

Cena 13,00 zł
(VAT 8%)

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

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

CZERWIEC / JUNE
ROK / VOLUME 65

2020 | 6

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

POLSKIE TOWARZYSTWO STATYSTYCZNE
POLISH STATISTICAL ASSOCIATION



WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

CZERWIEC / JUNE
ROK / VOLUME 65

2020 | 6 (709)

RADA NAUKOWA / SCIENTIFIC COUNCIL

dr Dominik Rozkrut (przewodniczący/chairman) – Uniwersytet Szczeciński, Prof. Anthony Arundel – University of Maastricht, dr hab. Bożena Balcerzak-Paradowska, Prof. Eric Bartelsman, PhD – Vrije Universiteit Amsterdam, prof. dr hab. Czesław Domański – Uniwersytet Łódzki, prof. dr hab. Elżbieta Gołata – Uniwersytet Ekonomiczny w Poznaniu, Prof. Semen Matkovskiy, PhD – 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, Prof. Juan Manuel Rodríguez Poo, PhD – University of Cantabria, Assoc. Prof. Iveta Stankovičová, BEng, PhD – 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

Prof. Tudorel Andrei, PhD – Bucharest Academy of Economic Studies, mgr Renata Bielak – Główny Urząd Statystyczny, dr Marek Cierpień-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, mgr Władysław Wiesław Łagodziński – Polskie Towarzystwo Statystyczne, dr Grażyna Marciniak, dr hab. Andrzej Młodak, prof. PWSZ – Państwowa Wyższa Szkoła Zawodowa im. Prezydenta Stanisława Wojciechowskiego w Kaliszu, dr Stanisław Paradysz, dr hab. Mateusz Pipień, prof. UEK – Uniwersytet Ekonomiczny w Krakowie, Marek Rojčček, BEng, PhD – University of Economics, Prague, Assoc. Prof. Anna Shostya, PhD – 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 Cierpień-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 Zygmont

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 Journal 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

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 nabywania czasopism / Information on purchasing 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

PEŁNA LISTA RECENZENTÓW OCENIAJĄCYCH ARTYKUŁY W 2019 ROKU FULL LIST OF REVIEWERS WHO REFEREED MANUSCRIPTS IN 2019

Składamy serdeczne podziękowania Recenzentom, którzy służyli nam wiedzą i doświadczeniem i poświęcili czas na wykonanie recenzji artykułów zgłoszonych do naszego czasopisma. W 2019 r. w tym gronie znaleźli się:

We would like to thank the following Reviewers, who in 2019 shared their expertise and knowledge with us and devoted their valuable time to assess the articles submitted for publication in our journal:

Marta Anacka	Magdalena Knapieńska	Eliza Przeździecka
Andrzej Bąk	Ryszard Kokoszczynski	Grażyna Rosa
Maciej Beręsewicz	Adam Kubów	Anna Sączewska-Piotrowska
Katarzyna Bień-Barkowska	Leszek Kucharski	Sven Schreiber
Tadeusz Borys	Janusz Kudła	Teresa Słaby
Marek Bryx	Tadeusz Kudłacz	Beata Stachowiak
Ewa Chodakowska	Emanuel Kulczycki	Iveta Stankovičová
Dariusz Chojceki	Sławomir Kurek	Monika Stanny
Beata Czarnacka-Chrobot	Ireneusz Kuroпка	Małgorzata Stec
Elżbieta Czarny	Joanna Landmesser	Bogdan Stefanowicz
Barbara Dańska-Borsiak	Janusz Łyko	Mirosław Szreder
Wiesław Dębski	Grzegorz Maciejewski	Urszula Sztandar-Sztanderska
Czesław Domański	Sebastian Majewski	Izabela Sztangret
Piotr Drygaś	Anna Malina	Piotr Szukalski
Andrzej Dudek	Grażyna Marciniak	Tomasz Przemysław Śleszyński
Grzegorz Dudek	Małgorzata Markowska	Dominik Śliwicki
Hanna Dudek	Ewa Mazurek-Krasodomska	Grzegorz Ślusarz
Tadeusz Dudyc	Grzegorz Mentel	Michał Twardochleb
Włodzimierz Dzun	Wawrzyniec Michalczyk	Joanna Tyrowicz
Józefa Famielc	Ewa Miklaszewska	Piotr Tyszko
Stanisław Flejterski	Andrzej Miszczuk	Marek Walesiak
Małgorzata Fronczek	Jacek Mizerka	Agnieszka Wałęga
Eugeniusz Gatnar	Ruslan Motoryn	Stanisław Wanat
Jan Gawelko	Krzysztof Najman	Zenon Wiśniewski
Ewa Giermanowska	Joanicjusz Nazarko	Dorota Witkowska
Michał Goliński	Wojciech Niemiro	Bogdan Wojtyniak
Wiesław Golnau	Andrzej Ochocki	Feliks Wysocki
Piotr Guzowski	Józef Oleński	Dorota Wyszkowska
Mariusz Hamulczuk	Witold Orzeszko	Artur Zaborski
Magdalena Jaciow	Walenty Ostasiewicz	Anna Zalcewicz
Piotr Jeżowski	Izabela Ostoj	Małgorzata Zaleska
Alina Jędrzejczak	Anna Pajor	Henryk Zawadzki
Witold Jurek	Andrzej Paliński	Wojciech Zieliński
Paweł Kaczmarczyk	Tomasz Panek	Wojciech Ziętara
Ewa Kaznowska	Jan Paradysz	Maciej Żukowski
	Robert Perdał	

SPIS TREŚCI CONTENTS

Od redakcji	6
From the editorial team	
Zaproszenie do nadsyłania artykułów	8
Call for papers	
Studia metodologiczne	
Methodological studies	
Kamil Wilak	
Ocena skali bezrobocia biernego w Polsce	11
Estimation of the scale of inactive unemployment in Poland	
Statystyka w praktyce	
Statistics in practice	
Zofia Mielecka-Kubień, Andrzej Wójcik	
Zanieczyszczenie powietrza a stan zdrowia mieszkańców dużych miast województwa śląskiego	39
Air pollution and health condition of inhabitants of big cities in Śląskie Voivodship	
Agnieszka Chłoń-Domińczak, Aneta Ptak-Chmielewska	
Wykorzystanie danych rejestrowych do oceny sytuacji ekonomicznej przedsiębiorstw mikro, małych i średnich na przykładzie województwa kujawsko-pomorskiego	52
The use of register data to monitor micro, small and medium-sized enterprises' economic situation at the local level on the basis of the example of Kujawsko-Pomorskie Voivodship	
Edukacja statystyczna	
Statistical education	
Zbigniew Śleszyński	
Wyznaczanie współczynników korelacji liniowej – podstawy	69
The basics of determining the coefficients of a linear correlation	
Dyskusje. Recenzje. Informacje	
Discussions. Reviews. Information	
Justyna Gustyn	
Wydawnictwa GUS. Maj 2020	88
Publications of Statistics Poland. May 2020	
Dla autorów	90
For the authors	
Zakres tematyczny działów	99
Thematic scope of sections	

OD REDAKCJI

Przez kilka ostatnich miesięcy na łamach naszego czasopisma zachęcaliśmy autorów do podjęcia tematu wyzwań badawczych związanych ze stale rosnącym zapotrzebowaniem na dane statystyczne. W związku z upływającym terminem zgłoszeń jeszcze raz zapraszamy do nadsyłania prac.

W czerwcowym wydaniu publikujemy artykuły wpisujące się w rozważania i analizy z obszaru metodologii, praktyki i edukacji statystycznej.

Dr Kamil Wilak w artykule *Ocena skali bezrobocia biernego w Polsce* poświęca uwagę problematyce tytułowego zagadnienia oraz metodom badania tego zjawiska. Autor szacuje liczbę osób biernie bezrobotnych oraz ich udział wśród osób zarejestrowanych w powiatowych urzędach pracy na podstawie danych jednostkowych z Badania Aktywności Ekonomicznej Ludności za lata 2010–2018. W tym celu stosuje podejście kalibracyjne, z wykorzystaniem zmiennych pomocniczych wynikających z danych o bezrobociu rejestrowanym. Z badania wynika, że znaczna część osób zarejestrowanych jako bezrobotne charakteryzuje się biernością zawodową, przy czym poziom bezrobocia biernego jest różnicowany ze względu na płeć, wiek i poziom wykształcenia.

Zanieczyszczenie powietrza a stan zdrowia mieszkańców dużych miast województwa śląskiego to problem zgłębiany przez prof. dr hab. Zofię Mielecką-Kubień i dr. Andrzeja Wójcika. Autorzy, posługując się danymi z Głównego Inspektoratu Ochrony Środowiska oraz Głównego Urzędu Statystycznego, oceniają poziom wybranych rodzajów zanieczyszczeń powietrza w miastach woj. śląskiego liczących ponad 100 tys. mieszkańców. Zestawiają te oceny z niektórymi charakterystykami stanu zdrowia ludności tego regionu w latach 2014–2016 i stwierdzają, że w miastach o najwyższym poziomie zanieczyszczenia powietrza wartości mierników umieralności również były najwyższe, a wartości przeciętnego dalszego trwania życia noworodka – najniższe. Najgorszą sytuację pod względem obydwu czynników odnotowano w Chorzowie, Dąbrowie Górniczej i Rybniku, najlepszą zaś – w Bielsku-Białej i Tychach.

W pracy *Wykorzystanie danych rejestrowych do oceny sytuacji ekonomicznej mikro, małych i średnich przedsiębiorstw na przykładzie województwa kujawsko-pomorskiego* dr hab. Agnieszka Chłoń-Domińczak, prof. SGH, i dr hab. Aneta Ptak-Chmielewska, prof. SGH, badają, w jakim stopniu dane gromadzone w rejestrach publicznych w Polsce, w tym w rejestrach prowadzonych przez Zakład Ubezpieczeń Społecznych, mogą stanowić źródło wiedzy o stanie i rozwoju sektora mikro, małych i średnich przedsiębiorstw (MŚP) na poziomie lokalnym. Poddają analizie informacje dotyczące tego sektora w woj. kujawsko-pomorskim według stanu na grudzień 2018 r. Wynika z niej, że dane rejestrowe umożliwiają ocenę sytuacji ekonomicznej sektora MŚP na poziomie gmin, zgodnie z rekomendacjami Organizacji Współpracy Gospodarczej i Rozwoju. Autorki uwzględniają też wnioski wynikające z monitorowania tego sektora w Unii Europejskiej.

Dr Zbigniew Śleszyński w artykule *Wyznaczanie współczynników korelacji liniowej – podstawy* przedstawia podstawowe miary współzależności cech ilościowych, w tym współczynniki: korelacji liniowej Pearsona, korelacji wielowymiarowej, determinacji, korelacji cząstkowej i semikorelacji cząstkowej, zdefiniowane jako współczynniki korelacji liniowej odpowiednich wektorów otrzymanych na podstawie równań regresji. W ich wyznaczeniu autor wykorzystuje macierze brzegowe. Zastosowanie omawianych współczynników ilustruje na przykładzie modelu regresji wzrostu płac w Polsce w latach 2001–2019 zawierającego cztery zmienne objaśniające, szacowanego metodą najmniejszych kwadratów.

Czerwcowe wydanie „WS” kończy omówienie wybranych nowości wydawniczych GUS.
Zapraszamy do lektury.

FROM THE EDITORIAL TEAM

For the last few months we have been encouraging authors to undertake the topic of research challenges relating to the constantly growing demand for statistical data. As the deadline for submitting applications is nearing, once more we would like to invite you to send in your work.

The June issue of *Wiadomości Statystyczne. The Polish Statistician* contains articles reflecting on and analysing the issues from the realms of methodology, practice, and education in statistics.

The article opening the issue, *Estimation of the scale of inactive unemployment in Poland* by Kamil Wilak, PhD, DSc, focuses on the phenomenon of inactive unemployment and methods of researching it. The author estimates the number of inactive unemployed and their share in the amount of registered unemployed in the Poviats Labour Offices, on the basis of microdata obtained from the Labour Force Survey covering the years 2010–2018. The author uses a calibration approach that involves auxiliary variables relating to data on registered unemployment. The results indicate that a significant proportion of persons registered as unemployed are economically inactive, and the level of inactive unemployment varies depending on the sex, age and education of the survey participants.

Air pollution and health condition of inhabitants of big cities in Śląskie Voivodship is a topic researched by Zofia Mielecka-Kubieñ, PhD, DSc, ProfTit, and Andrzej Wójcik, PhD, DSc. The authors utilised data obtained from the Chief Inspectorate of Environmental Protection and Statistics Poland in order to assess the level of selected types of air pollution in cities of Śląskie Voivodship with over 100,000 inhabitants. They compare the assessments against selected reports on the health condition of the population of the region in the years 2014–2016 and come to the conclusion that in cities where air pollution reached the highest levels, mortality indicators were also the highest, while the indicators of life expectancy of newborn babies were the lowest. The worst situation relating to both air pollution and health condition was reported in Chorzów, Dąbrowa Górnicza and Rybnik, and the best was in Bielsko-Biała and Tychy.

Agnieszka Chłoń-Domińczak, PhD, DSc, Professor at the Warsaw School of Economics, and Aneta Ptak-Chmielewska, PhD, DSc, Professor at the Warsaw School of Economics, are authors of the study *The use of register data to monitor micro, small and medium-sized enterprises' economic situation at the local level on the basis of the example of Kujawsko-Pomorskie Voivodship*, in which they investigate as to what extent data collected in public registers in Poland, including the information compiled in the Social Insurance Institution (ZUS) register, can constitute a source of information about the situation and development of the micro, small and medium-sized enterprise (SME) sector at a local level. The authors analyse information relating to this sector from Kujawsko-Pomorskie Voivodship as of December 2018. The results of the analysis show that register data enables the assessment of the economic situation of the SME sector at a local (gmina) level, as per the OECD recommendations. Additionally, the study uses conclusions drawn from monitoring the SME sector in the EU.

Zbigniew Śleszyński, PhD, DSc, in his paper entitled *The basics of determining the coefficients of a linear correlation*, presents the basic measures related to the analysis of relationships between quantitative variables, including: the Pearson correlation coefficient, the multivariate correlation coefficient, coefficient of determination, partial correlation coefficient, and semi-partial correlation coefficient, defined as linear correlation coefficients of relevant vectors derived from regression equations. The author applies bordered matrices to the calculations. The application of the discussed variables is illustrated in the model of regression of salary growth in Poland in the years 2001–2019 with four covariates, estimated using the least squares method.

The June issue of *WS* closes with a description of selected new releases of Statistics Poland publications.

We wish you pleasant reading.

ZAPROSZENIE DO NADSYŁANIA ARTYKUŁÓW

Zbliżające się Powszechny Spis Rolny 2020 oraz Narodowy Spis Powszechny Ludności i Mieszkań 2021 dają asumpt do poświęcenia szczególnej uwagi na łamach naszego czasopisma tym złożonym przedsięwzięciom oraz przyjrzenia się im w wymiarze metodologicznym, analitycznym i organizacyjnym. Stanowią również istotny impuls do podjęcia w szerszym kontekście tematu wyzwań badawczych związanych ze stale rosnącym zapotrzebowaniem na dane statystyczne.

Jedno z głównych wyzwań tego rodzaju dotyczy skutecznego zapewnienia ochrony tajemnicy statystycznej. Dane gromadzone w trakcie badań statystycznych czy ujmowane w rejestrach administracyjnych i stamtąd pozyskiwane zawierają wiele informacji dotyczących indywidualnych cech jednostek, w tym ich bezpośrednich identyfikatorów. W dzisiejszych realiach zwykle usuwanie takich elementów przed udostępnieniem wyników informacji statystycznych w celu wyeliminowania – lub przynajmniej maksymalnego zminimalizowania – ryzyka ujawnienia bądź odtworzenia przez użytkowników udostępnianych zasobów danych wrażliwych identyfikujących jednostkę statystyczną (np. respondenta) to o wiele za mało, a utrwalone w wieloletniej praktyce proste reguły ukrywania danych wrażliwych okazują się dalece niewystarczające, zwłaszcza w przypadku mikrodanych (tzn. odpersonalizowanych danych jednostkowych) czy wielowymiarowych kostek danych OLAP. Ponadto potencjalny użytkownik może dysponować również innymi, niezależnymi zasobami danych, które mogą tę identyfikację ułatwić. Co więcej, z uwagi na obowiązujące regulacje prawne dane teleadresowe podmiotów gospodarczych (także osób fizycznych prowadzących działalność gospodarczą) są jawne, a zatem w ich przypadku ochrona danych wrażliwych musi być zapewniona w inny sposób. Konieczne jest także zachowanie równowagi między minimalizacją ryzyka ujawnienia informacji wrażliwych a minimalizacją straty informacji powstającej na skutek stosowanych procedur. Stąd potrzeba ciągłego doskonalenia metodologii kontroli ujawniania danych wykorzystującej zaawansowane narzędzia probabilistyczne i ekonometryczne na podstawie doświadczeń nabywanych w tym zakresie.

Wobec znacznej liczby odmów uczestnictwa w badaniu bądź nieprawidłowo lub niekompletnie podanych informacji ważny problem stanowi też obciążenie odpowiedzi w badaniach statystycznych. Niezbędne jest więc doskonalenie narzędzi oceny skali obciążeń odpowiedzi i ich wpływu na jakość danych wynikowych z badania, co w znacznej mierze ułatwi ocenę efektywności badania i jego planowanie w kolejnych okresach. Luki w danych mogą być wypełnione dzięki imputacji prowadzonej różnymi metodami (np. w sposób losowy, oparty na zapożyczaniu danych lub na modelach ekonometrycznych). Rozwój narzędzi imputacji oraz optymalizacja ich zastosowania w konkretnych przypadkach, a także ocena wpływu imputacji na jakość informacji wynikowych staje się zatem nieodzownym elementem działań statystycznych.

Zachęcamy do zgłaszania artykułów, które wniosą wkład w rozwój badań nad wymienionymi zagadnieniami, dostarczając zarówno nowych, jak i udoskonalonych rozwiązań metodologicznych oraz praktycznych. Prace w języku polskim lub angielskim należy przysyłać na adres redakcja.ws@stat.gov.pl do **30 czerwca 2020 r.** Prace, które zostaną zgłoszone jako artykuły naukowe, będą poddane recenzji. Prosimy o zapoznanie się ze wskazówkami dla autorów oraz uwzględnienie wymogów redakcyjnych podanych w rubryce Dla autorów (s. 90) i na stronie internetowej czasopisma ws.stat.gov.pl/ForAuthors.

CALL FOR PAPERS

The upcoming Agricultural Census 2020 and the National Population and Housing Census 2021 motivate researchers to devote special attention to these complex undertakings and to examine them in relation to methodology, analysis and organisation. They are also an impulse to tackle, in a wider scope, research challenges connected to the constantly growing demand for statistical data.

One of the main such challenges is the protection of statistical confidentiality. Data collected in statistical surveys or stored in administrative registers, and then drawn from there, contain abundant information on personal attributes of individuals, including direct identifiers. Nowadays, the mere deletion of such elements before disclosing statistical information in order to either fully eliminate, or to minimise the risk of revealing or recreating sensitive data identifying a given statistical unit (e.g. a survey respondent) by the users of the available data sets, proves inadequate. By the same token, other simple methods of suppressing sensitive data, which have been firmly established over many years' practice, turn out to be by far insufficient, especially when it comes to microdata (i.e. depersonalised unit data) or multidimensional OLAP cubes. It also cannot be ruled out that a potential data user has access to other data sets, independent from the set in question, that might help identify the units on the basis of which the data set in question has been created. In addition, according to the current legal regulations, contact details of companies (and sole trader businesses) are public, therefore in their case, the protection of sensitive data has to be ensured in different ways. Last but not least, it is crucial to keep balance between minimising the risk of disclosing sensitive data and minimising the loss of data that results from the adopted procedures. All the above necessitates the constant improvement of the methodology used for supervising disclosure of data, which utilises advanced probabilistic and econometric tools, to the extent indicated by experiences in this field.

Given large numbers of refusals to take part in surveys or frequent provision of incomplete or incorrect data, the problem of response burden in statistical surveys is aggravating. This creates the need to improve tools for measuring its scale and influence on the quality of data yielded by surveys. By doing so, it will be possible to facilitate the evaluation of surveys' effectiveness and additionally to make the process of planning their continuation in subsequent periods much easier. The gaps in data sets can be filled by means of imputation carried out according to various methods (e.g. in a random way, based on borrowing of data, or following econometric models). The development of imputation tools and the optimization of their application to concrete cases, as well as the assessment of their influence on the quality of the result data, have thus become indispensable elements of statistical activity.

In this connection, we would like to warmly encourage interested researchers to submit papers that could contribute to the research into the above-mentioned issues, especially by, providing new methodological and practical tools, or presenting improved versions of the already-existing ones. Works in Polish or English should be sent to: redakcja.ws@stat.gov.pl by **30 June 2020**. Papers submitted as scientific articles will be subject to the reviewing process. We would like to kindly refer you to the guidelines for authors, and to recommend that you follow the editorial requirements, both of which are available on our journal's web page: ws.stat.gov.pl/ForAuthors.

Ocena skali bezrobocia biernego w Polsce

Kamil Wilak^a

Streszczenie. Celem badania opisanego w artykule jest oszacowanie poziomu bezrobocia biernego w Polsce. Estymacji podlegała liczba osób biernie bezrobotnych oraz ich udział wśród osób zarejestrowanych w powiatowych urzędach pracy. Szacunki przeprowadzono na podstawie danych jednostkowych z Badania Aktywności Ekonomicznej Ludności za lata 2010–2018. Zastosowano podejście kalibracyjne z uwzględnieniem zmiennych pomocniczych dotyczących bezrobocia rejestrowanego.

Otrzymane wyniki wskazują, że znaczna część osób zarejestrowanych jako bezrobotne charakteryzuje się biernością zawodową. W badanym okresie odsetek biernie bezrobotnych wynosił od 29,8% do 53,3%, przy czym od 2014 r. widoczny jest trend rosnący tego wskaźnika. Zaobserwowano także zróżnicowanie poziomu bezrobocia biernego ze względu na najważniejsze cechy demograficzne, tj. płeć, wiek i poziom wykształcenia. Największe różnice występują w przypadku płci – kobiety zarejestrowane jako bezrobotne znacznie częściej niż mężczyźni cechują się biernością zawodową.

Słowa kluczowe: bezrobocie bierne, bezrobocie rejestrowane, Badanie Aktywności Ekonomicznej Ludności, podejście kalibracyjne

JEL: E24, C13

Estimation of the scale of inactive unemployment in Poland

Abstract. The aim of the study described in the paper is to estimate the level of inactive unemployment in Poland. This required the estimation of the number of inactive unemployed and their percentage in the total number of persons registered as unemployed. The estimation was based on unit-level data from the Polish Labour Force Survey (2010–2018). A calibration approach was applied that involved auxiliary variables relating to registered unemployment.

The results indicate that a significant proportion of persons registered as unemployed are economically inactive. The percentage of inactive unemployed ranged from 29.8 to 53.3 percent over the studied period, with an upward trend observed since 2014. It was also demonstrated that the level of inactive unemployment varies among groups defined by basic demographic indicators, i.e. sex, age and education. The most pronounced differences can be observed between groups determined by sex – women registered as unemployed are economically inactive much more often than men.

Keywords: inactive unemployment, registered unemployment, Labour Force Survey, calibration approach

^a Uniwersytet Ekonomiczny w Poznaniu, Instytut Informatyki i Ekonomii Ilościowej.
ORCID: <https://orcid.org/0000-0002-4305-6202>.

1. Wprowadzenie

W powiatowych urzędach pracy (PUP) jako bezrobotne mogą rejestrować się osoby niezatrudnione i niewykonujące pracy zarobkowej, które poszukują pracy, są gotowe i zdolne do jej podjęcia oraz spełniają formalne warunki określone w ustawie o promocji zatrudnienia i instytucjach rynku pracy¹. Powszechnie uważa się jednak, że nie wszyscy zarejestrowani jako bezrobotni charakteryzują się tymi cechami. Część nie wykazuje zainteresowania znalezieniem pracy, a powodem rejestracji jest chęć korzystania ze świadczeń dla bezrobotnych. Warunków statusu osoby bezrobotnej nie spełniają również ci, którzy podejmują pracę, często w szarej strefie. Tego rodzaju zjawisko nazywane jest bezrobociem pozornym (a także fikcyjnym lub symulowanym) (Czapiński i Panek, 2013; Kwiatkowski, 2013). W literaturze pojęcie bezrobocia pozornego obejmuje też zjawisko polegające na rejestracji w powiatowych urzędach pracy osób, które nie chcą pracować albo nie są gotowe podjąć pracy (Wojciechowska, 2006). Ze względu na odmienny stosunek do pracy Wilak (2018) określa tę grupę mianem osób biernie bezrobotnych.

Informacje na temat bezrobocia biernego mogą być pomocne w wypracowaniu rozwiązań stosowanych w polityce społecznej wobec bezrobotnych. Ze względu na duże zróżnicowanie zjawiska bezrobocia w Polsce – zarówno przestrzenne, jak i pomiędzy różnymi grupami osób określonymi przez cechy demograficzne i społeczne – ważne jest, aby informacja na temat charakteru bezrobocia była dostępna nie tylko dla całej populacji, lecz także na niższym stopniu agregacji. Tymczasem opracowań dotyczących badań poziomu bezrobocia biernego w Polsce jest niewiele, a publikowane szacunki często są zagregowane w wysokim stopniu.

Celem badania omawianego w artykule jest oszacowanie poziomu bezrobocia biernego w Polsce. Estymacji podlegała liczba osób biernie bezrobotnych oraz ich udział wśród osób zarejestrowanych w PUP. Szacunki przeprowadzono ogółem dla całej populacji bezrobotnych, a także w podziale na domeny określone przez: płeć (kobiety, mężczyźni), wiek (18–24, 25–34, 35–44, 45–54, 55 lat i więcej), miejsce zamieszkania (miasto, wieś) i poziom wykształcenia (co najwyżej gimnazjalne, zasadnicze zawodowe, średnie, wyższe). Przeprowadzona analiza obejmowała lata 2010–2018.

