznych.

Podanie źródła jest konieczne niezależnie od tego, czy informacje lub stwierdzenia są ujęte w ramy cytatu, czy przedstawione bez dosłownego przytoczenia, np. w formie parafrazy. Jeżeli stwierdzenie może budzić jakiegokolwiek wątpliwości odbiorców, autor powinien wskazać stosowne źródło podawanej informacji.

22. Przypisy rzeczowe, słownikowe lub informacyjne należy umieszczać na dole strony. Przypisy bibliograficzne, zgodnie ze standardem APA (American Psychological Association), należy podawać w tekście głównym.
23. Bibliografię należy przygotować zgodnie ze standardem APA.

4.3. Zasady przywoływania publikacji w treści artykułu

Wyszczególnienie	Przykład przywołania	
	w odsyłaczu	w treści zdania
Autor indywidualny		
Jeden autor	(Iksiński, 2001)	Iksiński (2001)
Dwóch autorów	(Iksiński i Nowak, 1999)	Iksiński i Nowak (1999)
Trzech autorów lub więcej	(Jankiewicz i in., 2003)	Jankiewicz i in. (2003)
Autor instytucjonalny		
Nazwa funkcjonuje jako powszechnie znany skrótowiec: pierwsze przywołanie w tekście	(International Labour Organization [ILO], 2020)	International Labour Organization (ILO, 2020)
kolejne przywołanie	(ILO, 2020)	ILO (2020)
Pełna nazwa	(Stanford University, 1995)	Stanford University (1995)
Typ publikacji		
Publikacja bez ustalonego autorstwa	(Skrócony tytuł ..., 2015)	Pełny tytuł (2015)
Publikacja bez roku wydania	(Iksiński, b.r.)	Iksiński (b.r.)
Akt prawny	(Pełny tytuł)	Pełny tytuł
Strona internetowa / Zbiór danych: znana data publikacji	(Iksiński, 2020) / (Nazwa instytucji, 2020)	Iksiński (2020) / Nazwa instytucji (2020)
nieznana data publikacji	(Iksiński, b.r.) / (Nazwa instytucji, b.r.)	Iksiński (b.r.) / Nazwa instytucji (b.r.)
Rodzaj przywołania		
Przywoływanie kilku prac (porządek prac – chronologiczny, porządek autorów – alfabetyczny)	(Iksiński, 1997, 1999, 2004a, 2004b; Nowak, 2002)	Iksiński (1997, 1999, 2004a, 2004b) i Nowak (2002)
Przywoływanie publikacji za innym autorem (uwaga: w bibliografii należy wymienić tylko pracę cytowaną)	(Nowakowski, 1990, za: Zieniecka, 2007)	Nowakowski (1990, za: Zieniecka, 2007)

Źródło: opracowanie na podstawie: American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th edition). <https://doi.org/10.1037/0000165-000>.

4.4. Przykłady opisu bibliograficznego

Bibliografia powinna być zamieszczona na końcu opracowania. Prace należy podać alfabetycznie według nazwiska pierwszego autora. W przypadku dwóch lub więcej prac tego samego autora / tych samych autorów trzeba je uporządkować chronologicznie według roku publikacji. Jeśli kilka prac tego samego autora / tych samych autorów zostało opublikowanych w tym samym roku, należy ułożyć je alfabetycznie według tytułu i odpowiednio oznaczyć literami a, b, c itd.