2. Problematyka bezrobocia biernego

2.1. Bezrobocie w polskim systemie prawnym

Pojęcie *bezrobocie* w polskim systemie prawnym funkcjonuje od rozpoczęcia w 1989 r. transformacji systemowej². Od tego czasu uchwalono cztery ustawy, które definiowały rolę państwa w zwalczaniu tego zjawiska i przeciwdziałaniu jego skutkom. Obec-

¹ Ustawa z dnia 20 kwietnia 2004 r. o promocji zatrudnienia i instytucjach rynku pracy (Dz.U. 2004 nr 99, poz. 1001).

² Na początku lat 90. ubiegłego wieku bezrobocie było nowym zjawiskiem na polskim rynku pracy. W tym czasie powstało wiele opracowań poświęconych temu problemowi oraz polityce rynku pracy w okresie transformacji systemowej. Wśród nich można wymienić opracowania Góry, Sochy i Sztanderskiej (1995), Kotowskiej i Strzeleckiego (1993), Kryńskiej (1993), Kwiatkowskiego (1992) oraz Witkowskiego (1994).

nie obowiązuje Ustawa z dnia 20 kwietnia 2004 r. o promocji zatrudnienia i instytucjach rynku pracy (Dz.U. 2004 nr 99, poz. 1001)³. Zgodnie z jej zapisami poprzez rejestrację w PUP właściwym ze względu na zameldowanie lub faktyczne miejsce zamieszkania status bezrobotnego mogą otrzymać osoby pełnoletnie, które nie osiągnęły wieku emerytalnego⁴. Niezbędne jest w tym celu spełnienie warunków, które można podzielić na dwie grupy. Warunki z pierwszej grupy określają stosunek do pracy osób ubiegających się o przyznanie statusu osoby bezrobotnej. Są to:

- brak zatrudnienia i niewykonywanie pracy zarobkowej;
- zdolność i gotowość do podjęcia zatrudnienia albo innej pracy zarobkowej;
- poszukiwanie pracy.

Pierwszy warunek wyklucza możliwość zarejestrowania się jako bezrobotne przez te osoby, które wykonują pracę na podstawie stosunku pracy, stosunku służbowego lub umowy o pracę nakładczą, oraz te, które wykonują lub świadczą usługi na podstawie umów cywilnoprawnych, w tym umowy agencyjnej, umowy-zlecenia czy umowy o dzieło. Należy zwrócić uwagę, że w ustawie nie wprowadzono wyjątków dla niepełnego wymiaru czasu pracy czy też jej doraźnego charakteru, który często występuje w przypadku umów cywilnoprawnych – podjęcie jakiegokolwiek pracy wiąże się z utratą statusu osoby bezrobotnej. Może to zniechęcać do podejmowania prac o charakterze krótkotrwałym lub w niepełnym wymiarze czasu, a także sprzyjać pracy w szarej strefie.

Drugi warunek określa wymiar czasu pracy, w jakim osoby ubiegające się o status bezrobotnego powinny być gotowe i zdolne podjąć pracę. W przypadku zatrudnienia ustawodawca zastrzega, że zdolność i gotowość do jego podjęcia dotyczy pełnego wymiaru czasu pracy obowiązującego w danym zawodzie albo służbie. Wyjątkiem są osoby niepełnosprawne, które powinny cechować się zdolnością i gotowością do zatrudnienia w co najmniej połowie wymiaru czasu pracy. W przypadku innej pracy zarobkowej ustawodawca nie precyzuje wymiaru czasu, w jakim osoba bezrobotna powinna być gotowa i zdolna do jej podjęcia. Z powyższego opisu wynika, że osoba zdolna i gotowa podjąć jakąkolwiek pracę dorywczą, bez względu na jej wymiar, również może być uznana za osobę bezrobotną (Góral, 2011).

Pojęcia *zdolność* i *gotowość* nie zostały zdefiniowane w ustawie. Staszewska (2012) rozpatruje zdolność w dwóch płaszczyznach: prawnej i biologicznej. Pierwsza dotyczy prawnej zdolności do bycia pracownikiem w rozumieniu kodeksu pracy bądź stroną umowy cywilnoprawnej, druga – stanu zdrowia, który pozwala na podjęcie pracy. Oznacza to, że osoba bezrobotna nie może mieć stwierdzonej zaświadczeniem lekarskim niezdolności do wykonywania pracy. Natomiast przez gotowość należy rozumieć rzeczywistą wolę wykonywania pracy i znajdowanie się w sytuacji umożliwiającej jej wykonywanie, czego potwierdzeniem jest pozostawanie do dyspozycji PUP. Obowiązkiem bezrobotnego jest zgłaszanie się do PUP w wyznaczonym termi-

³ W dalszej części pracy określana jako ustawa o promocji zatrudnienia i instytucjach rynku pracy.

⁴ Do grudnia 2012 r. wiek emerytalny wynosił 65 lat dla mężczyzn i 60 lat dla kobiet. Od stycznia 2013 r. na skutek reformy systemu emerytalnego nastąpił stopniowy wzrost wieku emerytalnego – docelowo miał wynosić 67 lat (w przypadku mężczyzn w 2020 r., a w przypadku kobiet w 2040 r.). W październiku 2017 r. przywrócono wiek emerytalny sprzed reformy: 65 lat dla mężczyzn i 60 lat dla kobiet.

nie w celu potwierdzenia gotowości do podjęcia pracy bądź przyjęcia propozycji odpowiedniej pracy lub innej formy pomocy oferowanej przez urząd.

Ustawodawca nie zdefiniował również określenia *poszukiwanie pracy*. Nie określił form tej aktywności ani jej intensywności. Niemniej jednak pojęcie to można wiązać z chęcią i determinacją do znalezienia i podjęcia pracy, a im są one większe, tym aktywniej osoba bez pracy jej poszukuje.

Druga grupa obejmuje warunki o charakterze formalnym. Ich celem jest zawężenie podmiotowego zakresu pojęcia bezrobotnego, a w związku z tym ograniczenie bezrobocia w rozumieniu prawnym. Warunki formalne to:

- niepobieranie nauki;
- brak źródeł dochodu;
- niepobieranie określonych świadczeń.

Pierwszy z tych wymogów wyłącza z grupy bezrobotnych m.in. wszystkie osoby pobierające naukę, z wyjątkiem osób uczących się w szkołach dla dorosłych lub szkołach wyższych w systemie wieczorowym albo zaocznym. Pozostałe warunki dotyczą głównie braku dochodu po stronie bezrobotnego. Osiągnięcie przychodu lub sama możliwość osiągnięcia go wyklucza uzyskanie statusu osoby bezrobotnej. W tym kontekście uznaje się, że przychód osiągają m.in. właściciele lub posiadacze nieruchomości rolnej o powierzchni użytków rolnych przekraczającej 2 ha przeliczeniowe. Warunkiem posiadania statusu osoby bezrobotnej jest również niepobieranie określonych przez ustawodawcę świadczeń, np. zasiłku chorobowego lub macierzyńskiego. W przypadku niektórych, takich jak emerytura czy renta z tytułu niezdolności do pracy, sam fakt nabycia prawa do ich pobierania wyklucza możliwość posiadania statusu osoby bezrobotnej, a więc nie ma znaczenia, czy te świadczenia są faktycznie pobierane.

2.2. Status na rynku pracy według rekomendacji Międzynarodowej Organizacji Pracy

Źródłem definicji stosowanych w statystyce polskiego rynku pracy, oprócz ustawy o promocji zatrudnienia i instytucjach rynku pracy, są również akty prawne Unii Europejskiej (UE) dotyczące badania siły roboczej. Opierają się one na rekomendacjach Międzynarodowej Organizacji Pracy (MOP), które zostały określone w *Rezolucji dotyczącej definiowania aktywnych zawodowo, pracujących, bezrobotnych oraz niepełnozatrudnionych* przyjętej na XIII Międzynarodowej Konferencji Statystyków Pracy (International Conference of Labour Statisticians) w październiku 1982 r.⁵ Zapisy rezolucji szczegółowo objaśnili Hussmanns, Mehran i Verma (1992).

Definicje rekomendowane przez MOP zostały doprecyzowane w aktach prawnych Komisji Europejskiej (KE). Główne regulacje w zakresie definicji osoby bezrobotnej zawiera rozporządzenie Komisji (WE) nr 1897/2000. Poniżej przytoczono definicje stosowane w badaniach siły roboczej krajów członkowskich UE, w tym w Polsce w Badaniu Aktywności Ekonomicznej Ludności (BAEL).

⁵ Do dzisiaj definicje te uległy tylko niewielkim zmianom, które zostały ujęte w rezolucjach przyjętych na XVI i XVII ICLS.

Do pracujących zalicza się osoby w wieku 15 lat i więcej, które w badanym tygodniu wykonywały przez co najmniej jedną godzinę pracę przynoszącą zarobek lub dochód, tzn. były zatrudnione w charakterze pracownika najemnego, pracowały we własnym lub w dzierżawionym gospodarstwie rolnym, prowadziły własną działalność gospodarczą poza rolnictwem, pomagały (bez wynagrodzenia) w prowadzeniu rodzinnego gospodarstwa rolnego lub rodzinnej działalności gospodarczej poza rolnictwem. Definicja pracującego obejmuje również te osoby, które miały pracę, ale jej nie wykonywały z powodu choroby, urlopu macierzyńskiego lub wypoczynkowego, a także z innych powodów, przy czym w okresie niewykonywania pracy otrzymywały co najmniej 50% dotychczasowego wynagrodzenia lub czas niewykonywania pracy był krótszy niż trzy miesiące. Do pracujących zaliczani są również uczniowie, z którymi zakłady pracy lub osoby fizyczne zawarły umowę o naukę zawodu lub przyuczenie do określonej pracy, o ile otrzymywali wynagrodzenie.

Za bezrobotne uznaje się osoby wieku 15–74 lat, które spełniły jednocześnie trzy warunki: (1) w badanym tygodniu nie były osobami pracującymi (według powyższych kryteriów), (2) aktywnie poszukiwały pracy, tzn. w ciągu czterech tygodni (wliczając jako ostatni tydzień badany) podjęły konkretne działania, aby znaleźć pracę, oraz (3) były gotowe (zdolne) podjąć pracę w ciągu dwóch tygodni następujących po badanym tygodniu. Do bezrobotnych zalicza się także osoby, które nie poszukiwały pracy, ponieważ już ją znalazły i oczekiwały na jej rozpoczęcie przez okres nie dłuższy niż trzy miesiące oraz były gotowe tę pracę podjąć.

Natomiast biernymi zawodowo są osoby w wieku 15 lat i więcej niezaklasyfikowane jako pracujące lub bezrobotne, a więc osoby (1) niepracujące (według powyższych kryteriów), które w badanym tygodniu (2) nie poszukiwały pracy lub (3) nie były zdolne (gotowe) do jej podjęcia w ciągu dwóch tygodni następujących po tygodniu badanym. Do biernych zawodowo zalicza się także te osoby bez pracy, które jej nie poszukiwały, ponieważ już ją znalazły i oczekiwały na jej rozpoczęcie w okresie do trzech miesięcy, ale nie były gotowe jej podjąć. Wśród biernych zawodowo wyróżniono grupę osób zniechęconych, które nie poszukują pracy, ponieważ są przekonane, że jej nie znajdą.

2.3. Źródła danych na temat bezrobocia w Polsce

Jednym z głównych źródeł danych o bezrobociu w Polsce jest rejestr osób bezrobotnych⁶. Badanie bezrobocia rejestrowanego jest prowadzone przez Główny Urząd Statystyczny (GUS) wspólnie z Ministerstwem Rodziny, Pracy i Polityki Społecznej (MRPiPS). Podstawę prawną stanowi coroczne rozporządzenie Rady Ministrów w sprawie programu badań statystycznych statystyki publicznej. Staty-

⁶ Przegląd źródeł danych na temat rynku pracy wykorzystywanych przez polską statystykę publiczną można znaleźć w opracowaniu GUS (2018b).

styki bezrobocia rejestrowanego bazują na definicji osoby bezrobotnej z ustawy o promocji zatrudnienia i instytucjach rynku pracy. Dane zbiera się na podstawie trzech sprawozdań sporządzanych przez każdy PUP: MPiPS-01 wraz z załącznikami, MPiPS-02 i MPiPS-07⁷. Informacje na temat bezrobocia rejestrowanego publikowane są przez GUS m.in. w kwartalnych opracowaniach *Bezrobocie rejestrowane* (np. GUS, 2019b).

Drugim źródłem wiedzy na temat bezrobocia w Polsce jest reprezentacyjne badanie BAEL (ang. Labour Force Survey, LFS), prowadzone przez GUS co kwartał od 1992 r.⁸ Ma ono na celu rozpoznanie bieżącej sytuacji na polskim rynku pracy oraz zachodzących na nim zmian. Podstawę prawną do realizacji tego badania – podobnie jak w przypadku bezrobocia rejestrowanego – stanowi coroczne rozporządzenie Rady Ministrów w sprawie programu badań statystycznych statystyki publicznej. Podstawowym aktem prawnym dotyczącym LFS w krajach UE jest Rozporządzenie Rady Unii Europejskiej nr 577/98 z dnia 9 marca 1998 r.⁹

Metodyka BAEL opiera się na definicjach rekomendowanych przez MOP, które pozwalają podzielić badane osoby na trzy rozłączne grupy: pracujących, bezrobotnych i biernych zawodowo. Status respondenta jest określany na podstawie udzielonych przez niego odpowiedzi na pytania zawarte w ankiecie. Reprezentacyjny charakter BAEL pozwala na uogólnianie jego wyników na całą populację ludności Polski w wieku 15 lat i więcej¹⁰. Informacje na temat gospodarstw domowych i ich członków są zbierane przez ankietatorów przy wykorzystaniu dwóch kwestionariuszy: ZG i ZD. Za pomocą kwestionariusza ZG zdobywa się podstawowe informacje na temat wylosowanych do próby gospodarstw domowych i ich członków, ankieta ZD zawiera zaś pytania skierowane do poszczególnych członków tych gospodarstw (w wieku 15 lat i więcej) dotyczące ich aktywności na rynku pracy. Szczegółowe informacje na temat metodyki BAEL można znaleźć w opracowaniu GUS (2018a). Zmiany zachodzące w metodyce od początku prowadzenia badania, które mogą mieć znaczenie przy analizowaniu szeregów czasowych i porównywalności wyników, zostały wypunktowane przez Zgierską (2017). Wyniki badania publikowane są m.in. w kwartalnych opracowaniach *Aktywność ekonomiczna ludności Polski* (np. GUS, 2019a).

⁷ Wzory formularzy są ustalane w corocznym rozporządzeniu Rady Ministrów w sprawie określenia wzorów formularzy sprawozdawczych, objaśnień co do sposobu ich wypełniania oraz wzorów kwestionariuszy i ankiet statystycznych stosowanych w badaniach statystycznych ustalonych w programie badań statystycznych statystyki publicznej.

⁸ Z wyjątkiem II i III kwartału 1999 r., w których badanie nie zostało zrealizowane ze względu na wprowadzane zmiany metodologiczne.

⁹ Wszystkie akty prawne, które mają wpływ na kształt BAEL, zostały wymienione w załączniku nr 1 do opracowania GUS (2018a).

¹⁰ Z populacji wyłączone są osoby przebywające za granicą powyżej 12 miesięcy (do II kwartału 2012 r. – co najmniej trzy miesiące) oraz osoby mieszkające w obiektach zbiorowego zakwaterowania, takich jak hotele pracownicze, akademiki czy koszary wojskowe.

Przy porównywaniu obu wymienionych źródeł informacji na temat bezrobocia w Polsce należy mieć na uwadze ich odmiennosc. Rejestr bezrobotnych służy do ewidencji osób, które z powodu bezrobocia zgłaszają się po pomoc do PUP. Natomiast BAEL ma na celu pokazanie faktycznej sytuacji na rynku pracy i zachodzących na nim zmian.

Poziom bezrobocia rejestrowanego zależy od obowiązującego prawa, którego regulacje wpływają na wielkość populacji osób uprawnionych do posiadania statusu bezrobotnego, a także na skłonność do rejestracji w PUP zarówno przez osoby uprawnione, jak i nieuprawnione. Zmniejszenie korzyści przysługujących osobom zarejestrowanym jako bezrobotne bądź utrudnienie dostępu do tych korzyści może powodować odpływ osób z bezrobocia rejestrowanego. Jak pisał Janukowicz (2010):

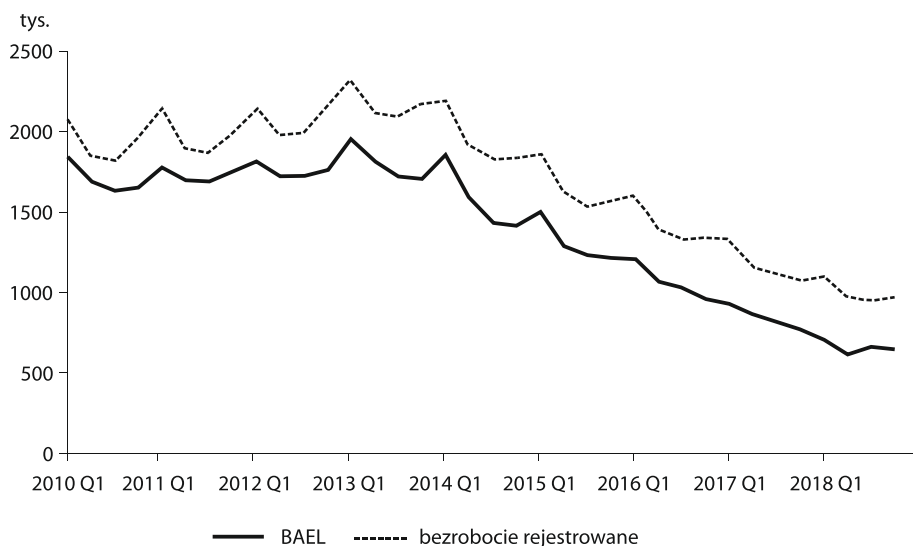
Liczba bezrobotnych ujęta w statystyce urzędów pracy w dużej mierze zależy od stopnia motywacji do zarejestrowania się jako osoby bezrobotne. Niechęć do rejestracji w urzędzie, spowodowana na przykład znikomymi szansami na znalezienie pracy bądź też brakiem otrzymania uprawnień do pobierania zasiłku, uniemożliwia w rzeczywistości poznanie prawdziwej skali problemu.

Zjawisko polegające na nierejestrowaniu się w PUP przez ludzi spełniających warunki uzyskania statusu osoby bezrobotnej jest określone w literaturze jako bezrobocie nierejestrowane (Kwiatkowski, 2005). Z kolei regulacje prawne korzystne dla osób ubiegających się o pomoc PUP mogą powodować napływ do bezrobocia rejestrowanego także osób niespełniających warunków uzyskania statusu bezrobotnego, tj. pracujących w szarej strefie bądź niezainteresowanych podjęciem pracy.

Klasyfikacja do grona bezrobotnych według metodyki BAEL nie jest uzależniona od figurowania w rejestrach administracyjnych. Osoby bez pracy, które jej poszukują oraz są gotowe i zdolne ją podjąć, ale nie rejestrują się w PUP jako bezrobotne, w badaniu bezrobocia rejestrowanego nie są zaliczane do grona bezrobotnych, podczas gdy w BAEL zostają przypisane właśnie do tej grupy. Inne różnice pomiędzy bezrobociem rejestrowanym a ustalonym na podstawie BAEL to m.in. odmienne definicje osoby bezrobotnej, różne metody pomiaru, a także różne okresy referencyjne¹¹.

W latach 2010–2018 poziom bezrobocia rejestrowanego znacznie różnił się od poziomu bezrobocia oszacowanego na podstawie BAEL (wykr. 1). Kwartalna liczba bezrobotnych wyznaczona na podstawie rejestru osób bezrobotnych niezmiennie przewyższała szacunki z BAEL, przeciętnie o 318,4 tys. osób. Natomiast tendencje były podobne – od 2013 r. obserwowano systematyczny spadek liczby bezrobotnych.

¹¹ Dane z badania bezrobocia rejestrowanego prezentowane są według stanu na koniec okresu (miesiąc/kwartal/roku), natomiast dane z BAEL – z uwagi na równomierne rozłożenie próby badawczej w kwartale – prezentowane są jako przeciętna w kwartale (ew. przeciętna w roku).

Wykr. 1. Bezrobotni w Polsce

Uwaga. Q1 – pierwszy kwartał.

Źródło: opracowanie własne na podstawie wyników BAEL oraz zagregowanych danych ze sprawozdań MPIPS-01.

Przyczyn wyższego poziomu bezrobocia rejestrowanego można doszukiwać się – oprócz wymienionych różnic pomiędzy źródłami danych – w niespełnianiu warunków statusu bezrobotnego przez osoby widniejące w ewidencji PUP. Powszechnie uważa się, że wśród osób zarejestrowanych jako bezrobotne znajdują się takie, które pracują w szarej strefie bądź nie są zainteresowane podjęciem pracy. Według metodyki BAEL osoby te zasilają grono, odpowiednio, pracujących i biernych zawodowo.

2.4. Wpływ polityki społecznej wobec bezrobotnych na rynek pracy

Ważnym obszarem podlegającym ustawom regulującym politykę rynku pracy i politykę społeczną są działania polegające na udzielaniu pomocy osobom bezrobotnym i ich rodzinom; celem tych działań jest m.in. zmniejszenie ekonomicznej uciążliwości braku dochodu. Służy temu np. wypłacanie zasiłku dla bezrobotnych. Do pobierania tego świadczenia uprawnione są osoby zarejestrowane jako bezrobotne we właściwym PUP, dla których urząd nie ma propozycji odpowiedniej pracy lub innej formy aktywizacji zawodowej. Istotne są również staż pracy bezrobotnego oraz wysokość jego wynagrodzenia w okresie poprzedzającym rejestrację. Szczegółowe warunki uzyskania tego świadczenia, jego wysokość, jak również okresy jego pobierania są określone w ustawie o promocji zatrudnienia i instytucjach rynku pracy. Zapisy w ustawie uzupełniono o regulacje wynikające z rozpo-

rządzenia ministra właściwego ds. pracy w sprawie szczegółowego trybu przyznawania zasiłku dla bezrobotnych, dodatku szkoleniowego, stypendium i dodatku aktywizacyjnego.

Zmiany prawa w tym zakresie w różnym stopniu motywowały osoby pozostające bez pracy do rejestracji w PUP. W badanym okresie zasiłku dla bezrobotnych nie można jednak uznać za główną przyczynę rejestracji – odsetek osób pobierających to świadczenie wśród zarejestrowanych w latach 2010–2018 kształtował się między 12,6% a 20,4%.

Innym przywilejem przysługującym bezrobotnym jest darmowe ubezpieczenie zdrowotne. Zgodnie z Ustawą z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych¹² osoby zarejestrowane w PUP jako bezrobotne i ich rodziny są zgłaszane przez te urzędy do ubezpieczenia zdrowotnego. W ostatnich latach trwa dyskusja na temat słuszności tego rozwiązania (Popiołek i Piątkowska, 2013; Rzemek, 2014; Topolska, 2014). Zauważa się, że znaczna część osób rejestrujących się jako bezrobotne nie kieruje się chęcią znalezienia pracy, lecz chęcią uzyskania ubezpieczenia zdrowotnego, w związku z czym pracownicy PUP zajmują się obsługą osób niezainteresowanych podjęciem pracy.

Kolejna korzyść to możliwość otrzymania okresowych świadczeń z jednostek organizacyjnych pomocy społecznej. Bezrobocie jest, oprócz ubóstwa, najczęstszym powodem przyznawania takich świadczeń¹³. Osoby ubiegające się o świadczenia okresowe z tytułu bezrobocia zobowiązane są do przedstawienia dokumentu poświadczającego rejestrację w PUP, mimo że taki wymóg nie został przewidziany w ustawie o pomocy społecznej (Hryniewicka, 2011).

Świadczenia socjalne z tytułu bezrobocia spełniają dwie podstawowe funkcje: dochodową i motywacyjną. Zapewnienie dochodu w okresie bezrobocia ma w pewnym stopniu zrekompensować utratę wynagrodzenia wynikającą z utraty zatrudnienia i zabezpiecza przed zbyt drastycznym obniżeniem standardu życia z powodu braku pracy. Funkcja ta jest spełniona tym lepiej, im wyższa jest stopa zastąpienia, czyli stosunek zasiłku do wcześniej otrzymywanego wynagrodzenia. Natomiast realizacja funkcji motywacyjnej, polegającej na zachęceniu bezrobotnego do aktywnego poszukiwania pracy, jest efektywna w sytuacji stosunkowo dużej uciążliwości ekonomicznej bezrobocia, a więc gdy stopa zastąpienia pozostaje na dość niskim poziomie. Pomiędzy funkcją dochodową a motywacyjną istnieje zatem pewna wymienność, tj. realizacja jednej funkcji może stać w sprzeczności ze spełnieniem drugiej.

¹² Dz.U. 2004 nr 210, poz. 2135.

¹³ Informacje o liczbie rodzin i liczbie osób w rodzinach, którym przyznano świadczenia społeczne, dostępne są na stronie www.gov.pl/web/rodzina/statystyki-pomocy-spoecznej (dane zagregowane ze sprawozdań MRPiPS-03).

Świadczenia socjalne dla bezrobotnych niewątpliwie oddziałują na rynek pracy, w tym na wielkość zatrudnienia, bezrobocia i płac, poprzez wpływ na postępowanie i decyzje podmiotów związanych z rynkiem pracy: pracodawców, pracujących, bezrobotnych i biernych zawodowo (Kwiatkowski, 2013). Podstawowe pytanie stawiane w naukach ekonomicznych w kwestii świadczeń dla bezrobotnych dotyczy tego, czy motywują one do szukania pracy, czy też przeciwnie – sprawiają, że bezrobotni niezbyt pręźnie starają się o znalezienie nowego zatrudnienia.

Stawiana w literaturze przedmiotu hipoteza głosi, że zasiłki osłabiają motywację do aktywnego poszukiwania pracy i wydłużają okres bezrobocia (Kwiatkowski, 2013). W skrajnych przypadkach bezrobotni usatysfakcjonowani wysokością dochodu ze świadczeń dla bezrobotnych mogą całkowicie zaprzestać szukania pracy, godząc się na trwałe pozostawanie bez zatrudnienia. Utrwalenie biernej postawy wśród bezrobotnych prowadzi do tego, że świadczenia stają się dla nich głównym źródłem utrzymania, a bezrobocie – sposobem na życie. Uzależnienie się bezrobotnych i ich rodzin od pomocy z zewnątrz grozi dziedziczeniem ubóstwa i trwałą marginalizacją (Lewandowski, 2006). Świadczenia dla bezrobotnych mogą także zachęcać osoby bierne zawodowo do pozorowanego szukania pracy w celu korzystania z przywilejów wiążących się ze statusem osoby bezrobotnej, co może prowadzić do powiększania się grupy osób uważanych za bezrobotne.

Efekty systemu zasiłkowego w Polsce badali m.in. Kaczorowski, Kubiak, Kwiatkowski, Pawlega i Uścińska (2009). Na podstawie analizy danych z BAEL za lata 2004–2008 sformułowali oni wniosek, że bezrobotni z prawem do zasiłku częściej wykazują bierną postawę w poszukiwaniu pracy. Jest on zgodny ze stanowiskiem, według którego świadczenia dla bezrobotnych zmniejszają aktywność w staraniach o znalezienie nowego zatrudnienia.

2.5. Definicja bezrobocia biernego

Zjawisko polegające na posiadaniu statusu osoby bezrobotnej w PUP pomimo niespełniania co najmniej jednego z podstawowych warunków posiadania statusu osoby bezrobotnej znane jest w literaturze jako bezrobocie pozorne. Wojciechowska (2006) wyróżnia wśród pozornie bezrobotnych dwie grupy: osoby, które pracują w szarej strefie, w związku z czym powinny zostać zaliczone do grona pracujących, oraz osoby, które zaniechały poszukiwań pracy, a więc powinny być uznane za bierne zawodowo. Podobnie Czapiński i Panek (2013) dzielą pozornie bezrobotnych na tych, którzy nie są zainteresowani pracą (nie szukają jej i/lub nie są gotowi jej podjąć), oraz pracujących na czarno lub w inny sposób osiągających dochód nie mniejszy niż minimalne wynagrodzenie za pracę.

Ludzie mający status osoby bezrobotnej, którzy nie pracują, a jednocześnie albo nie szukają pracy, albo nie chcą lub nie są gotowi jej podjąć, mogą znacznie różnić się od tych, którzy pracują w szarej strefie. Stąd też w celu przeciwdziałania tym zjawiskom należy stosować różne środki. W związku z tym Wilak (2018) zaproponował, aby od pozornie bezrobotnych oddzielić grupę osób biernych zawodowo i określać je mianem biernie bezrobotnych. Zgodnie z tą propozycją biernie bezrobotnymi nazywane będą te osoby zarejestrowane w PUP jako bezrobotne, które nie pracują, ale nie szukają pracy bądź nie są gotowe lub zdolne do jej podjęcia, a zatem nie powinny posiadać statusu osoby bezrobotnej.

2.6. Szacunki bezrobocia pozornego i biernego w Polsce

Lista opracowań podejmujących temat bezrobocia biernego w Polsce jest skromna. Do publikacji, w których częściowo omawia się to zagadnienie, należą raporty *Warunki i jakość życia Polaków* z projektu badawczego *Diagnoza społeczna*, realizowanego przez Radę Monitoringu Społecznego (Czapiński i Panek, 2013). Badanie było prowadzone z dwuletnią częstotliwością¹⁴ od 2000 r. do 2015 r. Jedną z jego części dotyczyła wykluczenia społecznego, m.in. bezrobocia. W ramach badania szacowano odsetek osób spełniających główne kryteria zawarte w ustawowej definicji osoby bezrobotnej. Na podstawie tych szacunków można obliczyć odsetek zarejestrowanych w PUP, którzy nie spełniają jednego z trzech głównych warunków posiadania statusu osoby bezrobotnej (tabl. 1)¹⁵.

Tabl. 1. Odsetek osób zarejestrowanych w PUP niespełniających warunków posiadania statusu osoby bezrobotnej