Typ publikacji	Przykład opisu bibliograficznego
Artykuł w czasopiśmie	
W wersji drukowanej	Nazwisko, X. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik</i> (zeszyt), strona początku–strona końca.
Dostępny w internecie, z DOI	Nazwisko, X., Nazwisko 2, Y. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik</i> (zeszyt), strona początku–strona końca. https://doi.org/xxx .
Dostępny w internecie, bez DOI	Nazwisko, X., Nazwisko 2, Y., Nazwisko 3, Z. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik</i> (zeszyt), strona początku–strona końca. https://xxx .
Maszynopis	
Niepublikowany / w przygotowaniu / zgłoszony do publikacji	Nazwisko, X. (rok). <i>Tytuł [maszynopis niepublikowany / w przygotowaniu / zgłoszony do publikacji]</i> . Nazwa instytucji, w której powstaje lub powstał maszynopis.
Opublikowany nieformalnie	Nazwisko, X., Nazwisko 2, Y. (rok). <i>Tytuł artykułu</i> . https://xxx .
Książka	
W wersji drukowanej	Nazwisko, X. (rok). <i>Tytuł książki</i> . Miejsce wydania: Wydawnictwo.
Dostępna w internecie, z DOI	Nazwisko, X. (rok). <i>Tytuł książki</i> . Miejsce wydania: Wydawnictwo. https://doi.org/xxx .
Dostępna w internecie, bez DOI	Nazwisko, X. (rok). <i>Tytuł książki</i> . Miejsce wydania: Wydawnictwo. https://xxx .
W przekładzie	Nazwisko, X. (rok). <i>Tytuł książki</i> (tłum. Y. Nazwisko). Miejsce wydania: Wydawnictwo.
Wydanie wielotomowe: tom zatytułowany	Nazwisko, X. (rok). <i>Tytuł książki: nr tomu. Tytuł tomu</i> . Miejsce wydania: Wydawnictwo.
tom niezatytułowany	Nazwisko, X. (rok). <i>Tytuł książki</i> (nr tomu). Miejsce wydania: Wydawnictwo.
Kolejne wydanie	Nazwisko, X. (rok). <i>Tytuł książki</i> (nr wydania). Miejsce wydania: Wydawnictwo.
Pod redakcją: w języku polskim	Nazwisko, X. (red.). (rok). <i>Tytuł książki</i> . Miejsce wydania: Wydawnictwo.
w języku angielskim	Nazwisko, X. (Ed.). (rok). <i>Tytuł książki</i> . Miejsce wydania: Wydawnictwo.
Rozdział w pracy zbiorowej	Nazwisko, X. (rok). Tytuł rozdziału. W: Y. Nazwisko, Z. Nazwisko 2 (red.), <i>Tytuł książki</i> (s. strona początku–strona końca). Miejsce wydania: Wydawnictwo. https://doi.org/xxx lub https://xxx .
Inne prace	
Raport: autor indywidualny	Nazwisko, X. (rok). <i>Tytuł raportu</i> . Miejsce wydania: Wydawnictwo.
autor instytucjonalny	Nazwa instytucji. (rok). <i>Tytuł raportu</i> . Miejsce wydania: Wydawnictwo.
Working Papers	Nazwisko, X. (rok). <i>Tytuł pracy</i> (nazwa serii i numer). https://doi.org/xxx lub https://xxx .
Materiały z konferencji: nieopublikowane	Nazwisko. X. (rok, dzień i miesiąc). <i>Tytuł pracy</i> [typ wystąpienia, np. referat]. Nazwa konferencji, miejsce konferencji.
opublikowane	Nazwisko. X. (rok). <i>Tytuł pracy</i> . Nazwa konferencji, miejsce konferencji.
Rozprawa doktorska: nieopublikowana	Nazwisko, X. (rok). <i>Tytuł pracy</i> [niepublikowana rozprawa doktorska]. Nazwa instytucji nadającej tytuł doktorski.
opublikowana	Nazwisko, X. (rok). <i>Tytuł pracy</i> [rozprawa doktorska, nazwa instytucji nadającej tytuł doktorski]. https://xxx .
Akt prawny	Pełny tytuł aktu prawnego wraz z datą publikacji w dzienniku urzędowym.

Typ publikacji	Przykład opisu bibliograficznego
Strona internetowa	
Znana data publikacji, zawartość strony się nie zmienia	Nazwisko, X. (rok, dzień i miesiąc). <i>Tytuł</i> . https://xxx .
Nieznana data publikacji, zawartość strony się zmienia	Nazwa instytucji. (b.r.). <i>Tytuł</i> . Pobrane dzień, miesiąc i rok pobrania z https://xxx .
Zbiór danych	
Surowe dane nieopublikowane	Nazwisko, X. (rok wydania pracy, w której dane są wykorzystywane) [opis danych, np. surowe dane nieopublikowane dotyczące...]. Źródło danych (np. nazwa uniwersytetu).
Dane opublikowane: znana data publikacji, zawartość zbioru się nie zmienia	Nazwisko, X. (rok). <i>Nazwa zbioru danych</i> [zbiór danych]. Wydawca. https://xxx .
nieznana data publikacji, zawartość zbioru się zmienia	Nazwa instytucji. (b.r.). <i>Nazwa zbioru danych</i> [zbiór danych]. Wydawca. Pobrane dzień, miesiąc i rok pobrania z https://xxx .

Źródło: opracowanie na podstawie: American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th edition). <https://doi.org/10.1037/0000165-000>.

Praca przygotowana w sposób niezgodny z powyższymi wskazówkami będzie odesłana do Autora z prośbą o dostosowanie formy artykułu do wymogów redakcyjnych.