Zarejestrowani w PUP jako bezrobotni, którzy:	2003	2005	2007	2009	2011	2013
~~~~~						
Nie są gotowi do podjęcia pracy .....	15,3	16,5	28,8	27,3	17,4	18,0
Nie szukają pracy .....	9,2	7,4	10,4	6,0	12,9	18,0
Pracują lub osiągają dochód ^a .....	6,6	8,5	8,8	15,2	10,1	2,2

a Pracują w pełnym wymiarze czasu lub osiągają dochód powyżej połowy minimalnego wynagrodzenia.

Źródło: opracowanie własne na podstawie: Czapiński i Panek (2013).

Według szacunków Czapińskiego i Panka osoby, które nie były gotowe podjąć pracy w latach 2003–2013, stanowiły od 15,3% do 28,8%, a po dodaniu osób, które nie szukały pracy – od 24,5% do 39,2%. Duży wzrost odsetka osób, które nie są gotowe do podjęcia pracy, obserwowano pomiędzy latami 2005 a 2007.

¹⁴ Wyjątkiem jest trzyletni okres pomiędzy pierwszym a drugim badaniem.

¹⁵ Ze względu na brak możliwości wyznaczenia odpowiednich wartości nie zaprezentowano danych za rok 2015.

Informacje na temat poziomu bezrobocia biernego znajdują się również w opracowaniach GUS (2019c, 2019d). Na podstawie BAEL zaprezentowano szacunki liczby osób zarejestrowanych w PUP w podziale na bezrobotnych, pracujących i biernych zawodowo zgodnie z definicjami zalecanymi przez MOP. Według tych szacunków w IV kwartale 2017 i 2018 r. odsetek biernych zawodowo wśród zarejestrowanych jako bezrobotni stanowił odpowiednio 50,4% i 51,8%.

Szeroką analizę bezrobocia pozornego i biernego w Polsce w latach 2001–2011 przeprowadził Wilak (2018). Przy wykorzystaniu danych jednostkowych z BAEL oszacowane zostały liczba i odsetek osób pozornie i biernie bezrobotnych w podziale na płeć, wiek, miejsce zamieszkania (miasto, wieś) i poziom wykształcenia, a także w przekroju województw. W wypadku bezrobocia biernego zaobserwowano m.in. duże zróżnicowanie względem płci – znacznie częściej cechowało ono kobiety. Odsetek biernie bezrobotnych wśród kobiet zarejestrowanych jako bezrobotne wynosił od 23,5% do 50,6%, podczas gdy wśród mężczyzn wahał się między 10,5% a 33,7%.

### **3. Metoda badania**

#### **3.1. Identyfikacja biernie bezrobotnych w BAEL**

Podstawą estymacji bezrobocia biernego są respondenci biorący udział w BAEL, którzy deklarują posiadanie statusu osoby bezrobotnej w PUP. Do ich wyodrębnienia wykorzystano pytanie: „Czy jest Pan(i) zarejestrowany(-na) w powiatowym urzędzie pracy jako bezrobotny(-na)?”, znajdujące się na końcu ankiety ZD. Trzeba mieć na uwadze, że odpowiedź na to pytanie jest nieweryfikowaną deklaracją osoby ankietowanej, a więc może nie być zgodna ze stanem faktycznym.

Identyfikacji biernie bezrobotnych dokonano za pomocą procedury zastosowanej przez Wilaka (2018), według której respondenci deklarujący posiadanie statusu bezrobotnego w PUP są podzieleni na trzy grupy: faktycznie, pozornie i biernie bezrobotnych¹⁶. Procedura ta bazuje na zmiennej wyznaczającej status respondenta na rynku pracy określony według rekomendowanych przez MOP definicji osób pracujących, bezrobotnych i biernych zawodowo. Respondenci deklarujący posiadanie statusu osoby bezrobotnej zarejestrowanej w PUP, mimo że w świetle odpowiedzi zawartych w ankiecie ich status określono jako biernych zawodowo, zgodnie z wyżej przyjętą definicją zaliczeni zostali do grona osób biernie bezrobotnych.

¹⁶ Według takiej samej procedury w opracowaniu GUS (2019d) zestawiono liczbę osób deklarujących rejestrację w PUP jako bezrobotne w podziale na pracujących, bezrobotnych i biernych zawodowo zgodnie z definicjami stosowanymi w BAEL.

**Tabl. 2.** Podstawowe charakterystyki rozkładu liczby osób zidentyfikowanych jako biernie bezrobotne w kwartalnych próbach BAEL w latach 2010–2018

Domeny	min	$Q_1$	$Me$	$\bar{x}$	$Q_3$	max
O g ó ł e m .....	684	1104	1467	1332	1576	1764
<b>Płeć</b>						
Mężczyźni .....	221	366	500	458	543	665
Kobiety .....	463	737	950	875	1031	1115
<b>Wiek</b>						
18–24 lata .....	72	122	205	192	262	292
25–34 .....	179	284	380	353	419	473
35–44 .....	152	234	257	252	282	329
45–54 .....	123	209	295	271	336	381
55 lat i więcej .....	144	223	267	264	311	347
<b>Miejsce zamieszkania</b>						
Miasto .....	381	620	761	709	830	895
Wieś .....	303	473	691	624	737	934
<b>Poziom wykształcenia</b>						
Co najwyżej gimnazjalne .....	121	201	313	276	345	398
Zawodowe .....	245	370	528	474	563	643
Średnie .....	243	400	498	464	547	634
Wyższe .....	75	109	120	119	129	157

Uwaga. min – minimum,  $Q_1$  – kwartył 1,  $Me$  – mediana,  $\bar{x}$  – średnia,  $Q_3$  – kwartył 3, max – maksimum.

Źródło: opracowanie własne na podstawie jednostkowych danych z BAEL.

Liczba reprezentantów osób biernie bezrobotnych w kwartalnych próbach BAEL waha się pomiędzy 684 a 1764; jej przeciętna wartość wynosi 1332 (tabl. 2). Najliczniej reprezentowaną domenę stanowią kobiety, których w kwartalnej próbie jest przeciętnie 875, a najmniejszą liczbę reprezentantów wśród analizowanych domen mają osoby z wykształceniem wyższym, których minimalna liczba w próbach kwartalnych wynosi 75, a przeciętna wartość jest równa 119.

### 3.2. Idea kalibracji wag

Do estymacji bezrobocia biernego wykorzystano podejście kalibracyjne. Metoda kalibracji wag jest techniką powszechnie wykorzystywaną w praktyce badań reprezentacyjnych. Opiera się na korekcie wag z badania, w wyniku której przeważone wartości zmiennych pomocniczych w próbie sumują się do ich znanych wartości globalnych. Zapewnia to często postulowaną w badaniach statystycznych zgodność rozumianą jako otrzymywanie takich samych szacunków w różnych badaniach statystycznych. Co więcej, w przypadku gdy zmienne pomocnicze są silnie skorelowane z inną zmienną, można się spodziewać, że oszacowania jej parametrów przy zastosowaniu skalibrowanych wag będą się charakteryzowały wyższą jakością.



Podejście kalibracyjne stosowane jest także jako środek zaradczy na powszechnie spotykane w badaniach ankietowych braki odpowiedzi, które mają negatywny wpływ na proces estymacji. Problem ten dotyczy również BAEL, w którym odsetek zbadanych mieszkań w ostatnich latach systematycznie malał – w roku 2018 wynosił poniżej 60%. Braki odpowiedzi nie tylko zmniejszają efektywną liczebność próby – co przekłada się na zmniejszenie precyzji szacunków – lecz także są źródłem błędów nielosowych. Wynika to z faktu, że osoby odmawiające odpowiedzi mogą różnić się od respondentów pod względem pewnych kluczowych cech. W związku z tym braki odpowiedzi mogą powodować również zniekształcenie rozkładu badanej cechy i w konsekwencji obciążenie oszacowań jego parametrów. Wykorzystanie podejścia kalibracyjnego w sytuacji braków odpowiedzi służy takiej modyfikacji wag, żeby z jednej strony zrekomensować utratę informacji, a z drugiej – zredukować obciążenie szacunków wynikające z nielosowego charakteru braków odpowiedzi.

Teoretyczne podstawy kalibracji zostały sformułowane przez Deville'a i Särndala (1992), którzy przedstawili sposób konstrukcji estymatora kalibracyjnego wartości globalnej przy założeniu, że znane są wartości zmiennych pomocniczych dla wszystkich jednostek w próbie. Z kolei temat podejścia kalibracyjnego w sytuacji braków odpowiedzi podejmowany był m.in. w pracach Särndala i Lundströma (2005) oraz Szymkowiaka (2019).

### 3.3. Formalne ujęcie kalibracji wag

W niniejszym artykule rozważana jest sytuacja braków odpowiedzi w badanej próbie. Niech  $U$  oznacza populację, a  $s \subseteq U$  – próbę wylosowaną z tej populacji. Oznaczmy przez  $r \subseteq s$  tę część próby wylosowanej do badania, w której znajdują się tylko zbadane jednostki. Niech  $d_j: j \in r$  będą wagami, które podlegają kalibracji, a  $w_j: j \in r$  – szukanymi wagami kalibracyjnymi. Niech  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_K)$  będzie wektorem zmiennych pomocniczych o znanych wartościach jednostkowych  $\mathbf{x}_j = (x_{1j}, \dots, x_{Kj})$ ,  $j \in r$ , a także znany wektorze wartości globalnych  $\mathbf{X} = \sum_{j \in U} \mathbf{x}_j = (\sum_{j \in U} x_{1j}, \dots, \sum_{j \in U} x_{Kj})$ . Ponadto niech  $G(\cdot)$  będzie funkcją odległości pomiędzy ilorazem  $\frac{w_j}{d_j}$  a 1. Funkcja  $G(\cdot)$  powinna spełniać następujące warunki: ściśła wypukłość, dwukrotna różniczkowalność,  $G(\cdot) \geq 0$ ,  $G'(1) = 0$ ,  $G''(1) = 1$ .

Wyznaczanie wag kalibracyjnych  $w_j: j \in r$  przy wykorzystaniu wektora zmiennych pomocniczych  $\mathbf{x}$  sprowadza się do:

- minimalizacji funkcji odległości:

$$\sum_{j \in r} d_j G\left(\frac{w_j}{d_j}\right) \rightarrow \min \quad (1)$$

- spełnienia równań kalibracyjnych:

$$\sum_{j \in r} w_j x_j = X \quad (2)$$

- spełnienia warunków ograniczających:

$$L_1 \leq \frac{w_j}{d_j} \leq L_2 \quad L_1 < 1 < L_2 \quad j \in r \quad (3)$$

Minimalizacja funkcji odległości (1) ma na celu otrzymanie takich wag kalibracyjnych, które przy przyjętej funkcji odległości są możliwie najbliższe wagom podlegającym kalibracji. Dzięki spełnieniu równań kalibracyjnych (2) wagi wyznaczone są w taki sposób, aby przy ich zastosowaniu wszystkie przeważone zmienne pomocnicze sumowały się w próbie do wartości globalnych. Ponieważ niektóre funkcje odległości mogą generować wagi ujemne bądź ekstremalne, na wagi kalibracyjne  $w_j$  często wprowadza się ograniczenia. Mówi o tym warunek (3), z którego wynika, że  $d_j L_1 \leq w_j \leq d_j L_2$ . Zbyt restrykcyjne przyjęcie parametrów  $L_1$  i  $L_2$  może jednak powodować brak zbieżności algorytmu iteracyjnego i w konsekwencji niespełnienie równań kalibracyjnych.

Przy wyborze funkcji  $G(\cdot)$  istnieje pewna dowolność. W literaturze omawiane są takie jej postaci, jak m.in. metody:

- liniowa:

$$G_1(x) = \frac{1}{2}(x - 1)^2 \quad (4)$$

- wykładnicza (raking¹⁷):

$$G_2(x) = x(\log x - 1) + 1 \quad (5)$$

- logitowa:

$$G_3(x) = \left[ (x - L_1) \log \left( \frac{x - L_1}{1 - L_1} \right) + (L_2 - x) \log \left( \frac{L_2 - x}{L_2 - 1} \right) \right] \frac{(1 - L_1)(L_2 - 1)}{L_2 - L_1} \quad (6)$$

¹⁷ Raking (także Raking Ratio Estimation) jako metoda poszukiwania liczebności w tablicy kontyngencji bazująca na algorytmie IPF (Iterative Proportional Fitting) jest szczególnym przypadkiem kalibracji wag (Szymkowiak, 2019). W polskiej literaturze nie ma adekwatnego tłumaczenia nazwy tej metody, dlatego w jej opisie często pozostawia się oryginalną nazwę.

- z sinusem hiperbolicznym:

$$G_4(x) = \frac{1}{2\alpha} \int_1^x \sin h\left(\alpha - \frac{\alpha}{t}\right) dt, \quad \text{gdzie } \sin h(x) = \frac{e^x - e^{-x}}{2} \quad (7)$$

Zastosowanie metody liniowej daje takie same wyniki jak estymator GREG z identycznym zestawem zmiennych pomocniczych, problemem jest jednak możliwość występowania wag ujemnych. Raking daje wagi dodatnie, mogą natomiast wystąpić wagi ekstremalne, które znacznie różnią się od wag wejściowych. W przypadku metod logitowej i z sinusem hiperbolicznym można nałożyć ograniczenia ujęte w warunku ograniczającym (3), tak aby wagi kalibracyjne nie różniły się zbyt od wag wejściowych. Parametr  $\alpha$  w funkcji odległości  $G_4$  przyjmuje wartości dodatnie i pozwala sterować stopniem rozrzutu wag kalibracyjnych w stosunku do wag wyjściowych. W omawianym badaniu zastosowano wszystkie cztery funkcje, spośród których wybrano tę, której użycie daje wagi kalibracyjne o najlepszych własnościach.

Wyznaczenie wag kalibracyjnych  $w_j$  sprowadza się do znalezienia rozwiązania zadania minimalizacji opisanego w warunku (1) przy jednoczesnym spełnieniu równań kalibracyjnych (2). Można w tym celu wykorzystać metodę czynników nieoznaczonych Lagrange'a, służącą do znajdowania ekstremum warunkowego funkcji różniczkowalnej. Funkcja Lagrange'a wyraża się wtedy wzorem:

$$\mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}; \mathbf{d}) = \sum_{j \in r} d_j G\left(\frac{w_j}{d_j}\right) - \boldsymbol{\lambda}^T \left( \sum_{j \in r} w_j \mathbf{x}_j - \mathbf{X} \right) \quad (8)$$

gdzie  $\boldsymbol{\lambda}^T = (\lambda_1, \dots, \lambda_K)^T$  jest wektorem mnożników (czynników nieoznaczonych) Lagrange'a. Równania Lagrange'a wyrażają się zatem wzorem:

$$\frac{\partial \mathcal{L}}{\partial w_j} = d_j \frac{\partial}{\partial w_j} G\left(\frac{w_j}{d_j}\right) - \boldsymbol{\lambda}^T \mathbf{x}_j = 0 \quad (9)$$

Ostatecznie wagi kalibracyjne wyznaczone za pomocą metody czynników nieoznaczonych Lagrange'a można przedstawić w postaci:

$$w_j = d_j F(\mathbf{x}_j^T \boldsymbol{\lambda}) = d_j g_j \quad (10)$$

gdzie:

$g_j = F(\mathbf{x}_j^T \boldsymbol{\lambda})$  – tzw. mnożniki wagowe (kalibracyjne),

$F(\cdot) = G'^{-1}(\cdot)$  – funkcja kalibracyjna, która jest funkcją odwrotną do pierwszej pochodnej funkcji odległości  $G(\cdot)$ .

Postaci funkcji kalibracyjnych  $F(\cdot)$  dla wybranych funkcji odległości  $G(\cdot)$  można znaleźć m.in. w pracach Deville'a i Särndala (1992) czy Szymkowiaka (2019).

Do wyznaczenia wag kalibracyjnych (10) w dalszym ciągu niezbędna jest znajomość wektora mnożników Lagrange'a  $\lambda$ , którego postać można znaleźć, rozwiązując równanie kalibracyjne (2):

$$\sum_{j \in r} w_j F(x_j^T \lambda) x_j = X \quad (11)$$

Wyznaczenie wektora  $\lambda$  czynników nieoznaczonych Lagrange'a, a w konsekwencji wag kalibracyjnych (10), zależy od postaci funkcji  $G(\cdot)$ . W metodzie liniowej istnieje możliwość jawnego przedstawienia wag kalibracyjnych w postaci odpowiedniego wzoru (Szymkowiak, 2019). W przypadku funkcji  $G_2$ ,  $G_3$  i  $G_4$  wymagane jest zastosowanie podejścia iteracyjnego. Algorytmy wyznaczania wektora  $\lambda$  dla tych funkcji można znaleźć m.in. w pracy Szymkowiaka (2019).

### 3.4. Ocena precyzji estymacji za pomocą bootstrapu

Precyzja estymacji należy do głównych aspektów jej jakości. Najczęściej stosowaną miarą precyzji estymacji jest wariancja estymatorów, a także obliczany na jej podstawie względny błąd szacunku REE (ang. *relative estimation error*).

Precyzja estymacji w badaniach reprezentacyjnych zależy od liczebności próby, schematu losowania próby oraz zastosowanego estymatora. W praktyce, w tym także w BAEL, najczęściej stosuje się złożone metody doboru próby, uwzględniające takie zabiegi jak warstwowanie, losowanie wielostopniowe czy nadreprezentację. Również procedury estymacji stosowane w badaniach statystycznych są często złożone i mogą obejmować m.in. kalibrację wag. W BAEL próby wybierane są za pomocą losowania dwustopniowego z warstwowaniem, a estymację charakterystyk przeprowadza się z wykorzystaniem wag konstruowanych w dwukrokowej procedurze. W takich przypadkach analityczne wyprowadzenie wariancji jest często bardzo trudne lub wręcz niemożliwe, a zatem należy ją wyznaczyć w inny sposób.

W literaturze można znaleźć propozycje różnych metod szacowania wariancji w przypadku złożonych procedur estymacji; jedną z najczęściej stosowanych jest zainicjowana przez Efrona (1979) metoda bootstrap. Opiera się ona na założeniu, że rozkład badanego parametru w próbie jest taki sam jak w populacji. W procedurze bootstrapu do szacowania rozkładu estymatorów wykorzystuje się podejście Monte Carlo – próbę traktuje się jak populację, z której za pomocą przyjętego w badaniu schematu losowania generuje się wielokrotnie próby bootstrapowe,

a następnie na ich podstawie oblicza się wartości estymatorów. Za oszacowanie rozkładu estymatora przyjmuje się jego empiryczny rozkład w próbie otrzymany w wyniku zastosowania metody Monte Carlo.

W badaniu omawianym w niniejszym artykule do estymacji wariancji wykorzystano wersję bootstrapu dostosowaną przez Rao i Wu (1988) do losowania dwustopniowego z warstwowaniem, z uwzględnieniem modyfikacji wprowadzonych przez Rao, Wu i Yue (1992). Metoda ta znana jest w literaturze jako Rao-Wu bootstrap lub rescaling bootstrap (RBS), a wykorzystywana (m.in. w BAEL) do wyznaczania względnego błędu szacunku. Ogólna procedura jest następująca:

1. Losowanie próby bootstrapowej. Z każdej warstwy  $h$  niezależnie, spośród wszystkich  $n_h$  wylosowanych do próby jednostek pierwszego stopnia (JPS), wybiera się  $m_h$  jednostek za pomocą losowania prostego ze zwracaniem. Wszystkie jednostki drugiego stopnia z próby, które znajdują się w wylosowanych JPS, tworzą próbę bootstrapową.
2. Wyznaczenie wag bootstrapowych. Niech  $d_j$  oznacza wagę pierwotną  $j$ -tej jednostki drugiego stopnia, zaś  $c_j$  liczbę wylosowań tej jednostki do próby bootstrapowej. Waga bootstrapowa  $j$ -tej jednostki drugiego stopnia przyjmuje postać:

$$d_j^* = \left[ \left( 1 - \left( \frac{m_h}{n_h - 1} \right)^{1/2} \right) + \left( \frac{m_h}{n_h - 1} \right)^{1/2} \left( \frac{n_h}{m_h} \right) c_j \right] d_j \quad (12)$$

3. Wyznaczenie oceny estymatora bootstrapowego. Na podstawie wylosowanej próby bootstrapowej, przy wykorzystaniu wag bootstrapowych, oblicza się wartość estymatora bootstrapowego  $\hat{Y}^*$ . W tym celu na wagach bootstrapowych wykonuje się wszystkie niezbędne zabiegi, aby z wag pierwotnych otrzymać wartość oryginalnego estymatora  $\hat{Y}$ , w tym kalibrację wag.
4. Wielokrotne próbkowanie. Kroki 1–3 powtarza się  $B$  razy, w wyniku czego otrzymuje się  $B$  ocen estymatora bootstrapowego:  $\hat{Y}^{*1}, \dots, \hat{Y}^{*B}$ .
5. Estymacja wariancji. Oszacowanie wariancji estymatora  $\hat{Y}$  ma postać:

$$D^2(\hat{Y}) = \frac{1}{B-1} \sum_{b=1}^B (\hat{Y}^{*b} - \bar{Y}^*)^2, \quad \text{gdzie } \bar{Y}^* = \frac{1}{B} \sum_{b=1}^B \hat{Y}^{*b} \quad (13)$$

Implementacja powyższej procedury wiąże się z wyborem liczby losowanych jednostek  $m_h$ . Kolenikov (2010) oraz Bruch, Münnich i Zins (2011) zalecają, aby z każdej warstwy losować  $m_h = n_h - 1$  jednostek losowania pierwszego stopnia. Wagi bootstrapowe (12) przyjmują wtedy postać:

$$d_j^* = \left( \frac{n_h}{n_h - 1} c_j \right) d_j \quad (14)$$

Należy również wybrać odpowiednio dużą liczbę replikacji  $B$ , ponieważ im większa liczba, tym bardziej zwiększa się dokładność bootstrapu, choć towarzyszy temu wzrost czasochłonności obliczeń. W wielu przypadkach wystarcza już 100 powtórzeń (Bruch i in., 2011). W BAEL stosuje się 500 replikacji.

### 3.5. Procedura estymacji

W badaniu omawianym w niniejszym artykule do estymacji bezrobocia pozornego i biernego wykorzystano podejście kalibracyjne. Zastosowano procedurę dwukrokową, podobną do stosowanej w BAEL (GUS, 2019a). Różnica polega na uwzględnieniu w drugim kroku kalibracji dodatkowych zmiennych pomocniczych dotyczących bezrobocia rejestrowanego, a także z wykorzystania różnych funkcji odległości.

W pierwszym kroku wagi pierwotne, będące odwrotnością prawdopodobieństwa wylosowania do próby  $j$ -tej jednostki, korygowane są ze względu na braki odpowiedzi. Na podstawie wag pierwotnych wylosowanych do próby mieszkań wyznacza się szacunki liczby mieszkań, które powinny zostać zbadane. Oblicza się je w przekroju województw i klas miejscowości. Następnie szacunki te wykorzystuje się jako wartości globalne w procesie kalibracji wag pierwotnych. Wynikiem tego zabiegu są tzw. wagi pośrednie, będące jednocześnie finalnymi wagami mieszkań, które z kolei przypisuje się osobom zamieszkującym zbadane mieszkania.

W drugim kroku modyfikacji podlegają wagi pośrednie przypisane wspomnianym osobom. Są one korygowane w taki sposób, aby odtwarzały znaną ze źródeł zewnętrznych strukturę demograficzną badanej populacji, a także liczby osób zarejestrowanych w PUP jako bezrobotne. Finalne wagi osób sumują się do wartości globalnych liczby osób w 102 domenach jednocześnie:

- 48 domen osób w wieku 15 lat i więcej w podziale na płeć  $\times$  miejsce zamieszkania (miasto, wieś)  $\times$  wiek (15–17, 18–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65 lat i więcej);
- 54 domeny osób zarejestrowanych w PUP jako bezrobotne w podziale na:
  - płeć  $\times$  wiek (mniej niż 25, 25–34, 35–44, 45–54, 55 lat i więcej);
  - płeć  $\times$  miejsce zamieszkania (miasto, wieś);
  - płeć  $\times$  poziom wykształcenia (co najwyżej gimnazjalne, zasadnicze zawodowe, średnie, wyższe);
  - płeć  $\times$  województwo.

Wektor zmiennych pomocniczych  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_{102})$  składa się ze zmiennych indykatorych, gdzie zmienna  $\mathbf{x}_k: k = 1, \dots, 102$  jest równa 1 dla osób należących do  $k$ -tej domeny, a dla pozostałych osób przyjmuje wartość 0. Wartości globalne  $X_k: k = 1, \dots, 48$  dla pierwszych 48 zmiennych pomocniczych przyjęto takie, jakie zostały zastosowane przy konstrukcji wag w BAEL. Dla pozostałych zmiennych pomocniczych jako wartości globalne  $X_k: k = 49, \dots, 102$  przyjęto liczbę osób zarejestrowanych w PUP

jako bezrobotne według stanu na koniec kwartału. Dane na ten temat, pochodzące ze sprawozdań MPiPS-01, zaczerpnięto ze strony internetowej Banku Danych Lokalnych (BDL) GUS.

Powyższa procedura jest przykładem podejścia dwukrokowego typu B, w którym w procesie estymacji uwzględnia się zmienne pomocnicze o wartościach znanych na poziomie próby (pierwszy krok) oraz zmienne pomocnicze o wartościach znanych na poziomie całej populacji (drugi krok) (Särndal i Lundström, 2005).

W procesie estymacji zastosowano wszystkie cztery funkcje odległości opisane wzorami (4)–(7). W przypadku metod logitowej i z sinusem hiperbolicznym zastosowano ograniczenia na wagi kalibracyjne (3) w postaci¹⁸  $L_1 = 0,25$ , a  $L_2 = 4$ . Wyboru ostatecznej postaci funkcji  $G(\cdot)$  dokonano na podstawie analizy statystycznych własności uzyskanych wag kalibracyjnych.

Za pomocą wag kalibracyjnych otrzymanych przy zastosowaniu wybranej funkcji odległości wyznaczono oceny estymatorów kalibracyjnych liczby biernie bezrobotnych i ich odsetka wśród osób zarejestrowanych w PUP jako bezrobotne. Szacunki przeprowadzono dla lat 2010–2018 w ujęciu kwartalnym, w podziale na:

- płeć (kobiety, mężczyźni);
- pięć grup wieku (mniej niż 25, 25–34, 35–44, 45–54, 55 lat i więcej);
- miejsce zamieszkania (miasto, wieś);
- poziom wykształcenia (co najwyżej gimnazjalne, zasadnicze zawodowe, średnie, wyższe).

Łącznie otrzymano po 468 (36 kwartałów  $\times$  13 domen) ocen estymatora liczby i odsetka biernie bezrobotnych, które dla domeny  $d$  i kwartału  $t$  zostały wyznaczone za pomocą wzorów:

$$\hat{Y}_{dt} = \sum_{j \in r_{dt}} 1_B(j) w_j \quad (15)$$

$$\hat{p}_{dt} = \frac{\hat{Y}_{dt}}{Y_{dt}^{rej}} \quad (16)$$

gdzie:

$r_{dt}$  – zbiór respondentów,

$1_B(j)$  – funkcja charakterystyczna zbioru osób zidentyfikowanych jako biernie bezrobotne,

$Y_{dt}^{rej}$  – liczba osób zarejestrowanych w PUP jako bezrobotne.

¹⁸ Ograniczeń  $(L_1, L_2)$  szukano w zbiorze  $\left\{ \left( \frac{1}{k}, k \right) : k \in N \right\}$ . Ostatecznie wybrano możliwie najbardziej restrykcyjne ograniczenia, a więc zastosowano najmniejsze  $k$ , dla którego osiągnięta była zbieżność iteracyjnego algorytmu wyznaczania wag.

Jakość otrzymanych ocen estymatora kalibracyjnego przeanalizowano za pomocą metody Rao-Wu bootstrap. Dla każdej z 468 ocen estymatora liczby i odsetka osób bezrobotnych wyznaczono względny błąd szacunku REE w postaci:

$$REE(\hat{Y}_{dt}) = REE(\hat{p}_{dt}) = \frac{\sqrt{D^2(\hat{Y}_{dt})}}{|\hat{Y}_{dt}|} \quad (17)$$

gdzie:

$\hat{Y}_{dt}$  – ocena estymatora liczby biernie bezrobotnych w domenie  $d$  ( $d = 1, 2, \dots, 13$ ) i kwartale  $t$  ( $t = 1, 2, \dots, 36$ ),

$D^2(\hat{Y}_{dt})$  – oszacowanie jego wariancji według wzoru (14).

Do ich wyznaczenia dla każdego z kwartałów  $t$  ( $t = 1, \dots, 36$ ) wykonano  $B = 500$  replikacji. Równość  $REE(\hat{Y}_{dt}) = REE(\hat{p}_{dt})$  w równaniu (17) wynika stąd, że w mianowniku estymatora frakcji we wzorze (16) liczba zarejestrowanych w PUP jako bezrobotni  $Y_{dt}^{rej}$  jest wartością stałą, tzn. nie jest obciążona błędem estymacji, a zatem:

$$\begin{aligned} REE(\hat{p}_{dt}) &= \frac{\sqrt{D^2(\hat{p}_{dt})}}{|\hat{p}_{dt}|} = \frac{\sqrt{D^2(\hat{Y}_{dt}/Y_{dt}^{rej})}}{|\hat{Y}_{dt}/Y_{dt}^{rej}|} = \frac{\sqrt{D^2(\hat{Y}_{dt})/Y_{dt}^{rej}/Y_{dt}^{rej}}}{|\hat{Y}_{dt}/Y_{dt}^{rej}|} = \frac{\sqrt{D^2(\hat{Y}_{dt})}}{|\hat{Y}_{dt}|} \end{aligned} \quad (18)$$

## 4. Wyniki badań

### 4.1. Analiza wyników kalibracji

W ocenie wyników kalibracji wag wzięto pod uwagę następujące aspekty: stopień odтворzenia znanych wartości globalnych, występowanie wag ujemnych, występowanie wag ekstremalnych oraz stopień odchylenia wag kalibracyjnych od pierwotnych. Z oceny wynika, że:

- wszystkie zastosowane podejścia pozwoliły osiągnąć cel kalibracji, jakim jest odтворzenie znanych wartości globalnych zmiennych pomocniczych;
- zastosowanie podejścia liniowego skutkuje otrzymaniem ujemnych wag (tabl. 3), co sprawia, że w tym przypadku nie jest ono zalecane;
- za pomocą rakingu otrzymuje się wagi najbardziej odstające od przeciętnych, co jest własnością niepożądaną;



- w przypadku rakingu występują znaczne odchylenia od wag pierwotnych, co również jest niewskazane; najmniejsze odchylenia wag kalibracyjnych uzyskuje się za pomocą metod logitowej i z sinusem hiperbolicznym.

**Tabl. 3.** Podstawowe charakterystyki rozkładu wag i mnożników kalibracyjnych

Metoda	min	$Q_1$	$Me$	$\bar{x}$	$Q_3$	max	$Sd$
<b>Wagi kalibracyjne</b>							
Liniowa .....	-1038,4	245,8	378,3	447,5	564,1	3993,3	299,1
Raking .....	28,2	249,4	374,6	447,5	552,7	5072,8	302,2
Logitowa .....	27,5	247,0	374,3	447,5	557,9	3825,1	300,1
Z sinusem hiperbolicznym .....	29,7	242,2	373,8	447,5	567,3	3605,9	300,2
<b>Mnożniki kalibracyjne</b>							
Liniowa .....	-3,550	1,333	1,677	1,840	2,188	9,288	0,791
Raking .....	0,185	1,332	1,636	1,841	2,106	14,116	0,803
Logitowa .....	0,411	1,312	1,640	1,840	2,151	9,383	0,791
Z sinusem hiperbolicznym .....	0,417	1,268	1,669	1,840	2,244	9,293	0,793

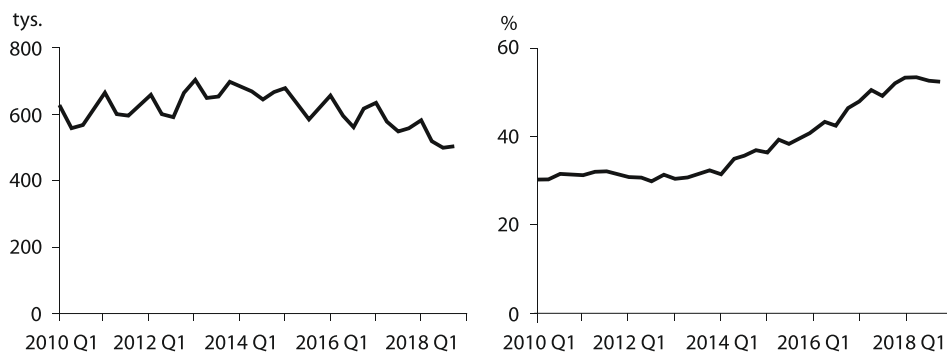
Uwaga. Jak przy tabl. 2.  $Sd$  – odchylenie standardowe.

Źródło: opracowanie własne na podstawie jednostkowych danych z BAEL oraz zagregowanych danych ze sprawozdań MPiPS-01.

Wagi otrzymane za pomocą metod logitowej i z sinusem hiperbolicznym charakteryzują się najlepszymi własnościami. Ze względu na szybkość osiągania zbieżności algorytmu wyznaczania wag kalibracyjnych, co jest szczególnie ważne w przypadku czasochłonnej procedury bootstrap, w dalszej części pracy zastosowano metodę logitową.

#### 4.2. Analiza wyników estymacji kalibracyjnej

W latach 2010–2013 liczba biernie bezrobotnych charakteryzowała się nieznacznym trendem rosnącym (wykr. 2, tabl. 4). Maksymalną wartość równą 699,3 tys. osiągnęła w I kwartale 2013 r. Z kolei od roku 2014 obserwuje się nieznaczną tendencję malejącą liczby biernie bezrobotnych. Jej najniższa wartość wynosząca 497,5 tys. została osiągnięta w III kwartale 2018 r. Tendencja spadkowa liczby biernie bezrobotnych po roku 2013 odpowiada zmianom dotyczącym liczby wszystkich osób zarejestrowanych w PUP jako bezrobotne (wykr. 1), jednak dynamika zmian liczby biernie bezrobotnych jest znacznie mniejsza. W efekcie po roku 2014 obserwuje się znaczący wzrost frakcji biernie bezrobotnych. Od II kwartału 2017 r. przekracza ona poziom 50% (z wyjątkiem III kwartału 2017 r.), osiągając w II kwartale 2018 r. maksymalną wartość 53,3%. Natomiast do roku 2014 odsetek biernie bezrobotnych utrzymywał się na poziomie niewiele wyższym niż 30%, raz tylko spadając nieznacznie poniżej tej granicy.

**Wykr. 2.** Liczba i odsetek biernie bezrobotnych

Uwaga. Jak przy wyk. 1.

Źródło: opracowanie własne na podstawie danych jednostkowych z BAEL oraz zagregowanych danych ze sprawozdań MPiS-01.

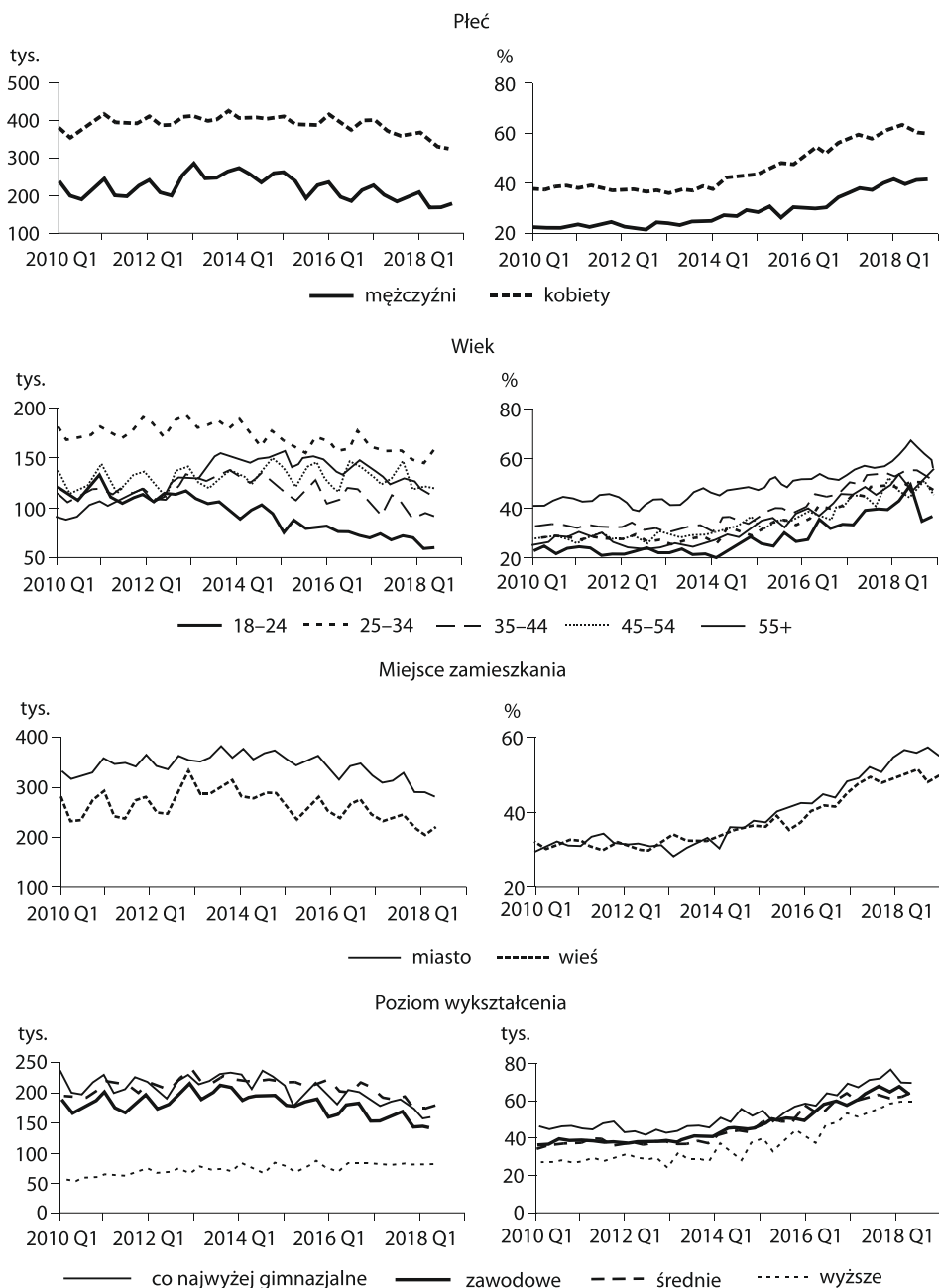
Bezrobocie bierne charakteryzuje się zróżnicowaniem ze względu na płeć (wykr. 3 i tabl. 4). Kobiety zarejestrowane jako bezrobotne znacznie częściej niż mężczyźni cechują się biernością w tym zakresie. Odsetek biernie bezrobotnych kobiet w analizowanych latach przyjmował wartości między 36,3% a 63,8%, a jego przeciętny poziom wyniósł 46,1%. Wśród mężczyzn zmieniał się między 21,6% a 42,0%, przyjmując przeciętną wartość równą 28,9%.

Na poziom bezrobocia biernego wpływają też wiek i poziom wykształcenia. Domenami, w których obserwuje się stosunkowo duży odsetek biernie bezrobotnych, są osoby w wieku 55 lat i więcej oraz osoby z wykształceniem co najwyżej gimnazjalnym (wykr. 3). Natomiast miejsce zamieszkania ma niewielki wpływ na poziom badanego zjawiska. W analizowanym okresie odsetek biernie bezrobotnych mieszkających na wsi i odsetek mieszkających w mieście były zbliżone (wykr. 3 i tabl. 4).

W niektórych domenach – m.in. w grupie osób w wieku 18–24 lat czy w grupie osób z wykształceniem co najwyżej gimnazjalnym – w ostatnich kwartałach analizowanego okresu można było zaobserwować dość znaczny spadek odsetka biernie bezrobotnych. Interpretacja tych wyników wymaga przeprowadzenia szacunków obejmujących kolejne kwartały.

Do oceny precyzji estymacji wykorzystano względny błąd szacunku REE wyznaczony za pomocą bootstrapu. Najmniejszą precyzją cechują się domeny o najniższej liczbie reprezentantów bezrobocia biernego w próbie, a więc grupy osób w wieku 55 lat i więcej oraz osób z wykształceniem wyższym, w których przeciętne wartości REE wyniosły odpowiednio 5,9% i 7,0% (tabl. 4). Natomiast w żadnym wypadku wartość REE nie przekroczyła 10%. Można zatem uznać, że szacunki liczby i odsetka biernie bezrobotnych, zarówno ogółem dla całej populacji, jak i we wszystkich badanych przekrojach, cechują się akceptowalną precyzją.

**Wykr. 3.** Liczba i odsetek biernie bezrobotnych według płci, wieku, miejsca zamieszkania i poziomu wykształcenia



Uwaga. Jak przy wykr. 1.

Źródło: opracowanie własne na podstawie danych jednostkowych z BAEL oraz zagregowanych danych ze sprawozdań MPIS-01.

**Tabl. 4.** Podstawowe charakterystyki rozkładu liczby i odsetka biernie bezrobotnych oraz względnego błędu szacunku REE w latach 2010–2018

Domeny	Liczba biernie bezrobotnych w tys.						Odsetek biernie bezrobotnych						REE					
	min	Q ₁	Me	$\bar{x}$	Q ₃	max	min	Q ₁	Me	$\bar{x}$	Q ₃	max	min	Q ₁	Me	$\bar{x}$	Q ₃	max
Ogółem .....	497,5	579,8	616,8	612,7	655,0	699,3	29,8	31,2	35,1	38,0	43,7	53,3	1,9	2,0	2,1	2,1	2,1	2,9
<b>Płeć</b>																		
Mężczyźni .....	168,6	197,7	220,4	221,5	245,4	285,6	21,6	23,6	26,8	28,9	31,8	42,0	3,4	3,5	3,7	3,9	4,2	5,6
Kobiety .....	325,8	378,2	396,3	391,1	408,7	429,7	36,3	37,8	42,6	46,1	54,9	63,8	2,0	2,2	2,3	2,4	2,4	3,2
<b>Wiek</b>																		
18–24 lat .....	45,5	64,6	88,5	85,1	102,8	128,3	23,1	25,5	27,5	30,8	35,7	51,4	4,4	4,9	5,4	5,9	6,6	9,6
25–34 .....	141,3	159,0	170,3	170,7	182,0	195,3	28,2	30,4	32,8	37,1	42,9	58,1	3,4	3,5	3,8	3,9	4,0	5,3
35–44 .....	96,9	110,8	116,9	120,1	130,7	147,2	26,2	28,7	32,4	36,2	42,8	55,9	4,0	4,5	4,6	4,8	5,0	6,6
45–54 .....	79,5	102,8	116,1	113,6	124,9	140,3	27,5	32,0	34,6	37,4	42,9	53,6	4,0	4,3	4,5	4,9	5,3	7,3
55 lat i więcej .....	78,3	106,9	125,7	123,2	144,0	156,8	40,1	45,9	49,5	51,1	55,5	68,8	3,5	3,8	4,1	4,2	4,4	5,4
<b>Miejsce zamieszkania</b>																		
Miasto .....	283,6	333,5	351,9	347,4	365,1	391,0	27,6	30,8	35,3	38,7	45,0	56,6	2,6	2,8	2,8	2,9	2,9	3,8
Wieś .....	204,1	240,1	268,7	265,3	287,1	338,9	29,1	31,7	34,5	37,1	42,1	50,8	2,8	3,0	3,2	3,3	3,4	4,7
<b>Poziom wykształcenia</b>																		
Co najwyżej gimnazjalne ...	142,4	178,3	193,1	193,4	210,0	227,0	33,4	36,5	39,3	43,1	48,0	62,1	3,6	3,8	4,0	4,2	4,5	5,8
Zawodowe .....	126,1	153,5	168,0	166,7	182,4	203,5	29,0	30,8	34,6	37,9	44,2	54,9	2,8	3,1	3,3	3,5	3,7	5,3
Średnie .....	160,4	186,6	202,1	196,6	207,9	228,5	27,2	29,7	33,9	37,1	44,7	51,8	2,7	3,2	3,3	3,4	3,5	4,7
Wyższe .....	35,7	49,0	57,3	55,9	63,9	69,9	18,1	21,7	25,2	29,0	35,0	47,5	6,1	6,7	7,0	7,0	7,4	8,5

Uwaga. Jak przy tabl. 2.

Źródło: opracowanie własne na podstawie danych jednostkowych z BAEL oraz zagregowanych danych ze sprawozdań MPIPS-01.

## 5. Podsumowanie

W artykule przedstawiono wyniki estymacji poziomu bezrobocia biernego – zjawiska polegającego na rejestrowaniu się w PUP jako bezrobotne przez osoby, które nie poszukują pracy bądź nie są gotowe lub zdolne do jej podjęcia. Przeprowadzone badanie jest kontynuacją pracy podjętej przez Wilaka (2018), w której dokonano estymacji bezrobocia pozornego i biernego w Polsce w latach 2002–2011 w przekroju domen określonych przez cechy demograficzno-ekonomiczne, a także w układzie województw.

W badaniu omawianym w niniejszym opracowaniu szacunkom podlegały liczba biernie bezrobotnych i ich odsetek wśród osób zarejestrowanych w PUP jako bezrobotne. Estymację przeprowadzono dla populacji ludności w Polsce, a także w podziale na płeć, grupy wieku, miejsce zamieszkania i poziom wykształcenia, na podstawie danych jednostkowych z BAEL za lata 2010–2018. Kluczowa była identyfikacja osób biernie bezrobotnych w bazie danych jednostkowych z tego badania. Przeprowadzono ją, opierając się na definicji osoby biernie bezrobotnej, która wynika z definicji zalecanych przez MOP.

Estymację przeprowadzono z wykorzystaniem podejścia kalibracyjnego. Zastosowano cztery metody, przy użyciu których korygowano wagi. Na podstawie analizy wag i mnożników kalibracyjnych oraz przy uwzględnieniu szybkości działania algorytmu za optymalną uznano metodę logitową. Wykorzystanie dodatkowych zmiennych pomocniczych z danych na temat bezrobocia rejestrowanego, oprócz stosowanych w BAEL zmiennych pomocniczych o charakterze demograficznym, pozwoliło na otrzymanie szacunków o akceptowalnej precyzji.

Otrzymane szacunki wskazują na znaczny odsetek osób zarejestrowanych w PUP jako bezrobotne, które charakteryzują się biernością zawodową. W ostatnich latach, pomimo dużego spadku liczby osób zarejestrowanych jako bezrobotne, liczba biernie bezrobotnych spadła nieznacznie, co przekłada się na duży wzrost ich odsetka. W analizowanym okresie ogólna liczba osób charakteryzujących się biernym bezrobociem mieściła się między 497,5 tys. a 699,3 tys. Z kolei odsetek osób biernie bezrobotnych wśród zarejestrowanych w PUP jako bezrobotne wynosił od 29,8% do 53,3%, przy czym od roku 2014 obserwuje się jego znaczny wzrost.

Problem bezrobocia biernego w większym stopniu dotyczy kobiet niż mężczyzn. Zróżnicowanie poziomu odsetka biernie bezrobotnych obserwuje się także w podziale na grupy wieku i poziom wykształcenia. Grupami o największym odsetku biernie bezrobotnych są osoby w wieku 55 lat i więcej, a także osoby o wykształceniu co najwyżej gimnazjalnym.

Przeprowadzone badanie pokazuje skalę problemu bezrobocia biernego. W 2018 r. znaczna część osób zarejestrowanych w PUP jako bezrobotne – ponad 50% – nie poszukiwała aktywnie pracy, nie była zainteresowana jej znalezieniem bądź nie była gotowa jej podjąć. W związku z tym zasoby powiatowych urzędów pracy w dużym stopniu wykorzystywane były do obsługi biernych zawodowo. Szacunki przedstawione w artykule mogą być pomocne w podejmowaniu decyzji na temat kierunku polityki społecznej wobec bezrobotnych. Zaproponowana metoda estymacji bezrobocia biernego może być implementowana systematycznie w kolejnych kwartałach. Oznacza to, że szacunki uzyskane w wyniku jej zastosowania mogą być traktowane jako miernik zmian zachodzących na rynku pracy.

## Bibliografia

- Bruch, C., Münnich, R., Zins, S. (2011). *Variance estimation for complex surveys*. Pobrane z: [https://www.uni-trier.de/fileadmin/fb4/projekte/SurveyStatisticsNet/Ameli_Delivrables/AMELI-WP3-D3.1-20110514.pdf](https://www.uni-trier.de/fileadmin/fb4/projekte/SurveyStatisticsNet/Ameli_Delivrables/AMELI-WP3-D3.1-20110514.pdf).
- Czapiński, J., Panek, T. (2013). Wykluczenie społeczne. *Contemporary Economics*, 7, 342–375. DOI: 10.5709/ce.1897-9254.115.
- Deville, J.-C., Särndal, C.-E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 87(418), 376–382. DOI: 10.1080/01621459.1992.10475217.
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1–26. DOI: 10.1214/aos/1176344552.
- Góra, M., Socha, M. W., Sztanderska, U. (1995). *Analiza polskiego rynku pracy w latach 1990–1994: Kierunki zmian i rola polityk rynku pracy*. Warszawa: Główny Urząd Statystyczny.
- Góral, Z. (red.). (2011). *Ustawa o Promocji Zatrudnienia i Instytucjach Rynku Pracy. Praktyczny Komentarz*. Warszawa: Wolters Kluwer Polska.
- GUS. (2018a). *Zeszyt metodologiczny. Badanie Aktywności Ekonomicznej Ludności*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2018b). *Zeszyt metodologiczny. Statystyka rynku pracy i wynagrodzeń*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2019a). *Aktywność Ekonomiczna Ludności Polski. IV kwartał 2018*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2019b). *Bezrobocie rejestrowane. I–IV kwartał 2018 r.* Warszawa: Główny Urząd Statystyczny.
- GUS. (2019c). *Rynek pracy w 2017 r.* Warszawa: Główny Urząd Statystyczny.
- GUS. (2019d). *Wybrane aspekty rynku pracy w 2018 r.* Warszawa: Główny Urząd Statystyczny.
- Hryniewicka, A. (2011). Bezrobotni jako klienci pomocy społecznej. W: *Bezrobocie i bezrobotni – pomoc społeczna – urzędy pracy: raport wstępny* (s. 96–138). Warszawa: WRZOS.
- Hussmanns, R., Mehran, F., Verma, V. (1992). *Surveys of economically active population, employment, unemployment and underemployment: An ILO manual on concepts and methods*. Genewa: International Labour Office.
- Janukowicz, P. (2010). Bezrobocie rejestrowane a bezrobocie według BAEL. *Polityka Społeczna*, (1), 18–20.

- Kaczorowski, P., Kubiak, P., Kwiatkowski, E., Pawlega, S., Uścińska, G. (2009). Nowoczesne systemy zabezpieczenia społecznego. W: E. Kryńska (red.), *Flexicurity w Polsce: diagnoza i rekomendacje* (s. 93–134). Warszawa: na zlecenie MPiPS Oficyna Wydawnicza ASPRA-JR.
- Kolenikov, S. (2010). Resampling variance estimation for complex survey data. *The Stata Journal*, 10(2), 165–199. DOI: 10.1177/1536867X1001000201.
- Kotowska, I. E., Strzelecki, Z. (1993). *Bezrobocie z punktu widzenia gospodarstw domowych*. Warszawa: Główny Urząd Statystyczny.
- Kryńska, E. (1993). *Bezrobocie a segmentacja rynku pracy*. Warszawa: Główny Urząd Statystyczny.
- Kwiatkowski, E. (1992). Bezrobocie w Polsce w okresie transformacji: podstawowe tendencje i ich determinanty. *Ekonomista*, (3), 439–451.
- Kwiatkowski, E. (2005). *Bezrobocie: podstawy teoretyczne*. Warszawa: Wydawnictwo Naukowe PWN.
- Kwiatkowski, E. (2013). *Ekonomiczne aspekty zasiłków dla bezrobotnych*. Warszawa: Wolters Kluwer Polska.
- Lewandowski, R. (2006). *Znaczenie i rola urzędów pracy w polityce społecznej państwa w świetle polskich doświadczeń*. Szczecin: „Pedagogium” Wydawnictwo OR TWP.
- Popiołek, A., Piątkowska, M. (2013). *Urzędy pracy bez ubezpieczenia zdrowotnego? Czy bezrobotnych ubędzie?* Pobrane z: <http://wyborcza.pl/1,76842,13859547,Urzed-y-pracy-bez-ubezpieczenia-zdrowotnego-Czy-bezrobotnych.html>.
- Rao, J. N. K., Wu, C. F. J. (1988). Resampling Inference with Complex Survey Data. *Journal of the American Statistical Association*, 83(401), 231–241. DOI: 10.2307/2288945.
- Rao, J. N. K., Wu, C. F. J., Yue, K. (1992). Some Recent Work on Resampling Methods for Complex Surveys. *Survey Methodology*, 18(2), 209–217.
- Rzemek, M. (2014). *Bezrobotni: ubezpieczenie zdrowotne bez rejestracji w urzędzie pracy*. Pobrane z: <http://www.rp.pl/artukul/1105056-Bezrobotni-ubezpieczenie-zdrowotne-bez-rejestracji-w-urzedzie-pracy.html>.
- Särndal, C.-E., Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. Chichester: John Wiley & Sons.
- Staszewska, E. (2012). *Środki prawne przeciwdziałania bezrobociu*. Warszawa: Wolters Kluwer Polska.
- Szymkowiak, M. (2019). *Podejście kalibracyjne w badaniach społeczno-ekonomicznych*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego.
- Topolska, K. (2014). *Resort pracy ma pomysł na ubezpieczenie zdrowotne dla bezrobotnych*. Pobrane z: <http://praca.gazetaprawna.pl/artykuly/793576,resort-pracy-ma-pomysl-na-ubezpieczenie-zdrowotne-dla-bezrobotnych.html>.
- Wilak, K. (2018). *Estymacja bezrobocia pozornego i biernego z wykorzystaniem strukturalnych modeli szeregów czasowych* (niepublikowana rozprawa doktorska). Poznań: Uniwersytet Ekonomiczny.
- Witkowski, J. (1994). *Podstawowe cechy bezrobocia w Polsce w okresie transformacji*. Warszawa: Główny Urząd Statystyczny.
- Wojciechowska, R. (2006). Problem bezrobocia pozornego. W: J. Ostaszewski (red.), *Bezrobocie w Polsce – diagnoza sytuacji, pożądane kierunki w ograniczaniu bezrobocia* (s. 191–202). Warszawa: Oficyna Wydawnicza SGH.
- Zgierska, A. (2017). 25 lat Badania Aktywności Ekonomicznej Ludności w Polsce – kamienie milowe i perspektywy rozwoju. *Wiadomości Statystyczne*, (12), 23–49. DOI: 10.5604/01.3001.0014.1089.

## Zanieczyszczenie powietrza a stan zdrowia mieszkańców dużych miast województwa śląskiego

Zofia Mielecka-Kubieñ^a, Andrzej Wójcik^b

**Streszczenie.** Duże miasta woj. śląskiego, przede wszystkim na obszarze byłego Górnośląskiego Okręgu Przemysłowego, charakteryzują się m.in. wysokim stopniem uprzemysłowienia, wyższym niż średni w kraju poziomem życia, a także bardzo wysokim stopniem zanieczyszczenia środowiska. Celem badania omawianego w artykule jest ocena poziomu wybranych rodzajów zanieczyszczeń powietrza w dużych miastach woj. śląskiego oraz ich porównanie z niektórymi charakterystykami stanu zdrowia ludności tego województwa w latach 2014–2016. W badaniu wykorzystano dane Głównego Inspektoratu Ochrony Środowiska oraz Głównego Urzędu Statystycznego.

Zestawienie wartości wybranych mierników zdrowia ludności badanych miast z poziomem ich zanieczyszczenia pozwoliło stwierdzić, że w miastach o najwyższym poziomie zanieczyszczenia powietrza wartości mierników umieralności również były najwyższe, a wartości przeciętnego dalszego trwania życia noworodka – najniższe. Najgorszą sytuację pod względem obydwu czynników zaobserwowano w Chorzowie, Dąbrowie Górniczej i Rybniku, natomiast najlepszą – w Bielsku-Białej i Tychach.

**Słowa kluczowe:** zanieczyszczenie powietrza, umieralność, trwanie życia, województwo śląskie

**JEL:** J11, Q56

## Air pollution and health condition of inhabitants of big cities in Śląskie Voivodship

**Abstract.** Big cities in Śląskie Voivodship, especially those from the territory of the former Upper Silesian Industrial District, are characterised by a high level of industrialization, relatively high standard of living and very high level of environmental pollution. The aim of the study described in this paper is to assess the levels of selected types of air pollution in big cities in Śląskie Voivodship and to compare them against chosen reports on the health condition of the population of the voivodship in the years 2014–2016. The study was based on data from the Chief Inspectorate of Environmental Protection and Statistics Poland.

The comparison of selected indicators of the health condition of the populations in the studied cities against their levels of air pollution demonstrated that in cities where air pollution reached the highest levels, the values of indicators of mortality were also the highest, whereas the values of indicators of life expectancy of newborn babies were the lowest. The worst situation regarding both the air pollution and health condition was observed in Chorzów, Dąbrowa Górnicza and Rybnik, and the best in Bielsko-Biała and Tychy.

**Keywords:** air pollution, mortality, life expectation, Śląskie Voivodship

^a Uniwersytet Ekonomiczny w Katowicach, Kolegium Zarządzania.  
ORCID: <https://orcid.org/0000-0001-5348-4940>.

^b Uniwersytet Ekonomiczny w Katowicach, Kolegium Zarządzania.  
ORCID: <https://orcid.org/0000-0003-3803-8840>.



## 1. Wprowadzenie

Rozwój społeczno-gospodarczy, jaki dokonał się w XX i XXI w., wywarł nieodwracalny wpływ na stan środowiska naturalnego. Ekspansja przemysłu powoduje coraz większe zapotrzebowanie na energię elektryczną. Z kolei rozwój energetyki skutkuje intensywną eksploatacją paliw kopalnych. Prowadzi to do zanieczyszczenia gleby, wody oraz powietrza.

Degradacja środowiska naturalnego, w tym zanieczyszczenie powietrza, oraz jej wpływ na zdrowie stanowią obecnie, zarówno w Polsce, jak i na świecie, przedmiot częstych debat. Zgodnie z wynikami badań Greenstone'a i Qing Fang (2018) zanieczyszczenie powietrza jest obecnie główną przyczyną przedwczesnej umieralności, skracając w skali świata przeciętne dalsze trwanie życia średnio o 1,8 roku. Według Światowej Organizacji Zdrowia (WHO) z powodu zanieczyszczenia powietrza umiera przedwcześnie od 6 mln do 7 mln ludzi rocznie na świecie, przy czym w Polsce jest to od 40 tys. do 45 tys. zgonów, z których ok. 70% jest spowodowane przede wszystkim chorobami serca i układu krążenia¹. Źródłami zanieczyszczenia powietrza są zarówno przemysł (w tym energetyka oparta na węglu), jak i gospodarstwa domowe oraz transport drogowy.

Szczególnie niebezpieczne dla zdrowia są pyły PM_{2,5} i PM₁₀, których cząsteczki powstają głównie podczas spalania węgla i drewna oraz pracy silników Diesla (Greenstone i Qing Fang, 2018) i wnikają do dróg oddechowych razem z tlenem. Im mniejsze cząsteczki pyłów PM (ang. *particulate matter*) zawarte w sadzy, dymie, kurzu itp., tym są one groźniejsze. Cząsteczki PM o średnicy poniżej 5 µm docierają do oskrzelików, a cząstki o średnicy 2,5 µm lub mniejszej – do pęcherzyków płucnych. Zwiększają w ten sposób ryzyko zachorowania na raka, a dostawszy się do krwi, docierają do mózgu i serca, przyczyniając się do udarów mózgu i zawałów serca.

Według Lekstona:

Polscy eksperci podkreślają, że wśród grup najbardziej narażonych na zanieczyszczenia atmosfery są dzieci, ludzie w podeszłym wieku, kobiety ciężarne oraz osoby z chorobami dróg oddechowych. Wskazują, że długotrwała obecność trujących gazów w atmosferze wywołuje choroby naczyniowe, zawały serca, udary, astmę i alergie już na etapie życia płodowego. Dodatkowo w rejonach zanieczyszczeń zwiększa się liczba infekcji, które pociągają za sobą dalsze obciążenia dla gospodarki, m.in. w postaci wizyt u lekarza czy nieobecności w pracy. Wpływ zanieczyszczeń powietrza na występowanie chorób układu oddechowego, układu krążenia, udarów mózgu, nowotworów płuc oraz na zwiększenie umieralności w populacji ogólnej został dobrze udokumentowany (Lekston, 2017).

¹ Z wywiadu z prof. dr. hab. n. med. Pawłem Buszmanem, kardiologiem Śląskiego Uniwersytetu Medycznego (Airly, 2019).

Wpływ zanieczyszczenia powietrza na zdrowie został omówiony m.in. w pracach Janka (2014), Kuchcik i Milewskiego (2018) oraz Mazurka i Badydy (2018).

Celem badania omawianego w artykule jest ocena poziomu wybranych rodzajów zanieczyszczeń powietrza w dużych miastach woj. śląskiego oraz ich porównanie z niektórymi charakterystykami stanu zdrowia ludności tego województwa w latach 2014–2016.

## 2. Metoda badania

Badaniem objęto miasta woj. śląskiego liczące co najmniej 100 tys. mieszkańców: Bielsko-Białą, Bytom, Chorzów, Częstochowę, Dąbrowę Górniczą, Gliwice, Katowice, Rudę Śląską, Rybnik, Sosnowiec, Tychy i Zabrze. Dla ograniczenia zakresu wahań losowych w obliczeniach brano pod uwagę średnie wielkości z lat 2014–2016.