ZAKRES TEMATYCZNY DZIAŁÓW

THEMATIC SCOPE OF SECTIONS

(for the English translation of the information given below, please visit ws.stat.gov.pl/AimScope)

Studia metodologiczne

W tym dziale zamieszczane są artykuły naukowe przedstawiające teoretyczne rozwiązania metodologiczne ze wskazaniem ich praktycznej użyteczności, w tym prace przeglądowe i porównawcze oraz dotyczące etyki w statystyce. Poruszane w nich zagadnienia obejmują różne dziedziny statystyki, ekonomii matematycznej i ekonometrii. Omawiane rezultaty badawcze mogą znaleźć efektywne zastosowanie w badaniach empirycznych oraz analizach statystycznych i służyć podnoszeniu ich jakości, jak również powiększeniu zasobu informacyjnego.

Statystyka w praktyce

Dział ten zawiera artykuły poświęcone nowatorskim zastosowaniom w praktyce znanych narzędzi i modeli statystycznych oraz analizie i ocenie statystycznej zjawisk społeczno-ekonomicznych i innych; zamieszczane tu prace opierają się w szczególności na danych pochodzących z zasobów statystyki publicznej. Zastosowania w praktyce obejmują również wykorzystanie narzędzi informatycznych do uzyskiwania i przetwarzania informacji statystycznych, naliczania danych wynikowych, ich prezentacji i rozpowszechniania. Może to też dotyczyć opracowań stosujących nowoczesne techniki programistyczne pozwalające na efektywną komunikację z systemami informacyjnymi oraz ułatwiające wykorzystanie danych wynikowych. Publikowane są także artykuły sygnalizujące problemy związane z projektowaniem badań statystycznych, uzyskiwaniem, integracją i przetwarzaniem danych oraz generowaniem wynikowych informacji statystycznych i kontrolą ich ujawniania wraz z propozycjami efektywnych rozwiązań w tym zakresie.

Studia interdyscyplinarne. Wyzwania badawcze

To blok tematyczny zawierający artykuły wskazujące i podejmujące wyzwania badawcze, które są szczególnie istotne ze względu na rosnące potrzeby współczesnych użytkowników danych statystycznych i wymagają zaangażowania znacznych nakładów pracy, środków oraz rozwiązań z różnych dziedzin nauki i techniki. W dziale tym publikowane są również opracowania dotyczące: wykorzystania technologii informacyjnych i komunikacyjnych (ICT), gospodarki opartej na wiedzy, problematyki innowacyjności, przepływu informacji we współczesnym społeczeństwie oraz przetwarzania i analizy zagadnień związanych z data science i big data, a zatem problematyki bardzo często powiązanej z działaniami interdyscyplinarnymi.

Edukacja statystyczna

W tym dziale zamieszczane są artykuły dotyczące metod i efektów nauczania statystyki oraz popularyzacji myślenia statystycznego. Odnosi się to zwłaszcza do problemów związanych z kształceniem w zakresie umiejętności stosowania statystyki na wszystkich poziomach edukacji, a także do wykorzystywania nowoczesnych koncepcji i metod dydaktycznych oraz pomocy naukowych w nauczaniu statystyki. Uwaga skoncentrowana jest na rozumieniu prawdopodobieństwa i statystyki, badaniach z zakresu nauczania statystyki, postaw i zachowań społecznych w odniesieniu do tej dziedziny wiedzy, jak również na rozumieniu informacji statystycznych. Ponadto ukazywane są problemy związane z prezentacją danych statystycznych oraz ich interpretacją w powszechnym obiegu informacyjnym, np. w środkach społecznego przekazu.

Z dziejów statystyki

Prace publikowane w tym dziale poświęcone są historii prowadzenia obserwacji statystycznych oraz rozwoju ich metodologii i narzędzi. Ponadto zamieszczane są tu informacje dotyczące życia i osiągnięć zawodowych wybitnych statystyków, jak również najważniejszych instytucji i organizacji statystycznych w Polsce i za granicą.

Dyskusje. Recenzje. Informacje

Jedyny dział zawierający teksty nierecenzowane i niemające charakteru artykułów naukowych. Obejmuje informacje o najważniejszych wydarzeniach dotyczących statystyki polskiej i międzynarodowej, a także sprawozdania z konferencji naukowych, recenzje książek i opracowań z zakresu statystyki i jej zastosowań, rekomendacje nowych, istotnych i ciekawych pozycji wydawniczych z tego obszaru wiedzy, jak również odpowiedzi autorów na recenzje oraz polemiki, dyskusje i sprostowania dotyczące artykułów zamieszczonych na łamach czasopisma.