W Polsce poziom zanieczyszczenia powietrza atmosferycznego jest systematycznie monitorowany przez Główny Inspektorat Ochrony Środowiska (GIOŚ) w stacjach pomiarowych, a wynik pomiaru przekazywany do informacji publicznej na Portalu Jakości Powietrza. W woj. śląskim znajduje się 30 stacji pomiarowych², w tym 10 w miastach liczących powyżej 100 tys. mieszkańców: w Częstochowie i Katowicach po dwie stacje, zaś w Bielsku-Białej, Dąbrowie Górniczej, Rybniku, Sosnowcu, Tychach i Zabrzu po jednej stacji. W pozostałych miastach będących przedmiotem badania: Bytomiu, Chorzowie, Gliwicach i Rudzie Śląskiej nie ma takich stacji³.

Stacje pomiarowe w wymienionych miastach mierzą różne rodzaje zanieczyszczeń. Na przykład w Dąbrowie Górniczej mierzonych jest 10 rodzajów zanieczyszczeń, a w Gliwicach – jedynie pięć. W konsekwencji nie ma porównywalnych informacji dla miast będących przedmiotem badania, więc nie można zastosować w proponowanej analizie wyników otrzymanych w stacjach pomiarowych GIOŚ. Informacje z tego źródła można natomiast wykorzystać do porównania poziomu zanieczyszczenia powietrza w różnych aglomeracjach.

Drugim podstawowym źródłem danych dotyczących zanieczyszczenia powietrza są roczniki *Ochrona środowiska* wydawane przez Główny Urząd Statystyczny (GUS) oraz bazy danych GUS⁴. Dostępne tam dane dla dużych miast woj. śląskiego obejmują emisję zanieczyszczeń pyłowych i gazowych z zakładów szczególnie uciążliwych⁵.

Badanie emisji zanieczyszczeń jest problematyczne ze względu na ich przemieszczanie się. Z tego powodu ocenę poziomu zanieczyszczeń powietrza w danym miejscu należy traktować jako szacunkową.

² <https://powietrze.gios.gov.pl/pjp/maps/measuringstation>.

³ <http://powietrze.katowice.wios.gov.pl/stacje>.

⁴ <https://bdl.stat.gov.pl/BDL/dane/podgrup/teryt>.

⁵ Zakłady szczególnie uciążliwe dla czystości powietrza to tzw. punktowe źródła emisji zanieczyszczeń, do których zaliczono wszystkie jednostki organizacyjne ustalone przez ministra właściwego do spraw środowiska (<https://bdl.stat.gov.pl/BDL5/dane/podgrup/temat/9/216/1651>).

Zanieczyszczenie powietrza jest monitorowane także przez różne firmy prywatne (np. Airly, 2019), ale wyniki tych pomiarów mają charakter jedynie informacyjny. Ocena jakości powietrza i wprowadzanie progów alarmowych możliwe są tylko na podstawie danych pochodzących z Państwowego Monitoringu Środowiska.

Na potrzeby badania omawianego w niniejszym artykule wykorzystano ogólnodostępne dane GUS oraz wyniki pomiarów stacji GIOŚ dotyczące aglomeracji woj. śląskiego, co wymusiło zawężenie zakresu analizy do danych dotyczących emisji pyłowych i gazowych z zakładów szczególnie uciążliwych na terenie dużych miast województwa. Emisja gazów obejmuje: wartości ogółem (bez dwutlenku węgla), emisję niezorganizowaną⁶ oraz emisję dwutlenku siarki, tlenków azotu, tlenku węgla, dwutlenku węgla, metanu i podtlenku azotu. W przypadku emisji pyłów nie wyróżniono ich rodzajów. Dla badanych miast nie ma opublikowanych informacji na temat innych rodzajów emisji, przede wszystkim emisji niskiej (spaliny emitowane z palenisk domowych) oraz emisji zanieczyszczeń ze środków transportu drogowego. Wielkość emisji gazów oraz pyłów odniesiono do powierzchni miast (Kukuła, 2019, s. 4).

Stan zdrowia ludności Polski jest systematycznie monitorowany przez GUS (np. GUS, 2016b, 2017b), lecz dane te są zbyt ogólne dla potrzeb proponowanej analizy; brakuje w nich wyszczególnienia miast będących przedmiotem badania. Współczynniki zachorowalności na wybrane choroby dla badanych miast podaje Śląski Urząd Wojewódzki w Katowicach (ŚUW) w raportach o stanie zdrowia mieszkańców woj. śląskiego. Są to jednak współczynniki niestandardyzowane według wieku (np. ŚUW, 2014), co ogranicza ich wartość poznawczą i uniemożliwia porównywanie poziomu zachorowalności w różnych populacjach. Szczegółowe dane dotyczące zachorowalności w badanych miastach nie są ogólnie dostępne.

W badaniu omawianym w niniejszym artykule stan zdrowia populacji miast woj. śląskiego liczących 100 tys. i więcej mieszkańców scharakteryzowano poprzez przeciętne dalsze trwanie życia noworodka i negatywne charakterystyki zdrowia, takie jak współczynnik zgonów niemowląt, oraz standaryzowane ze względu na wiek⁷ współczynniki zgonów: ogólny, z powodu chorób układu krążenia (według klasyfikacji ICD-10 I00-I99), nowotworów (C00-C97) i chorób układu oddechowego (J00-J99).

Źródłem danych wykorzystanych do oceny stanu zdrowia ludności w dużych miastach woj. śląskiego była Baza Demografia GUS⁸.

⁶ Emisja niezorganizowana to wprowadzenie do powietrza zanieczyszczeń w sposób niezorganizowany: z hałd, wysypisk, podczas przeładunku substancji sypkich lub lotnych, z hal produkcyjnych, poprzez wentylatory dachowe i okienne, w wyniku pożarów lasów itp. (Maciejowski, 1995).

⁷ Standaryzacja współczynników demograficznych została omówiona m.in. w pracy Kurkiewicz (2010).

⁸ <http://demografia.stat.gov.pl/bazademografia/Tables.aspx>.

### 3. Zanieczyszczenie powietrza w dużych miastach województwa śląskiego

Wyczerpujące omówienie stanu środowiska naturalnego w Polsce można znaleźć w opracowaniu GIOŚ (2017). Z kolei w pracy Mieleckiej-Kubień (2017) na podstawie analizy zanieczyszczenia środowiska i stanu zdrowia mieszkańców w województwach w roku 2013 wyciągnięto wniosek, że zanieczyszczenie środowiska w woj. śląskim jest znacznie wyższe niż w jakimkolwiek innym województwie. Również przedstawione tam charakterystyki stanu zdrowia ludności mają w przypadku woj. śląskiego wyjątkowo niekorzystne wartości. Wysoki poziom zanieczyszczenia środowiska naturalnego tego województwa potwierdzają dane statystyczne za lata 2014–2016 (tabl. 1) obejmujące emisję zanieczyszczeń pyłowych i gazowych z zakładów szczególnie uciążliwych.

**Tabl. 1.** Średnia emisja zanieczyszczeń pyłowych i gazowych z zakładów szczególnie uciążliwych w latach 2014–2016 według województw

Wyszczególnienie	Zanieczyszczenia pyłowe ogółem	Zanieczyszczenia gazowe				
		ogółem	dwutlenek węgla	dwutlenek siarki	tlenek węgla	tlenki azotu
w t/km ²						
Polska .....	0,14	673,2	668,3	1,1	1,1	0,8
Dolnośląskie .....	0,14	653,8	651,6	1,1	0,4	0,6
Kujawsko-pomorskie .....	0,15	478,7	475,9	0,9	0,8	0,7
Lubelskie .....	0,07	199,9	199,0	0,3	0,2	0,3
Lubuskie .....	0,07	147,6	146,4	0,2	0,7	0,3
Łódzkie .....	0,14	2274,9	2267,2	3,8	1,5	2,3
Małopolskie .....	0,17	683,6	677,2	1,6	1,0	1,2
Mazowieckie .....	0,11	804,1	800,7	1,7	0,5	1,0
Opolskie .....	0,17	1339,1	1333,8	1,0	2,4	1,6
Podkarpackie .....	0,08	156,6	155,7	0,3	0,2	0,3
Podlaskie .....	0,04	102,4	101,9	0,1	0,2	0,1
Pomorskie .....	0,12	361,4	360,0	0,6	0,3	0,4
Śląskie .....	0,80	3099,6	3038,9	4,9	12,8	3,7
Świętokrzyskie .....	0,17	1099,8	1092,8	1,4	3,5	1,8
Warmińsko-mazurskie .....	0,04	62,0	61,6	0,2	0,1	0,1
Wielkopolskie .....	0,16	537,1	534,6	0,9	0,3	0,8
Zachodniopomorskie .....	0,11	377,3	376,0	0,4	0,1	0,4

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

Jak wynika z danych zawartych w tabl. 1, poziom zanieczyszczenia powietrza w woj. śląskim pod względem emisji zanieczyszczeń pyłowych i gazowych z zakładów szczególnie uciążliwych znacznie przekraczał w analizowanym okresie średni poziom krajowy.

Poziom zanieczyszczenia powietrza w woj. śląskim cechuje się znacznym zróżnicowaniem przestrzennym (RWMS GIOŚ, 2019; WIOŚ, 2017). Podane w tabl. 1 in-

formacje o zanieczyszczeniu powietrza są wartościami średnimi dla całych województw, w przypadku woj. śląskiego obejmują więc, oprócz cechujących się wysokim stopniem zanieczyszczenia powietrza obszarów uprzemysłowionych, także znacznie mniej zanieczyszczone tereny zielone w okolicach Bielska-Białej czy Tychów. Można przypuszczać, że szczególnie niekorzystna sytuacja pod tym względem występuje na obszarze byłego Górnośląskiego Okręgu Przemysłowego (GOP) oraz byłego Rybnickiego Okręgu Węglowego (ROW) (Dziembała, 1991).

W rocznikach *Ochrona środowiska* (GUS, 2015, 2016a, 2017a) zamieszczono informacje o zanieczyszczeniu powietrza w różnych aglomeracjach, otrzymane na podstawie wyników uzyskanych w stacjach pomiarowych GIOŚ. Dwie z tych aglomeracji – górnośląska i rybnicko-jastrzębska – mieszczą się w woj. śląskim. Aglomeracja górnośląska obejmuje 19 graniczących ze sobą miast: Gliwice, Zabrze, Katowice, Bytom, Świętochłowice, Siemianowice Śląskie, Sosnowiec, Dąbrowę Górniczą, Jaworzno, Czeladź, Mysłowice, Będzin, Tychy, Rudę Śląską, Piekary Śląskie, Chorzów, Mikołów, Tarnowskie Góry i Knurów. Są to w większości silnie uprzemysłowione miasta byłego GOP. Aglomeracja rybnicko-jastrzębska obejmuje głównie miasta byłego ROW: Jastrzębie-Zdrój, Pszów, Radlin, Rybnik, Rydułtowy, Wodzisław Śląski i Żory.

W latach 2014–2016 średnie roczne stężenie pyłu PM_{2,5} ( $\mu\text{g}/\text{m}^3$ ) przybrało najwyższy poziom w aglomeracji krakowskiej ( $35,3 \mu\text{g}/\text{m}^3$ ), a następnie w aglomeracji górnośląskiej ( $32,3 \mu\text{g}/\text{m}^3$ ) i rybnicko-jastrzębskiej ( $28,0 \mu\text{g}/\text{m}^3$ ), przy poziomie dopuszczalnym według polskich norm⁹  $25 \mu\text{g}/\text{m}^3$  (GUS, 2015, tabl. 36 i 37, 2016a, tabl. 37, 2017a, tabl. 35). Najwyższe stężenie pyłu PM₁₀ stwierdzono w aglomeracji krakowskiej ( $50,6 \mu\text{g}/\text{m}^3$ ), a następnie rybnicko-jastrzębskiej ( $45,4 \mu\text{g}/\text{m}^3$ ) i górnośląskiej ( $42,6 \mu\text{g}/\text{m}^3$ ), przy dopuszczalnym poziomie  $40 \mu\text{g}/\text{m}^3$ . W tych aglomeracjach zaobserwowano szczególnie wysokie stężenie dwutlenku azotu, dwutlenku siarki, tlenku węgla, benzenu, ołowiu, arsenu i kadmu. Informacje te potwierdzają występowanie szczególnie niekorzystnej sytuacji pod względem zanieczyszczenia powietrza w woj. śląskim.

Dane statystyczne na temat zanieczyszczenia powietrza dostępne dla 12 miast woj. śląskiego liczących 100 tys. i więcej mieszkańców, z których tylko dwa: Bielsko-Biała i Rybnik nie należą do aglomeracji górnośląskiej, dotyczą – jak już podano wcześniej – poziomu emisji gazów i pyłów z zakładów szczególnie uciążliwych. Wartości te w analizowanych latach 2014–2016 w badanych miastach na tle województwa i kraju przedstawia tabl. 2.

⁹ W niektórych krajach normy te są znacznie niższe (Greenstone i Qing Fang, 2018, s. 30–31), a wielkości norm zalecane przez WHO (2006, s. 9) to średniorocznie  $10,0 \mu\text{g}/\text{m}^3$  dla PM_{2,5} oraz  $20,0 \mu\text{g}/\text{m}^3$  dla PM₁₀, „[...] zatem obecne normy nie odzwierciedlają prawdziwego, negatywnego oddziaływania tych substancji na organizm człowieka” (Kuchcik i Milewski, 2018, s. 7). Występowanie w długim okresie średniorocznych wartości PM_{2,5} równych  $35,0 \mu\text{g}/\text{m}^3$  oraz PM₁₀ równych  $70,0 \mu\text{g}/\text{m}^3$  zwiększa ryzyko zgonu o 15% w porównaniu z poziomem tych pyłów zgodnym z normami (WHO, 2006, s. 11).

**Tabl. 2.** Średnioroczna emisja zanieczyszczeń gazowych i pyłowych z zakładów szczególnie uciążliwych w miastach woj. śląskiego liczących 100 tys. i więcej mieszkańców w latach 2014–2016

Przekrój badania	Emisja gazów							Emisja pyłów			
	ogółem	ogółem bez dwutlenku węgla	dwutlenek węgla	metan	dwutlenek siarki	tlenek węgla	tlenki azotu	niezorganizowana	podtlenek azotu	Emisja pyłów	
				<b>W tys. t</b>							
Polska .....	210494,4	1535,5	208958,9	530,4	351,4	344,7	263,0	134,4	5,0	43418,0	
Woj. śląskie .....	38227,8	748,8	37479,0	474,6	61,1	158,4	45,8	75,7	1,4	9853,0	
Bielsko-Biała .....	236,1	0,7	235,4	0,0	0,2	0,1	0,2	0,0	0,0	71,0	
Bytom .....	302,7	2,6	300,1	0,0	1,4	0,5	0,6	0,0	0,0	208,3	
Chorzów .....	1610,0	5,0	1605,0	0,0	1,7	0,3	1,2	0,0	0,2	103,0	
Częstochowa .....	887,3	5,3	882,0	0,0	0,9	2,3	2,1	0,9	0,0	261,3	
Dąbrowa Górnicza .....	9564,0	156,3	9407,7	0,0	10,0	137,2	8,6	3,7	0,0	4170,7	
Gliwice .....	435,7	24,4	411,3	20,4	1,5	0,9	0,6	4,4	0,0	212,7	
Katowice .....	1073,4	56,2	1017,2	51,8	2,3	0,5	0,8	0,0	0,5	268,0	
Ruda Śląska .....	208,9	47,7	161,2	46,9	0,6	0,1	0,2	0,0	0,0	207,7	
Rybnik .....	7279,7	54,5	7225,3	22,0	17,1	3,6	10,8	0,0	0,1	953,0	
Sosnowiec .....	133,5	0,7	132,8	0,0	0,4	0,2	0,2	0,0	0,0	84,3	
Tychy .....	308,6	2,4	306,2	0,0	0,7	0,3	0,6	5,1	0,0	143,7	
Zabrze .....	334,0	7,5	326,5	4,5	1,4	0,4	1,0	31,0	0,0	161,7	
				<b>W t/km²</b>							
Polska .....	673,20	4,91	668,29	1,70	1,12	1,10	0,84	0,43	0,02	0,14	
Woj. śląskie .....	3099,64	6072	3038,92	3848	4,95	12,85	3,71	6,14	0,11	0,80	
Bielsko-Biała .....	1889,02	5,73	1883,30	0,0	1,54	1,14	1,58	0,19	0,23	0,57	
Bytom .....	4386,35	37,61	4348,74	0,27	20,64	7,50	8,91	0,39	0,01	3,02	
Chorzów .....	48786,93	150,93	48636,00	0,10	50,00	9,69	35,78	0,14	6,86	3,12	
Częstochowa .....	5545,44	33,02	5512,42	0,01	5,61	14,23	12,86	5,43	0,0	1,63	
Dąbrowa Górnicza .....	50602,96	826,83	49776,13	0,18	52,75	726,11	45,60	19,38	0,0	22,07	
Gliwice .....	3251,42	181,99	3069,43	152,06	11,32	6,79	4,29	32,86	0,02	1,59	
Katowice .....	6505,38	340,83	6164,55	314,16	13,90	3,18	5,03	0,0	3,19	1,62	
Ruda Śląska .....	2678,82	611,98	2066,84	600,76	7,09	1,61	2,44	0,0	0,0	2,66	
Rybnik .....	49187,26	367,97	48819,29	148,74	115,50	24,58	73,25	0,0	0,97	6,44	
Sosnowiec .....	1467,17	8,10	1459,08	0,0	4,18	1,86	1,73	0,05	0,0	0,93	
Tychy .....	3763,68	29,50	3734,17	0,0	8,75	3,92	6,88	62,11	0,24	1,75	
Zabrze .....	4174,63	93,56	4081,07	56,15	17,55	4,65	12,88	387,57	0,0	2,02	

Źródło: obliczenia własne na podstawie danych GUS (<https://bdl.stat.gov.pl/BDL/dane/podgrup/teryt>).

W większości badanych miast poziom emisji znacznie przekracza poziom krajowy; wyjątkiem jest emisja niezorganizowana oraz emisja metanu i podtlenku azotu w niektórych miastach. W analizowanym okresie w woj. śląskim wyemitowano średnio 18,2% ogółu krajowej emisji gazów i 22,7% pyłów. Warto zwrócić uwagę, że udział powierzchni woj. śląskiego w powierzchni kraju to ok. 3,9%. Poziom emisji gazów i pyłów jest więc w tym województwie wielokrotnie wyższy niż średnia krajowa.

Pod względem stosunku poziomu emisji gazów i pyłów do powierzchni miast w najgorszej sytuacji są Dąbrowa Górnicza, Chorzów i Rybnik, gdzie poziom emisji zanieczyszczeń w przeliczeniu na 1 km² był w analizowanym okresie wielokrotnie wyższy niż średnio w kraju i w całym województwie. Przykładowo emisja gazów ogółem (średniorocznie w t/km²) w Dąbrowie Górniczej była ponad 75 razy wyższa niż średnio w kraju, a emisja pyłów (średniorocznie w t/km²) ponad 157 razy wyższa niż w kraju. Należy przy tym wziąć pod uwagę, że do średniej krajowej wliczono zanieczyszczenia w omawianych miastach, zatem różnice między skrajnymi ośrodkami są jeszcze większe.

Można także zauważyć, że emisja gazów w Dąbrowie Górniczej stanowiła w badanym okresie 25,0% całej emisji gazów w woj. śląskim i 4,5% tej emisji w kraju, a emisja pyłów – 42,3% całej emisji pyłów w tym województwie i 9,6% emisji krajowej.

Najlepszą sytuację zaobserwowano w Bielsku-Białej, a w dalszej kolejności – w Sosnowcu i Tychach, gdzie poziom emitowanych zanieczyszczeń był na ogół o wiele wyższy niż średnia w kraju, ale zazwyczaj kształtował się poniżej średniej dla woj. śląskiego.

Warto podkreślić, że w Chorzowie, będącym od lat jednym z najbardziej zagrożonych ekologicznie miast w Polsce (Dziembała, 1991), nie ma stacji pomiarowej GIOŚ.

#### **4. Stan zdrowia mieszkańców dużych miast województwa śląskiego w świetle wybranych charakterystyk**

Analiza wartości przeciętnego dalszego trwania życia noworodka pokazuje (tabl. 3), że w populacji mężczyzn najwyższe wartości – przewyższające średnią krajową i średnią dla ogółu miast Polski – odnotowano w Bielsku-Białej, Tychach i Zabrze. W pozostałych miastach woj. śląskiego odpowiednie wartości okazały się niższe, a najniższe zaobserwowano w Chorzowie i Rudzie Śląskiej. Różnica między wartością najwyższą (Bielsko-Biała) i najniższą (Chorzów) wśród badanych miast wyniosła blisko cztery lata.

W populacji kobiet przeciętne dalsze trwanie życia noworodka w badanych miastach woj. śląskiego miało w każdym przypadku niższą wartość niż w kraju oraz w miastach Polski. Najniższą zaobserwowano w Chorzowie, Rudzie Śląskiej i Bytomiu, natomiast najwyższą w Bielsku-Białej. Różnica między wartością najwyższą (Bielsko-Biała) a najniższą (Chorzów) wyniosła ponad trzy lata¹⁰.

¹⁰ Podobne różnice w długości przeciętnego dalszego trwania życia w Chorzowie i Bielsku-Białej w populacjach mężczyzn i kobiet są obserwowane od wielu lat, (Mielecka-Kubień, 2004).

**Tabl. 3.** Przeciętne dalsze trwanie życia noworodka w miastach woj. śląskiego liczących 100 tys. i więcej mieszkańców w latach 2014–2016 według płci

Wyszczególnienie	Mężczyźni	Kobiety
Polska .....	73,5	81,0
Polska – wyłącznie miasta .....	73,9	80,9
Bielsko-Biała .....	74,5	80,0
Bytom .....	71,3	77,7
Chorzów .....	70,6	76,9
Częstochowa .....	72,2	79,0
Dąbrowa Górnicza .....	72,0	78,8
Gliwice .....	73,9	79,2
Katowice .....	73,0	78,8
Ruda Śląska .....	70,8	77,3
Rybnik .....	73,1	79,3
Sosnowiec .....	71,3	78,3
Tychy .....	74,2	79,7
Zabrze .....	74,0	79,5

Źródło: obliczenia własne za pomocą programu MORTPACK for Windows, Version 4.3, United Nations, New York na podstawie danych GUS (<http://demografia.stat.gov.pl/bazademografia>).

Współczynnik zgonów niemowląt (na 1000 urodzeń żywych) przybrał w badanym okresie (tabl. 4) najniższą wartość w Tychach i Bielsku-Białej, a najwyższą w Chorzowie i Rybniku. Różnica między wartością najwyższą a najniższą wyniosła 3,4, co oznacza, że na każde 10 tys. urodzeń żywych zmarło w tym czasie w Tychach 31 niemowląt, w Chorzowie natomiast 65, czyli ponaddwukrotnie więcej.

**Tabl. 4.** Współczynnik zgonów niemowląt (na 1000 urodzeń żywych) w miastach woj. śląskiego liczących 100 tys. i więcej mieszkańców w latach 2014–2016

Przekrój badania	Współczynnik zgonów niemowląt	Relacja względem	
		miast ogółem	Tychów
		w %	
Polska .....	4,2	2,4	35,5
Polska – wyłącznie miasta .....	4,1	.	32,3
Bielsko-Biała .....	3,8	-7,3	22,6
Bytom .....	5,0	22,0	61,3
Chorzów .....	6,5	58,5	109,7
Częstochowa .....	3,9	-4,9	25,8
Dąbrowa Górnicza .....	3,9	-4,9	25,8
Gliwice .....	3,9	-4,9	25,8
Katowice .....	4,2	2,4	35,5
Ruda Śląska .....	5,4	31,7	74,2
Rybnik .....	5,7	39,0	83,9
Sosnowiec .....	5,3	29,3	71,0
Tychy .....	3,1	-24,4	.
Zabrze .....	5,7	39,0	83,9

Źródło: obliczenia własne na podstawie danych GUS (<http://demografia.stat.gov.pl/bazademografia>).



Podobne prawidłowości można zaobserwować, analizując standaryzowane współczynniki zgonów ogółem oraz ze względu na wybrane przyczyny zgonów. W tabl. 5 przedstawiono wartości standaryzowanego współczynnika zgonów ogółem oraz z powodu niektórych chorób: nowotworowych, układu krążenia i układu oddechowego dla mężczyzn i kobiet.

**Tabl. 5.** Standaryzowany współczynnik zgonów z powodu wybranych przyczyn w miastach woj. śląskiego liczących 100 tys. i więcej mieszkańców w latach 2014–2016 według płci

Przekrój badania	Ogółem	Nowotwory	Choroby układu	
			krążenia	oddechowego
<b>Mężczyźni</b>				
Polska .....	0,0106	0,0030	0,0041	0,0007
Polska – wyłącznie miasta .....	0,0102	0,0030	0,0039	0,0006
Bielsko-Biała .....	0,0104	0,0032	0,0046	0,0004
Bytom .....	0,0125	0,0034	0,0048	0,0005
Chorzów .....	0,0133	0,0036	0,0053	0,0004
Częstochowa .....	0,0118	0,0031	0,0050	0,0005
Dąbrowa Górnicza .....	0,0126	0,0033	0,0057	0,0006
Gliwice .....	0,0106	0,0033	0,0039	0,0004
Katowice .....	0,0111	0,0033	0,0044	0,0004
Ruda Śląska .....	0,0132	0,0038	0,0049	0,0007
Rybnik .....	0,0111	0,0030	0,0046	0,0011
Sosnowiec .....	0,0127	0,0037	0,0056	0,0005
Tychy .....	0,0107	0,0033	0,0043	0,0005
Zabrze .....	0,0102	0,0029	0,0034	0,0005
<b>Kobiety</b>				
Polska .....	0,0053	0,0017	0,0022	0,0003
Polska – wyłącznie miasta .....	0,0053	0,0017	0,0021	0,0003
Bielsko-Biała .....	0,0061	0,0018	0,0031	0,0002
Bytom .....	0,0075	0,0022	0,0032	0,0002
Chorzów .....	0,0083	0,0024	0,0036	0,0003
Częstochowa .....	0,0069	0,0020	0,0032	0,0002
Dąbrowa Górnicza .....	0,0070	0,0018	0,0036	0,0003
Gliwice .....	0,0066	0,0021	0,0028	0,0002
Katowice .....	0,0067	0,0020	0,0029	0,0002
Ruda Śląska .....	0,0081	0,0023	0,0036	0,0002
Rybnik .....	0,0067	0,0017	0,0031	0,0006
Sosnowiec .....	0,0072	0,0019	0,0037	0,0003
Tychy .....	0,0063	0,0019	0,0030	0,0002
Zabrze .....	0,0063	0,0019	0,0025	0,0002

Źródło: obliczenia własne na podstawie danych GUS (<http://demografia.stat.gov.pl/bazademografia>).

W populacji mężczyzn najwyższy współczynnik zgonów ogółem wystąpił w Chorzowie, Rudzie Śląskiej i Dąbrowie Górniczej, podczas gdy najniższy – w Zabrze, Bielsku-Białej, Gliwicach i Tychach. Przy założonej standardowej strukturze wieku na każde 100 tys. mężczyzn w analizowanym okresie zmarło w Chorzowie 1330 mężczyzn, a w Zabrze – 1020, tj. o 310 osób mniej. Standaryzowany współczynnik

zgonów z powodu nowotworów osiągnął najwyższe wartości w Rudzie Śląskiej, Sosnowcu i Chorzowie, a z powodu chorób układu krążenia – w Dąbrowie Górniczej, Sosnowcu i Chorzowie. Pod względem zgonów na choroby układu oddechowego najgorszą sytuację zaobserwowano w Rybniku; w pozostałych miastach poziom zgonów z tej przyczyny nie odbiegał lub był nieco niższy od średniej w kraju.

W populacji kobiet w żadnym z wymienionych w tabl. 5 miast woj. śląskiego standaryzowane współczynniki zgonów ogółem oraz z powodu chorób układu krążenia i nowotworów nie były niższe od średniej w kraju lub w miastach Polski. Najwyższą wartość ogólnego współczynnika zgonów zaobserwowano w Chorzowie i Rudzie Śląskiej – przy założonej standardowej strukturze wieku średnio na 100 tys. kobiet w analizowanym okresie umarło w Chorzowie 830 kobiet, podczas gdy średnio w kraju – 530, tj. o 300 kobiet mniej. Spośród badanych miast najniższa wartość tego współczynnika wystąpiła w Bielsku-Białej, Tychach i Zabrze, lecz nawet tam umarło o ponad 100 kobiet na każde 100 tys. więcej niż w kraju i miastach Polski. Standaryzowany współczynnik zgonów z powodu nowotworów był najwyższy w Chorzowie i Rudzie Śląskiej, a z powodu chorób układu krążenia – w Sosnowcu, Chorzowie, Dąbrowie Górniczej i Rudzie Śląskiej. Podobnie jak w populacji mężczyzn współczynnik zgonów kobiet z powodu chorób układu oddechowego przybrał szczególnie wysoką wartość w Rybniku, gdzie przy założeniu standardowej struktury wieku w badanym okresie zmarło z tej przyczyny dwukrotnie więcej kobiet niż średnio w kraju.

## 5. Podsumowanie

Stan zdrowia populacji zależy od wielu czynników. Do najważniejszych należą: środowisko społeczno-ekonomiczne, środowisko naturalne oraz zachowania indywidualne. Jedną z determinant zdrowia jest zanieczyszczenie środowiska, w tym powietrza atmosferycznego. Wyniki wielu badań zarówno w Polsce, jak i za granicą wskazują na istnienie związku pomiędzy poziomem zanieczyszczenia powietrza a stanem zdrowia ludności.

Zgodnie z założonym celem badania dokonano oceny poziomu wybranych rodzajów zanieczyszczeń powietrza w dużych miastach woj. śląskiego oraz niektórych charakterystyk stanu zdrowia ludności w tych miastach, co pozwoliło na zaobserwowanie statystycznego związku pomiędzy miernikami obu rodzajów zjawisk. Miasta o najbardziej zanieczyszczonym powietrzu atmosferycznym: Chorzów, Dąbrowa Górnicza i Rybnik najczęściej cechowały się najwyższymi wartościami mierników umieralności i najniższymi wartościami przeciętnego dalszego trwania życia noworodka. Natomiast Bielsko-Biała i Tychy, leżące na terenach zielonych woj. śląskiego i najmniej zanieczyszczone spośród dużych miast w tym województwie, charakteryzowały się najkorzystniejszą sytuacją zdrowotną mieszkańców. Przedstawione wyniki wskazują zatem

na znaczne zróżnicowanie przestrzenne woj. śląskiego zarówno pod względem poziomu zanieczyszczenia powietrza, jak i stanu zdrowia mieszkańców.

Statystyczny związek pomiędzy poziomem zanieczyszczenia powietrza i wybranymi charakterystykami stanu zdrowia populacji wyraźnie uwidocznił się w przypadku przeciętnego dalszego trwania życia noworodka, współczynnika zgonów niemowląt, współczynnika zgonów z powodu chorób układu krążenia i nowotworów, a w mniejszym stopniu – z powodu chorób układu oddechowego. Można przypuszczać, że w tym ostatnim przypadku wpływ zanieczyszczenia powietrza byłby bardziej widoczny w badaniach zachorowalności niż umieralności.

Intencją autorów artykułu było zwrócenie uwagi na niepokojącą sytuację mieszkańców niektórych miast woj. śląskiego. Wyniki przeprowadzonej analizy wskazują, że część miast, przede wszystkim Chorzów, Dąbrowa Górnicza i Rybnik, cechuje bardzo wysoki poziom zanieczyszczenia powietrza. Uzupełnienie informacji o dane na temat niskiej emisji i emisji zanieczyszczeń ze środków transportu drogowego oraz innych rodzajów zanieczyszczeń prawdopodobnie jeszcze pogorszyłoby tę ocenę. Szacunkom tym towarzyszą bardzo niekorzystne wartości wybranych charakterystyk stanu zdrowia ludności. Jeżeli w dwóch miastach leżących w tym samym województwie, które dzieli odległość 53 km, różnica w długości przeciętnego dalszego trwania życia noworodka wynosi blisko cztery lata, muszą istnieć ważne przyczyny wystąpienia takiej sytuacji. Jedną z nich jest zapewne zanieczyszczenie środowiska naturalnego, w tym powietrza.

Wskazane byłoby bardziej zaawansowane monitorowanie zanieczyszczenia powietrza i ogólnie środowiska naturalnego w miastach najbardziej zagrożonych zanieczyszczeniami. Z powodu ograniczonej dostępności danych statystycznych przedmiotem badania były jedynie miasta liczące 100 tys. i więcej mieszkańców, co nie oznacza, że podobne problemy nie występują w mniejszych miejscowościach woj. śląskiego. Interesującym uzupełnieniem monitoringu jakości powietrza byłoby zbadanie sposobu migracji zanieczyszczeń emitowanych z zakładów szczególnie uciążliwych.

## Bibliografia

- Airly. (2019). *#Oddychaj Polsko. Raport o stanie powietrza w Polsce w latach 2017/2018 oraz 2018/2019*. Pobrane z: <https://airly.eu/pl/report/>.
- Dziembała, L. J. (1991). *Badanie zagrożenia życia na obszarze kłęski ekologicznej: studium statystyczno-demograficzne na przykładzie województwa katowickiego*. Katowice: Akademia Ekonomiczna.
- GIOŚ. (2017). *Stan środowiska w Polsce. Sygnały 2016*. Warszawa: Główny Inspektorat Ochrony Środowiska.
- Greenstone, M., Qing Fang, C. (2018). *AQLI. Introducing the Air Quality Life Index*. Pobrane z: <https://aqli.epic.uchicago.edu/wp-content/uploads/2018/11/AQLI-Report.111918-2.pdf>.

- GUS. (2015). *Ochrona środowiska 2015*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2016a). *Ochrona środowiska 2016*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2016b). *Stan zdrowia ludności Polski w 2014 r.* Warszawa: Główny Urząd Statystyczny.  
Pobrane z: <https://stat.gov.pl/obszary-tematyczne/zdrowie/zdrowie/stan-zdrowia-ludnosci-polski-w-2014-r-,6,6.html>.
- GUS. (2017a). *Ochrona środowiska 2017*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2017b). *Zdrowie i ochrona zdrowia w 2016 r.* Warszawa: Główny Urząd Statystyczny.  
Pobrane z: <https://stat.gov.pl/obszary-tematyczne/zdrowie/zdrowie/zdrowie-i-ochrona-zdrowia-w-2016-r-,1,7.html>.
- Janka, R. M. (2014). *Zanieczyszczenia pyłowe i gazowe: podstawy obliczenia i sterowania poziomem emisji*. Warszawa: Wydawnictwo Naukowe PWN.
- Kuchcik, M., Milewski, P. (2018). Zanieczyszczenie powietrza w Polsce – stan, przyczyny i skutki. *Studia KPZK*, 182(2), 341–364.
- Kukuła, K. (2019). Degree of environmental pollution in Poland – ranking of the provinces in 2016: Stopień zanieczyszczenia środowiska w Polsce – ranking województw w 2016 roku. *Economic and Regional Studies*, 12(1), 23–32. DOI: 10.2478/ers-2019-0003.
- Kurkiewicz, J. (red.). (2010). *Procesy demograficzne i metody ich analizy*. Kraków: Uniwersytet Ekonomiczny.
- Lekston, A. (2017). Zanieczyszczenie powietrza – jak ograniczyć skutki zdrowotne. Pobrane z: <https://www.mp.pl/pacjent/pulmonologia/aktualnosci/177874,zanieczyszczenie-powietrza-jak-ograniczyc-skutki-zdrowotne>.
- Maciejowski, M. (red.). (1995). *Leksykon Ochrony Środowiska ze słownikami polsko-angielskim, angielsko-polskim, polsko-niemieckim, niemiecko-polskim*. Gdańsk: Fundacja Ecobaltic.
- Mazurek, H., Badyda, A. (red.). (2018). *Smog. Konsekwencje zdrowotne zanieczyszczeń powietrza*. Warszawa: Wydawnictwo Lekarskie PZWL.
- Mielecka-Kubień, Z. (2004). O niektórych problemach mierzenia i porównywania umieralności (na przykładzie umieralności w Chorzowie i Bielsku-Białej w latach 1999–2001). W: A. S. Barczak (red.), *Postępy ekonometrii* (s. 165–177). Katowice: Akademia Ekonomiczna.
- Mielecka-Kubień, Z. (2017). Stan zdrowia mieszkańców a zanieczyszczenie środowiska w układzie województw. Analiza statystyczna. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, (318), 53–68.
- RWMŚ GIOŚ. (2019). *Pięcioletnia ocena jakości powietrza w województwie śląskim: Raport wojewódzki za lata 2014–2018*. Katowice: Główny Inspektorat Ochrony Środowiska, Regionalny Wydział Monitoringu Środowiska.
- ŚUW. (2014). *Raport o stanie zdrowia mieszkańców województwa śląskiego 2013*. Katowice: Śląski Urząd Wojewódzki.
- WHO. (2006). *WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: Global update 2005: Summary of risk assessment*. Geneva: World Health Organization.
- WIOŚ. (2017). *Stan środowiska w województwie śląskim w 2016 r.* Katowice: Wojewódzki Inspektorat Ochrony Środowiska.

## Wykorzystanie danych rejestrowych do oceny sytuacji ekonomicznej mikro, małych i średnich przedsiębiorstw na przykładzie województwa kujawsko-pomorskiego

Agnieszka Chłoń-Domińczak^a, Aneta Ptak-Chmielewska^b

**Streszczenie.** Mikro, małe i średnie przedsiębiorstwa (MŚP) odgrywają istotną rolę w gospodarce. Według danych Komisji Europejskiej sektor MŚP w 2018 r. skupiał 99,8% przedsiębiorstw w Europie, podobnie w Polsce, w związku z czym dostęp do informacji o tym sektorze i jego sytuacji ekonomicznej, również na najniższym poziomie agregacji, jest kluczowy. Celem badania omawianego w artykule jest zdiagnozowanie, w jakim stopniu dane gromadzone w rejestrach publicznych w Polsce, w tym w rejestrach prowadzonych w Zakładzie Ubezpieczeń Społecznych (ZUS), mogą stanowić źródło wiedzy o stanie i rozwoju sektora MŚP na poziomie lokalnym.

W artykule omówiono sytuację ekonomiczną sektora MŚP w Europie oraz zakres informacji gromadzonych w rejestrach ZUS na temat przedsiębiorstw mikro (w tym osób prowadzących działalność na własny rachunek), małych i średnich. Wykorzystano informacje dotyczące sektora MŚP w woj. kujawsko-pomorskim według stanu na grudzień 2018 r. Uzyskane wyniki potwierdzają, że dane rejestrowe umożliwiają ocenę sytuacji ekonomicznej sektora MŚP na poziomie gmin, zgodnie z rekomendacjami Organizacji Współpracy Gospodarczej i Rozwoju w zakresie zwiększania wiedzy o sytuacji tego sektora na poziomie lokalnym.

**Słowa kluczowe:** mikro, małe i średnie przedsiębiorstwa, rejestry publiczne, monitorowanie sektora MŚP

**JEL:** R12, L25

## The use of register data to monitor micro, small and medium-sized enterprises' economic situation at the local level on the basis of the example of Kujawsko-Pomorskie Voivodship

**Abstract.** Micro, small and medium-sized enterprises (SME) play a significant role in the economy. According to the European Commission data, the SME sector in 2018 represented 99.8% of all enterprises in Europe, and likewise in Poland, therefore it is vital to have access to information regarding this sector and its economic situation on the lowest possible level of aggregation. The aim of the study described in this paper is to assess as to what degree data collected in public registers in Poland, including the information compiled by the Social Insurance Institution (ZUS) register, can constitute a source of information about the situation and development of the SME sector at the local level.

The paper presents the economic situation of the SME sector in Europe, and also a range of information collected in ZUS registers relating to micro enterprises (including the self-employed entrepreneurs), small and medium-sized enterprises. Information concerning the SME sector in Kujawsko-Pomorskie Voivodship as of December 2018 is used in the research. The results confirm that register data enables the assessment of the economic situation of micro, small and medium-sized enterprises at the level of gminas (municipalities, communes), as per the OECD recommendations to broaden knowledge about the condition of the sector at a local level.

**Keywords:** micro, small and medium-sized enterprises, public registers, monitoring of SME sector

^a Szkoła Główna Handlowa w Warszawie, Kolegium Analiz Ekonomicznych.  
ORCID: <https://orcid.org/0000-0001-5068-0276>.

^b Szkoła Główna Handlowa w Warszawie, Kolegium Analiz Ekonomicznych.  
ORCID: <https://orcid.org/0000-0002-9896-4240>.

## 1. Wprowadzenie

Sytuacja ekonomiczna w Polsce w czasie kryzysu finansowego w 2009 r. i bezpośrednio po nim była stosunkowo stabilna w porównaniu do innych krajów europejskich. Polska miała pozycję lidera wśród krajów Organizacji Współpracy Gospodarczej i Rozwoju (OECD) pod tym względem (PARP, 2012, s. 9). Utrzymała wysoką wiarygodność finansową, unikając recesji i drastycznych kryzysów walutowych, a jej dystans do krajów Europy Zachodniej zmniejszył się w okresie kryzysu i bezpośrednio po nim szybciej niż we wcześniejszych latach. Jednym z filarów tego sukcesu były efektywnie działające przedsiębiorstwa, które zdołały dostosować swoją działalność do warunków kryzysowych.

Przeważająca większość przedsiębiorstw działających w Polsce to przedsiębiorstwa mikro, małe i średnie (MŚP). Na potrzeby niniejszego artykułu przyjęto następującą nomenklaturę:

- małe przedsiębiorstwo – do 49 pracujących, w tym:
  - mikroprzedsiębiorstwo – 0–9 pracujących (bez samozatrudnionych),
  - prowadzący działalność na własny rachunek (samozatrudniony);
- średnie przedsiębiorstwo – od 50 do 249 pracujących;
- duże przedsiębiorstwo – powyżej 250 pracujących.

W dalszej analizie przedsiębiorstwa małe i mikro oraz samozatrudnieni są traktowani rozłącznie, zgodnie z rozróżnieniem przyjętym w najnowszej literaturze przedmiotu.

Polskie MŚP radzą sobie równie dobrze jak duże przedsiębiorstwa zagraniczne, mające znaczne zasoby kapitałowe. Należy zauważyć, że zdolności dostosowawcze przedsiębiorstw z sektora MŚP są często większe niż dużych przedsiębiorstw. Przedsiębiorstwa reagują na kryzys ograniczaniem zatrudnienia (Orłowski, Pasternak, Flaht i Szubert, 2010), tymczasem większość MŚP w Polsce jest zaangażowana w działalność usługową, która najmniej cierpi w czasie kryzysu i nie musi podejmować drastycznych decyzji o zwalnianiu pracowników.

Z jednej strony MŚP są elastyczne i dostosowują się do zmian ekonomicznych głównie dzięki zatrudnianiu niewielkiej liczby osób i ograniczaniu wydatków oraz niskiemu poziomowi działalności eksportowej. W wąskim zakresie wykorzystują finansowanie zewnętrzne i mają konserwatywne podejście do ekspansji działalności, więc w konsekwencji nie są zaangażowane w ryzykowne operacje finansowe i duże projekty inwestycyjne. Z drugiej strony mają ograniczone możliwości angażowania się w duże projekty inwestycyjne, m.in. ze względu na mniejszą skalę produkcji oraz skromniejsze zasoby kapitałowe (zwłaszcza kapitału własnego) i mniejszą wiarygodność kredytową. Wynika to z cech jakościowych przedsiębiorstw tej wielkości.

W Europie sektor MŚP skupia 99,8% przedsiębiorstw i odgrywa istotną rolę w gospodarce, ale jego pozycja różni się pomiędzy krajami. Dostęp do danych na temat sytuacji MŚP, zarówno na poziomie krajowym, jak i regionalnym oraz lokalnym, jest więc kluczowy. Dane Głównego Urzędu Statystycznego (GUS) nie są wystarczającym źródłem informacji o sektorze MŚP na poziomie gmin, szczególnie w przypadku mikroprzedsiębiorstw, które nie są objęte pełnym badaniem w ramach statystyki publicznej. Ponadto informacje o stanie zatrudnienia w sektorze MŚP nie są aktualizowane w rejestrze REGON, co utrudnia możliwość prowadzenia bieżących analiz. Należy jednak zaznaczyć, że REGON stanowi jedynie podstawę do tworzenia właściwego rejestru przedsiębiorstw pn. Baza Jednostek Statystycznych (BJS), który jest na bieżąco aktualizowany na podstawie rejestrów administracyjnych oraz badań przedsiębiorstw prowadzonych przez GUS.

Sektor MŚP znacząco przyczynia się do rozwoju regionalnego i lokalnego przez tworzenie miejsc pracy, możliwości inwestowania oraz formowanie kapitału gospodarczego i potencjału rozwoju. Rozmieszczenie przestrzenne MŚP w Polsce jest nierównomierne, co wynika z lokalnego zróżnicowania poziomu rozwoju, dostępu przedsiębiorców do infrastruktury czy zasobów, a także rynku, na którym działają przedsiębiorcy.

Warto podkreślić, że samorzady w Polsce nie mają wypracowanych mechanizmów monitorowania rozwoju przedsiębiorczości na poziomie lokalnym. Kontrola Najwyższej Izby Kontroli z 2017 r. (NIK, 2017) wykazała, że w większości gmin, w których rada gminy przyjęła wieloletnie dokumenty strategiczne uwzględniające zadania z zakresu rozwoju przedsiębiorczości, nie prowadzono systematycznych działań monitorujących wdrażanie strategii lub prowadzono je niezgodnie z przyjętymi zasadami. Co więcej, ze względu na różnorodność przyjętych dokumentów strategicznych nie jest możliwe porównywanie stanu sektora pomiędzy poszczególnymi gminami, nawet jeżeli wdrożono monitorowanie strategii.

Celem badania omawianego w artykule jest zdiagnozowanie, w jakim stopniu dane gromadzone w rejestrach publicznych w Polsce, w tym w rejestrach prowadzonych w Zakładzie Ubezpieczeń Społecznych (ZUS), mogą stanowić źródło wiedzy o stanie rozwoju mikro, małych i średnich przedsiębiorstw na poziomie lokalnym.

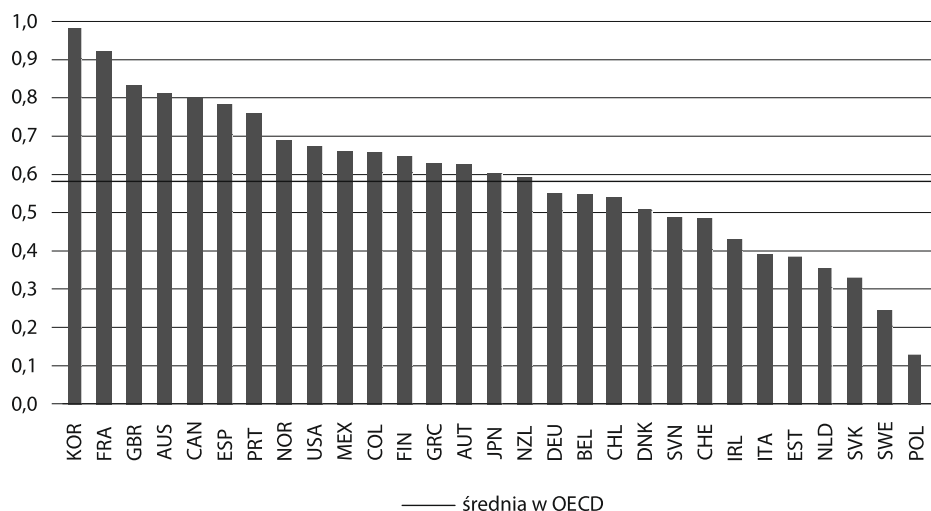
## **2. Monitorowanie sektora MŚP w Unii Europejskiej**

Regionalne nierówności rozwojowe sektora MŚP są naturalne, na co wskazuje m.in. OECD (2017). Niektóre regiony korzystają z lokalnych zasobów naturalnych, inne – z atrakcyjności turystycznej, korzystnych warunków pogodowych czy dobrze rozwiniętej struktury przemysłowej. Należy jednak monitorować zmiany zachodzące w sektorze MŚP, aby zbadać, czy (i na ile) prowadzą one do zwiększenia spójności

rozwoju gospodarczego, czy może wręcz przeciwnie – nierówności na poziomie lokalnym lub regionalnym narastają i jak są one duże. W tym celu należy dysponować danymi statystycznymi pozwalającymi na ocenę kierunku zmian. Ze względu na trudności w pozyskaniu danych bezpośrednio od przedsiębiorstw w ramach m.in. badań reprezentacyjnych, a także z uwagi na koszty pozyskiwania takich informacji, w statystyce w coraz większym stopniu wykorzystuje się dane administracyjne.

Państwo jest właścicielem ogromnych zasobów danych, które mogą być użyte w różny sposób do wspierania prowadzenia polityki publicznej (Bożykowski, Chłoń-Domińczak, Jasiński i Zajac, 2019). Instytucje administracji publicznej gromadzą różnego typu dane, ale bariery prawne oraz brak rozwiązań pozwalających na korzystanie z tych zasobów przez naukowców i praktyków rynkowych powodują, że nieczęsto są one wykorzystywane do celów analitycznych czy ewaluacyjnych. Przeprowadzone w 2014 r. badanie OECD dotyczące dostępu do danych administracyjnych wyraźnie pokazuje, że sytuacja w Polsce w tym zakresie pozostawia jeszcze wiele do życzenia. Na podstawie wartości wskaźnika OURdata (Open, Useful and Re-usable data), mierzącego otwartość, użyteczność i możliwość wtórnego użycia danych, Polska została oceniona najniżej ze wszystkich krajów OECD uwzględnionych w badaniu (wykr. 1).

**Wykr. 1.** Wskaźnik OURdata w wybranych krajach OECD w 2014 r.



Uwaga. KOR – Korea Płd., FRA – Francja, GBR – Wielka Brytania, AUS – Australia, CAN – Kanada, ESP – Hiszpania, PRT – Portugalia, NOR – Norwegia, USA – Stany Zjednoczone, MEX – Meksyk, COL – Kolumbia, FIN – Finlandia, GRC – Grecja, AUT – Austria, JPN – Japonia, NZL – Nowa Zelandia, DEU – Niemcy, BEL – Belgia, CHL – Chile, DNK – Dania, SVN – Słowenia, CHE – Szwajcaria, IRL – Irlandia, ITA – Włochy, EST – Estonia, NLD – Holandia, SVK – Słowacja, SWE – Szwecja, POL – Polska.

Źródło: opracowanie własne na podstawie: OECD (2015).



OECD (2019) wskazuje, że istnieje potrzeba ulepszenia systemów gromadzenia danych oraz krajowych informacji statystycznych dla monitorowania i ewaluacji polityki w zakresie MŚP i przedsiębiorczości. Zgodnie z rekomendacjami OECD dane powinny być dostępne w odpowiednich przedziałach czasowych i na poszczególnych poziomach dezagregacji oraz odnosić się do wskaźników, które są ważne dla oceny sytuacji sektora obecnie i w przyszłości. Rejestry administracyjne stanowią źródło danych, które spełnia tak określone warunki.

Jednym z dokumentów dotyczących wspierania rozwoju przedsiębiorczości w Unii Europejskiej (UE) jest komunikat Komisji Wspólnot Europejskich (COM(2008)394) – „Najpierw myśl na małą skalę” – Program „Small Business Act” dla Europy (SBA). Wskazano w nim następujące priorytety: promowanie przedsiębiorczości, poprawę dostępu do finansowania, zmniejszenie obciążeń regulacyjnych, poprawę dostępu do rynków oraz umiędzynarodowienie działalności przedsiębiorstw. Ujęto je w formie 10 zasad, które mają na celu ukierunkowanie polityki rozwoju na szczeblu unijnym i krajów UE na promowanie przedsiębiorczości i ułatwianie rozwoju przedsiębiorcom. Są to:

1. tworzenie warunków, w których przedsiębiorcy i przedsiębiorstwa rodzinne mogą dobrze prosperować, a przedsiębiorczość jest nagradzana;
2. zagwarantowanie, by uczciwi przedsiębiorcy, których przedsiębiorstwo zostało postawione w stan upadłości, szybko dostali drugą szansę;
3. opracowywanie przepisów zgodnie z zasadą „najpierw myśl na małą skalę”;
4. sprawienie, by organy administracji publicznej lepiej reagowały na potrzeby MŚP;
5. dostosowanie instrumentów polityki publicznej do potrzeb MŚP: ułatwienie MŚP udziału w zamówieniach publicznych oraz lepsze wykorzystanie możliwości pomocy państwa dla MŚP;
6. ułatwianie MŚP dostępu do finansowania i rozwijanie otoczenia prawnego i biznesowego sprzyjającego terminowym płatnościom w transakcjach handlowych;
7. pomaganie MŚP w szerszym korzystaniu z możliwości oferowanych przez jednolity rynek;
8. wspieranie podnoszenia kwalifikacji w MŚP i wszelkich form innowacji;
9. umożliwienie MŚP przekształcania wyzwań związanych z ochroną środowiska na nowe możliwości;
10. zachęcanie MŚP do czerpania korzyści z rozwoju rynków i wspieranie ich w tych działaniach.

Zasady te są monitorowane przez dziewięć grup wskaźników statystycznych (zasady 3 i 4 zostały połączone w jedną grupę). Monitoring, zgodnie z raportem Stano i Ghisetti (2016), daje dobrą podstawę do oceny realizacji priorytetów rozwoju przedsiębiorczości.

Według ostatniego raportu monitorującego (Komisja Europejska, 2019a, 2019b) udział przedsiębiorstw MŚP w 2018 r. w ogólnej liczbie przedsiębiorstw wyniósł w Polsce 99,8%, tak samo jak w UE, a udział zatrudnionych w tym sektorze wyniósł 67,1%, podczas gdy w UE – 66,6%.

Na podstawie danych pochodzących z rejestrów ZUS możliwe jest oszacowanie zarówno wartości wspomnianych wyżej wskaźników na poziomie lokalnym, jak i ich zmiany w czasie, w tym według sekcji Polskiej Klasyfikacji Działalności (PKD).

Dane z rejestrów ZUS mogą być wykorzystane do szacowania wskaźników służących do monitorowania realizacji pierwszej zasady SBA, w szczególności takich jak:

- całkowita aktywność przedsiębiorcza na wczesnym etapie (ang. *early stage entrepreneurial activity*) – procent dorosłych, którzy założyli firmę lub podejmują kroki, aby ją założyć;
- całkowita aktywność przedsiębiorcza na wczesnym etapie dla populacji kobiet w wieku produkcyjnym (ang. *early stage entrepreneurial activity for female population*) – procent kobiet, które założyły firmę lub podejmują kroki w celu jej założenia;
- stabilność przedsiębiorstw (ang. *established business ownership*) – odsetek osób w wieku 18–64 lat, które prowadzą działalność gospodarczą przez co najmniej 42 miesiące;
- udział przedsiębiorstw o wysokim wzroście (ang. *share of high growth enterprises*) – odsetek przedsiębiorstw, w przypadku których roczny wzrost zatrudnienia w okresie trzyletnim przekracza 20%.

Z perspektywy wspierania lokalnych strategii w zakresie rozwoju przedsiębiorczości dane rejestrowe pochodzące z ubezpieczeń społecznych mogą być wykorzystywane do monitorowania wybranych aspektów tworzenia warunków do powstawania przedsiębiorstw na poziomie lokalnym i regionalnym (gminy, powiatu, województwa).

### 3. Metoda badania

W badaniu omawianym w niniejszym artykule dokonano analizy zróżnicowania informacji dotyczących przedsiębiorstw mikro (w tym osób samozatrudnionych), małych i średnich na podstawie danych na temat płatników składek i ubezpieczonych w gminach w woj. kujawsko-pomorskim, zgromadzonych w rejestrach ZUS. Dane te zostały pozyskane w ramach projektu *Usytuowanie na poziomie samorządów lokalnych instrumentów wsparcia dla Małych i Średnich Przedsiębiorstw (MŚP), działających w oparciu o model wielopoziomowego zarządzania regionem (REGIOGMINA)*.¹

¹ Projekt zrealizowano w konsorcjum kierowanym przez samorząd woj. kujawsko-pomorskiego z udziałem Szkoły Głównej Handlowej w Warszawie i Uniwersytetu Mikołaja Kopernika w Toruniu, w ramach inicjatywy GOSPOSTRATEG ze środków Narodowego Centrum Badań i Rozwoju.

Analizowane dane obejmują: liczbę płatników składek, liczbę osób ubezpieczonych oraz średnią podstawę wymiaru składek na ubezpieczenia emerytalne i rentowe w gminach woj. kujawsko-pomorskiego w grudniu 2018 r.

Płatnicy składek są rozumiani jako jednostki organizacyjne bądź osoby zobligowane z jakiegokolwiek tytułu do obliczenia należnych składek, pobrania ich oraz przekazania tych składek do ZUS; są to przede wszystkim pracodawcy.

Osoby ubezpieczone to takie, które podlegają lub podlegały ubezpieczeniu emerytalnemu i ubezpieczeniom rentowym przynajmniej jeden dzień, np. jako pracownicy lub osoby prowadzące działalność gospodarczą.

Podstawę wymiaru składek na ubezpieczenia emerytalne i rentowe w przypadku pracowników stanowi przychód uzyskiwany u pracodawcy z umowy o pracę lub/i innych umów, a także wartość świadczeń uzyskiwanych w naturze (np. z wynajmu mieszkania służbowego). W przypadku zleceniobiorców – którzy nie podlegają ubezpieczeniom społecznym z innych tytułów – jest to kwota odpłatności określonej w umowie.

W badaniu analizowano rozkład liczby płatników składek według wielkości przedsiębiorstw, udział osób ubezpieczonych w wybranych sekcjach PKD według wielkości przedsiębiorstw, a także rozkłady podstaw wymiaru składek w poszczególnych gminach woj. kujawsko-pomorskiego. Analiza ta miała na celu ocenę, czy obserwowane zróżnicowanie informacji dotyczących płatników składek, osób ubezpieczonych i podstaw wymiaru składek na poziomie gmin wskazuje na zasadność monitorowania sytuacji ekonomicznej sektora MŚP na poziomie gmin.

#### **4. Dane o sektorze MŚP gromadzone w rejestrach ZUS**

Zgodnie z ustawą o systemie ubezpieczeń społecznych² ZUS prowadzi rejestr płatników składek, czyli gromadzi informacje o podmiotach opłacających składki na ubezpieczenia społeczne za osoby zatrudnione na umowę o pracę i umowy cywilnoprawne, a także za samych siebie. Płatnikami składek są zatem przedsiębiorcy opłacający składki za samych siebie, pracodawcy opłacający składki za swoich pracowników, a także inne podmioty objęte obowiązkiem ubezpieczenia społecznego. ZUS prowadzi też rejestr osób ubezpieczonych, zawierający informacje o osobach, za które odprowadzane są składki na ubezpieczenia społeczne.

W rejestrze płatników składek ZUS odnotowuje:

- rejestrację płatników wraz z powstaniem obowiązku opłacania składek, odpowiednio poprzez rejestrację w Centralnej Ewidencji i Informacji o Działalności Gospodarczej (dla przedsiębiorców) lub w Krajowym Rejestrze Sądowym (w przypadku spółek);

² Ustawa o systemie ubezpieczeń społecznych z dnia 13 października 1998 r. (Dz.U. 1998 nr 137, poz. 887).

- opłacanie składek za osoby ubezpieczone lub za samych siebie – raporty składane co miesiąc przez płatników zatrudniających osoby ubezpieczone;
- wyrejestrowanie płatników wraz z ustaniem obowiązku opłacania składek (podobnie jak w przypadku rejestracji).

W rejestrze osób ubezpieczonych gromadzone są:

- zgłoszenie osoby do ubezpieczenia, m.in. dane osobowe (imię, nazwisko, dane identyfikacyjne, rok urodzenia, płeć) osoby objętej ubezpieczeniami społecznymi;
- miesięczne informacje dotyczące opłacanych składek, w tym tytuł ubezpieczenia (umowa o pracę, umowa-zlecenie, inne tytuły) i podstawa wymiaru składek na ubezpieczenia emerytalne i rentowe;
- wyrejestrowanie osoby z ubezpieczenia, następujące na skutek rozwiązania umowy o pracę lub innej umowy rodzącej obowiązek ubezpieczenia społecznego.

Rejestry płatników składek i ubezpieczonych zawierają szczegółowe informacje o liczbie płatników składek, w tym również według liczby ubezpieczonych. Informacje te pozwalają na przybliżone oszacowanie wielkości przedsiębiorstw na podstawie kryterium liczby pracujących, w szczególności mikroprzedsiębiorstw, według sekcji PKD, a także miejsca (gminy, powiatu, województwa), w którym dane przedsiębiorstwo działa³.

Obowiązujące przepisy pozwalają na pozyskanie danych jednostkowych (nawet po anonimizacji) o płatnikach składek i osobach ubezpieczonych jedynie wskazanym ustawowo instytucjom (np. GUS), natomiast dla celów badań naukowych możliwe jest otrzymanie wskaźników zagregowanych, np. na poziomie gminy.

Dane gromadzone w rejestrach ZUS mogą posłużyć do określenia liczby pracujących, która stanowi jedno z kryteriów klasyfikacji przedsiębiorstw. Zgodnie z przyjętą w UE definicją MŚP do ich kwalifikacji stosowane są trzy kryteria: liczba pracujących, obrót roczny oraz roczna suma bilansowa, przy czym spełnienie kryterium liczby pracujących jest obowiązkowe, a w przypadku pułapu obrotu i sumy bilansowej jedno z kryteriów może być przekroczone (Komisja Europejska, 2015). Dane z rejestru ZUS nie pozwalają na uwzględnienie wszystkich wymienionych kryteriów, dlatego w dalszej analizie oparto kategoryzację przedsiębiorstw na liczbie osób ubezpieczonych – tj. takich, za które pracodawca (płatnik składek) odprowadza składki na ubezpieczenia emerytalne i rentowe – jako na przybliżonej informacji o liczbie pracujących. Przyjęto założenie, że jeżeli płatnik składek odprowadza składki za osobę ubezpieczoną, to oznacza, że taka osoba swoją pracą przyczynia się do funk-

³ Należy zauważyć, że w niektórych przypadkach informacja o siedzibie płatnika dotyczy np. centrali przedsiębiorstwa, jednak dla analizowanej populacji płatników można przyjąć, że miejsce siedziby firmy jest tożsame z miejscem jej działalności.

cjonowania przedsiębiorstwa, a zatem powinna być uwzględniona w liczbie pracujących w przedsiębiorstwie, nawet jeżeli nie ma podpisanej umowy o pracę i formalnie nie jest pracownikiem.

Precyzyjniejszą kategoryzację przedsiębiorstw mogłoby umożliwić wykorzystanie informacji o rocznych obrotach, które znajdują się w danych rejestrowych gromadzonych przez administrację skarbową.

Do statystyk, które mogą zostać oszacowane na podstawie danych gromadzonych w rejestrach ZUS, należą:

1. liczba płatników składek według sekcji PKD oraz liczby ubezpieczonych:
  - od 1 do 9 ubezpieczonych, będąca przybliżeniem liczby mikroprzedsiębiorstw;
  - od 10 do 49 ubezpieczonych, będąca przybliżeniem liczby małych przedsiębiorstw;
  - od 50 do 249 osób ubezpieczonych, będąca przybliżeniem liczby średnich przedsiębiorstw;
  - 250 i więcej osób ubezpieczonych, będąca przybliżeniem liczby dużych przedsiębiorstw;
2. liczba osób ubezpieczonych prowadzących działalność na własny rachunek i opłacających za siebie składki, czyli samozatrudnionych, w tym według płci i grup wieku oraz sekcji PKD;
3. liczba osób ubezpieczonych w przedsiębiorstwach w ramach grup określonych w pkt 1, za które płatnicy składek odprowadzają składki, będąca przybliżoną informacją o liczbie pracujących, według płci, grup wieku i sekcji PKD;
4. przeciętna podstawa wymiaru składki ubezpieczonych na ubezpieczenia emerytalne i rentowe u płatników według liczby ubezpieczonych, w tym według płci, grup wieku i sekcji PKD.

Wskaźniki te pozwalają w szczególności na oszacowanie stanu ekonomicznego sektora MŚP w ujęciu przestrzennym na poziomie gminy lub powiatu w danym okresie (np. na koniec roku kalendarzowego), w tym określenie roli MŚP na lokalnym rynku pracy. Ponadto informacje dotyczące rejestracji płatników składek umożliwiają monitorowanie dynamiki zmian. Odbywa się to poprzez zbieranie informacji na temat dokonywanych w wybranym okresie (np. roku kalendarzowym):

- rejestracji i wyrejestrowań płatników składek (prowadzących działalność na własnych rachunek i opłacających za siebie składki, od 1 do 9 ubezpieczonych, od 10 do 49 ubezpieczonych, od 50 do 249 ubezpieczonych) według sekcji PKD;
- zgłoszeń do ubezpieczenia i wyrejestrowań osób ubezpieczonych.

Na podstawie danych gromadzonych w rejestrach ZUS można również monitorować:

- przepływ płatników składek pomiędzy poszczególnymi kategoriami przedsiębiorstw;
- przeżywalność przedsiębiorstw – przez określenie liczby wyrejestrowanych płatników składek po upływie danego okresu od pierwszego zgłoszenia płatnika składek;
- tworzenie przedsiębiorstw – przez określenie liczby płatników składek, którzy zgłaszają się do ubezpieczenia społecznego;
- dynamikę zatrudnienia – przez określenie liczby osób ubezpieczonych, które są rejestrowane i wyrejestrowywane z ubezpieczeń społecznych.

W efekcie możliwe jest również oszacowanie wskaźników monitorowania pierwszej zasady SBA, zgodnie z opisem przedstawionym wcześniej.

Należy zwrócić uwagę na utrudnienia i ograniczenia wynikające z wykorzystywania danych rejestrowych. Po pierwsze, osoby ubezpieczone to zarówno pracujący na podstawie umowy o pracę, jak i osoby objęte ubezpieczeniem społecznym z innych tytułów (np. na podstawie umowy-zlecenia czy wykonywania pracy nakładczej). Identyfikacja tych ubezpieczonych, którzy pracują na podstawie umowy o pracę, jest możliwa dopiero za pomocą kodu tytułu ubezpieczenia. Po drugie, dane dla wybranych miesięcy mogą być obciążone zniekształceniami wynikającymi np. z faktu wypłacania dodatkowych składników wynagrodzeń, które mogą wpływać na podstawę wymiaru składek na ubezpieczenia emerytalne i rentowe.

## **5. Diagnoza sektora MŚP w województwie kujawsko-pomorskim na podstawie danych gromadzonych w rejestrach ZUS**

Zgodnie z danymi z rejestru płatników składek w grudniu 2018 r. w woj. kujawsko-pomorskim było 93116 płatników składek ZUS (których liczba jest na potrzeby niniejszej diagnozy traktowana jako przybliżona liczba przedsiębiorstw), z czego 99,68% stanowili:

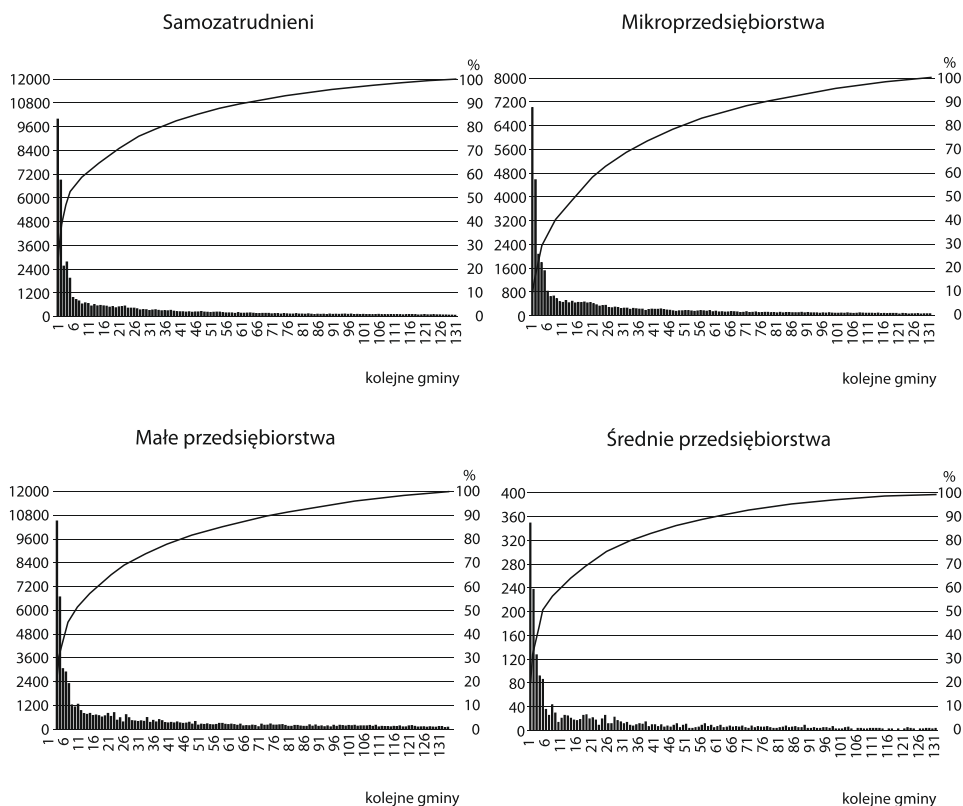
- 51,29% – samozatrudnieni;
- 38,30% – mikroprzedsiębiorstwa;
- 8,13% – małe przedsiębiorstwa;
- 1,91% – średnie przedsiębiorstwa.

Najwięcej przedsiębiorstw w sektorze MŚP związanych jest z sekcją PKD G (handel) – 22%, sekcją F (budownictwo) – 13% oraz sekcją C (przetwórstwo przemysłowe) – 10%.

Rozkład liczby przedsiębiorstw w poszczególnych gminach jest nierównomierny. W pięciu największych miastach w województwie: Bydgoszczy, Toruniu, Włocławku, Grudziądzu i Inowrocławiu działa 50,0% osób prowadzących działalność na własny rachunek, 47,0% mikroprzedsiębiorstw, 44,7% małych przedsiębiorstw

i 50,2% średnich przedsiębiorstw. Nierówności w zakresie liczebności przedsiębiorstw określonych na podstawie liczby płatników w gminach oraz skumulowane liczebności przedstawia wyk. 2. Wskazuje on na znaczące zróżnicowanie rozkładu przedsiębiorstw w sektorze MŚP w woj. kujawsko-pomorskim, co potwierdza, że gminy w tym województwie znacznie różnią się pod względem liczby przedsiębiorstw mikro, małych i średnich działających na ich terenie.

**Wykr. 2.** Przedsiębiorstwa (na podstawie liczby płatników składek) w gminach woj. kujawsko-pomorskiego w grudniu 2018 r. według liczby pracujących



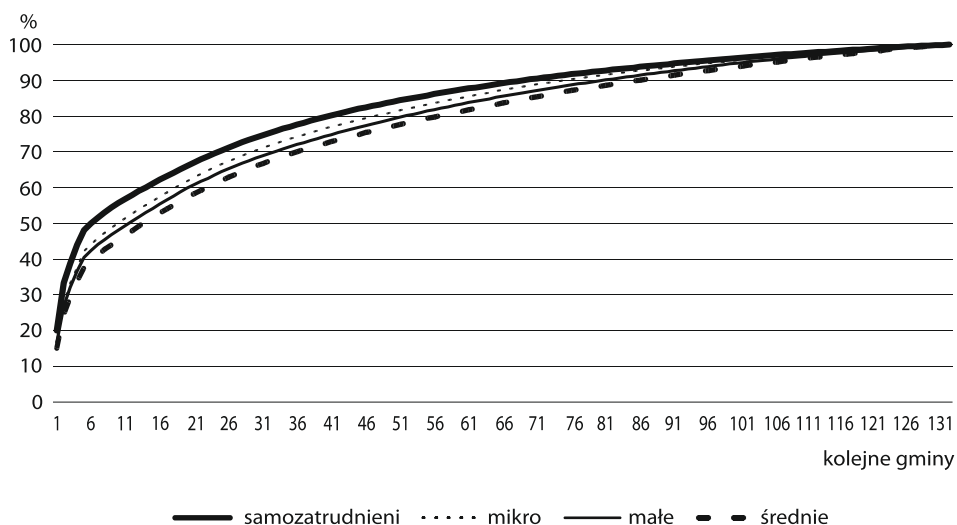
Uwaga. Słupki pokazują liczebność płatników składek w danej grupie w gminach w porządku malejącym, a wykres liniowy – skumulowane liczebności.

Źródło: opracowanie własne na podstawie zagregowanych danych ZUS uzyskanych w projekcie REGIOGMINA.

Nierównomierny rozkład liczebności przedsiębiorstw jest powiązany z nierównomiernym rozkładem liczby osób ubezpieczonych będącej przybliżeniem liczby pracujących. W woj. kujawsko-pomorskim w grudniu 2018 r. w przedsiębiorstwach mikro było ubezpieczonych 115,0 tys. osób, w małych – 166,1 tys. osób, a w średnich – 183,7 tys. osób.

Podobnie jak w przypadku przedsiębiorstw charakterystyczna jest koncentracja pracujących (których liczbę oszacowano na podstawie liczby ubezpieczonych) w największych miastach województwa, bez względu na wielkość przedsiębiorstw. W 12 największych miejscowości mieszka około połowy pracujących, a w 1/3 gmin skupiających najwięcej osób pracujących (łącznie z miastami na prawach powiatu) pracuje ok. 3/4 tej populacji (wykr. 3).

**Wykr. 3.** Pracujący (na podstawie liczby ubezpieczonych) w gminach woj. kujawsko-pomorskiego w grudniu 2018 r. według wielkości przedsiębiorstw w gminach – dystrybuanta empiryczna



Uwaga. Gminy zostały uporządkowane malejąco według liczby osób ubezpieczonych.

Źródło: opracowanie własne na podstawie zagregowanych danych ZUS uzyskanych w projekcie REGIOGINA.

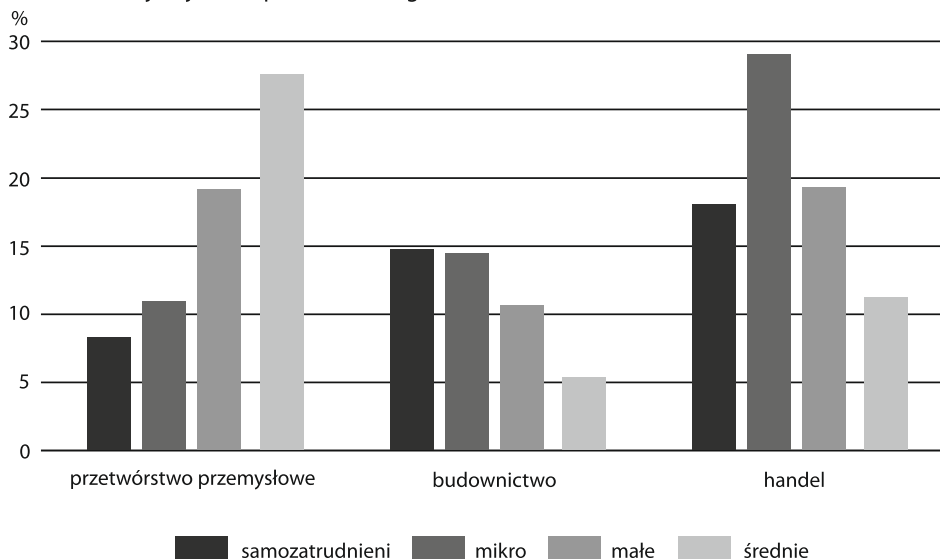
Obserwujemy również zróżnicowanie rozkładu przedsiębiorstw według sekcji PKD (wykr. 4 i tabl. 1):

- wśród samozatrudnionych 18% działa w handlu, a 15% – w budownictwie; ich udział w przetwórstwie przemysłowym jest relatywnie niewielki (8%);
- wśród mikroprzedsiębiorstw prawie 30% pracujących działa w handlu, ok. 15% – w budownictwie, a nieco ponad 10% – w przetwórstwie przemysłowym. Największe zróżnicowanie udziału pracujących według sektora obserwuje się w budownictwie;
- wśród małych przedsiębiorstw udział pracujących w handlu również jest najwyższy (19%), choć mniejszy niż w przypadku mikroprzedsiębiorstw, prawie taki sam w przetwórstwie przemysłowym i niższy w budownictwie (11%);



- wśród średnich przedsiębiorstw największy jest udział pracujących w przetwórstwie przemysłowym (28%), a oszacowane wskaźniki pozycyjne (mediana i kwartyle) wskazują na duże zróżnicowanie udziału pracujących w tym sektorze na poziomie gmin, następnie w budownictwie – 14% i najmniej w handlu – nieco ponad 10%.

**Wykr. 4.** Udział pracujących w przedsiębiorstwach według sekcji PKD z największym udziałem pracujących w ogólnej liczbie pracujących w woj. kujawsko-pomorskim w grudniu 2018 r.



Źródło: opracowanie własne na podstawie zagregowanych danych ZUS uzyskanych w projekcie REGIOGINA.

**Tabl. 1.** Udział pracujących w przedsiębiorstwach według sekcji PKD z największym udziałem pracujących w ogólnej liczbie pracujących w grudniu 2018 r. – wybrane statystyki opisowe danych dla gmin

Statystyki opisowe	Przetwórstwo przemysłowe	Budownictwo	Handel
	w %		
<b>Samozatrudnieni</b>			
Średnia ważona ^a .....	8,30	14,74	18,05
Średnia nieważona .....	21,53	18,85	7,39
Mediana .....	20,63	18,33	7,11
Kwartył 1 .....	16,73	15,78	5,08
Kwartył 3 .....	26,25	21,32	9,62
<b>Mikroprzedsiębiorstwa</b>			
Średnia ważona ^a .....	10,87	14,41	29,06
Średnia nieważona .....	11,02	16,39	29,21
Mediana .....	10,70	15,73	29,61
Kwartył 1 .....	9,30	12,73	26,69
Kwartył 3 .....	12,89	19,50	31,88

**Tabl. 1.** Udział pracujących w przedsiębiorstwach według sekcji PKD z największym udziałem pracujących w ogólnej liczbie pracujących w grudniu 2018 r.  
– wybrane statystyki opisowe danych dla gmin (dok.)

Statystyki opisowe	Przetwórstwo przemysłowe	Budownictwo	Handel
	w %		
<b>Małe przedsiębiorstwa</b>			
Średnia ważona ^a .....	19,12	10,61	19,30
Średnia nieważona .....	20,01	11,06	18,72
Mediana .....	19,78	10,26	18,53
Kwartył 1 .....	15,94	8,39	15,92
Kwartył 3 .....	23,14	12,82	21,75
<b>Średnie przedsiębiorstwa</b>			
Średnia ważona ^a .....	27,56	5,36	11,20
Średnia nieważona .....	30,11	5,91	11,24
Mediana .....	28,82	5,40	10,77
Kwartył 1 .....	23,60	3,12	7,67
Kwartył 3 .....	36,51	7,67	13,86

a Dane ważone liczebnością przedsiębiorstw w poszczególnych gminach.

Źródło: opracowanie własne na podstawie zagregowanych danych ZUS.

Zróznicowanie sektora MŚP obserwujemy również w przypadku średniej podstawy wymiaru składek na ubezpieczenia emerytalne i rentowe, co ilustruje tabl. 2. W tym zestawieniu pominięto osoby prowadzące działalność na własny rachunek, ze względu na zbliżoną wysokość podstawy wymiaru składek na ubezpieczenia społeczne płaconych ryczałtowo od deklarowanej kwoty równej 60% przeciętnego wynagrodzenia w gospodarce.

**Tabl. 2.** Średnia podstawa wymiaru składek w sektorze MŚP^a w grudniu 2018 r. według sekcji PKD oraz płci

Przedsiębiorstwa	Ogółem	Sekcje PKD			Płeć	
		przetwórstwo przemysłowe	budownictwo	handel	kobiety	mężczyźni
Mikro .....	2744	2424	2607	2397	2688	2795
Małe .....	3372	3232	3895	3243	3117	3601
Średnie .....	3637	3810	5084	3396	3178	4076

a Bez samozatrudnionych.

Źródło: jak przy tabl. 1.

Wysokość podstawy wymiaru składek rośnie wraz z wielkością przedsiębiorstw (płatników składek). Spośród trzech największych pod względem udziału pracujących (osób ubezpieczonych) sekcji PKD najwyższe wartości podstawy wymiaru składek obserwujemy w budownictwie, natomiast najniższe – w handlu. Różnica pomię-

dzy przeciętną podstawą wymiaru składek kobiet i mężczyzn (na korzyść mężczyzn) rośnie wraz z wielkością przedsiębiorstw. W przypadku mikroprzedsiębiorstw wynosi niecałe 4,0% średniej podstawy wymiaru składek kobiet, w grupie małych przedsiębiorstw – 15,5%, a w grupie średnich przedsiębiorstw – już 28,3%.

**Tabl. 3.** Średnia podstawa wymiaru składek w przedsiębiorstwach według sekcji oraz płci – wybrane statystyki opisowe dla gmin

Statystyki opisowe	Ogółem	Sekcje PKD			Płeć	
		przetwórstwo przemysłowe	budownictwo	handel	kobiety	mężczyźni
<b>Mikroprzedsiębiorstwa^a</b>						
Średnia nieważona .....	2114	2014	1985	1864	2080	2148
Mediana .....	2090	2033	1958	1876	2063	2124
Kwartył 1 .....	2008	1806	1812	1741	1958	2030
Kwartył 3 .....	2221	2169	2112	1965	2223	2276
<b>Małe przedsiębiorstwa</b>						
Średnia nieważona .....	2720	2423	2614	2398	2667	2769
Mediana .....	2705	2357	2542	2345	2677	2741
Kwartył 1 .....	2589	2205	2319	2196	2494	2584
Kwartył 3 .....	2850	2641	2869	2567	2808	2914
<b>Średnie przedsiębiorstwa</b>						
Średnia nieważona .....	3361	3228	3944	3250	3096	3600
Mediana .....	3354	3189	3870	3126	3078	3560
Kwartył 1 .....	3207	2969	3433	2912	2929	3399
Kwartył 3 .....	3513	3478	4351	3565	3278	3764

a Bez samozatrudnionych.

Źródło: jak przy tabl. 1.

Tablica 3 przedstawia rozkład kwartyłowy wartości średniej podstawy wymiaru składek na ubezpieczenia emerytalne i rentowe w gminach woj. kujawsko-pomorskiego. Wynika z niego, że zróżnicowanie średniej podstawy wymiaru składek pomiędzy gminami (mierzone odchyleniem ćwiartkowym) rośnie wraz z wielkością przedsiębiorstw, szczególnie w budownictwie i handlu.

## 6. Podsumowanie

Dane dotyczące płatników składek na ubezpieczenia społeczne w woj. kujawsko-pomorskim zagregowane na poziomie gmin pokazują, że sektor mikro, małych i średnich przedsiębiorstw jest bardzo zróżnicowany przestrzennie. W największych miejscowościach występuje koncentracja zarówno przedsiębiorstw, jak i pracujących. Zróżnicowanie wielkości i roli sektora MŚP na poziomie gmin zależy również od rodzaju działalności.

Uzyskane wyniki potwierdzają, że dane gromadzone w rejestrach publicznych umożliwiają dokonywanie oceny sytuacji ekonomicznej sektora MŚP na poziomie regionalnym i lokalnym. Pozwalają na przedstawienie informacji dotyczących tego sektora, w tym określenie liczby przedsiębiorstw (przybliżonych liczbą płatników składek), liczby pracujących (przybliżonych liczbą osób ubezpieczonych) na poziomie gmin, również według sekcji PKD. Oszacowania takie mogą być łatwo powtórzone, a uzyskanie informacji nie jest kosztowne.

Wykorzystanie danych gromadzonych w rejestrach ZUS może pozwolić na zastosowanie się do rekomendacji OECD w zakresie zwiększania wiedzy o rozwoju sektora MŚP na poziomie lokalnym, a także na bieżące monitorowanie wskaźników dotyczących realizacji pierwszej zasady programu SBA przyjętego przez Komisję Europejską.

Korzystanie jedynie z danych rejestrowych ma jednak ograniczenia. Informacje pochodzące z rejestrów ZUS pozwalają na kategoryzację przedsiębiorstw na podstawie liczby osób objętych ubezpieczeniami społecznymi. Liczba pracujących może być obliczona dopiero na podstawie kodu tytułu ubezpieczeń. Dane gromadzone w rejestrach ZUS nie obejmują informacji o wielkości obrotów przedsiębiorstw czy sumy bilansowej przedsiębiorstw. Takie dane są zawarte w rejestrach administracji skarbowej. Połączenie danych pochodzących z rejestrów ZUS i administracji skarbowej dla poszczególnych przedsiębiorstw wymaga odpowiednich rozwiązań prawnych. Ponadto dane administracyjne nie pozwalają na ocenę, jak na sytuację przedsiębiorstw wpływają działania ukierunkowane na wspieranie rozwoju przedsiębiorczości podejmowane przez lokalne lub regionalne władze samorządowe, a także inne podmioty powołane w tym celu. Wpływ ten można ocenić jedynie pośrednio przez zastosowanie odpowiednich modeli analitycznych.

Dalsze pogłębione analizy mogą umożliwić oszacowanie wskaźników monitorowania stanu i zmian zachodzących w sektorze MŚP na poziomie lokalnym zgodnie z SBA. Dane te, wraz z innymi informacjami (np. uwarunkowaniami rozwoju przestrzennego), mogą być również wykorzystane do identyfikacji czynników wpływających na strukturę sektora MŚP w poszczególnych gminach. W efekcie takich analiz będzie formułować rekomendacje dotyczące lokalnych instrumentów rozwoju tego sektora.

## Bibliografia

- Bożykowski, M., Chłoń-Domińczak, A., Jasiński, M., Zajac, T. (2019). *Dane publiczne – nowy impuls dla rozwoju Polski*. Warszawa: Polski Instytut Ekonomiczny.
- Komisja Europejska. (2015). *Poradnik dla użytkowników dotyczących definicji MŚP*. Luksemburg: Urząd Publikacji Unii Europejskiej. Pobrane z: [http://publications.europa.eu/resource/cellar/179c0ce87-f4dc-11e6-8a35-01aa75ed71a1.0005.01/DOC_1](http://publications.europa.eu/resource/cellar/179c0ce87-f4dc-11e6-8a35-01aa75ed71a1.0005.01/DOC_1).
- Komisja Europejska. (2019a). *2019 SBA Fact Sheet: Poland*. Pobrane z: <https://ec.europa.eu/docsroom/documents/38662/attachments/22/translations/en/renditions/pdf>.

- Komisja Europejska. (2019b). *Annual Report on European SMEs 2018/2019*. Pobrane z: <https://op.europa.eu/en/publication-detail/-/publication/cadb8188-35b4-11ea-ba6e-01aa75ed71a1/language-en>.
- NIK. (2017). *Informacja o wynikach kontroli: Wspieranie przedsiębiorczości przez gminy*. Pobrane z: <https://www.nik.gov.pl/plik/id,17394,vp,19963.pdf>.
- OECD. (2015). *Government at a Glance 2015*. Paris: OECD. Pobrane z: [https://www.oecd-ilibrary.org/governance/government-at-a-glance-2015_gov_glance-2015-en](https://www.oecd-ilibrary.org/governance/government-at-a-glance-2015_gov_glance-2015-en).
- OECD. (2017). Making policy evaluation work: The Case of Regional Development Policy. *OECD Science, Technology and Innovation Policy Papers*, (38), 1–35. Pobrane z: [https://www.oecd-ilibrary.org/science-and-technology/making-policy-evaluation-work_c9bb055f-en](https://www.oecd-ilibrary.org/science-and-technology/making-policy-evaluation-work_c9bb055f-en).
- OECD. (2019). Monitoring and evaluation of SME and entrepreneurship programmes. W: *Strengthening SMEs and Entrepreneurship for Productivity and Inclusive Growth: OECD 2018 Ministerial Conference on SMEs* (s. 127–132). Paris: OECD Publishing. Pobrane z: <http://www.oecd.org/publications/strengthening-smes-and-entrepreneurship-for-productivity-and-inclusive-growth-c19b6f97-en.htm>.
- Orłowski, W., Pasternak, R., Flaht, K., Szubert, D. (2010). *Procesy inwestycyjne i strategie przedsiębiorstw w czasach kryzysu*. Warszawa: Polska Agencja Rozwoju Przedsiębiorczości.
- PARP. (2012). *Raport o stanie sektora małych i średnich przedsiębiorstw w Polsce w latach 2010–2011*. Warszawa: Polska Agencja Rozwoju Przedsiębiorczości.
- Stano, P. M., Ghisett, C. (2016). *Monitoring Europe SMEs' performance in Europe: Methodological assessment of the SME Scoreboard 2016*. Luxembourg: Publications Office of the European Union. Pobrane z: <https://ec.europa.eu/jrc/en/publication/monitoring-smes-performance-europe-methodological-assessment-sme-scoreboard-2016>.

## Wyznaczanie współczynników korelacji liniowej – podstawy

Zbigniew Śleszyński^a

**Streszczenie.** Celem artykułu jest przedstawienie podstawowych miar współzależności cech ilościowych, które wykorzystuje się w modelowaniu ekonometrycznym, oraz ich wybranych zastosowań. Omówiono współczynniki: korelacji liniowej Pearsona, korelacji wielowymiarowej, determinacji, korelacji cząstkowej i semikorelacji cząstkowej. Zaprezentowane w artykule podejście do wymienionych miar jest jednorodne. Każdą zdefiniowano jako współczynnik korelacji liniowej odpowiednich wektorów otrzymanych na podstawie równań regresji. Podano też wzajemne zależności między współczynnikami. W obliczeniach wykorzystano macierze brzegowe, co znacznie ułatwiło ten proces. W celu sprawdzenia poprawności obliczeń posłużono się programem Statistica 13.3 PL. Zagadnienie zilustrowano na przykładzie modelu regresji wzrostu płac w Polsce w latach 2001–2019 zawierającego cztery zmienne objaśniające, szacowanego metodą najmniejszych kwadratów.

**Słowa kluczowe:** współczynnik korelacji liniowej, współczynnik korelacji wielowymiarowej, współczynnik determinacji, współczynnik korelacji cząstkowej, współczynnik semikorelacji cząstkowej, macierze brzegowe

**JEL:** C10, C20, C40, C50

## The basics of determining the coefficients of a linear correlation

**Abstract.** The aim of the paper is to present the basic measures related to the analysis of relationships between quantitative variables used in econometric modelling and their selected applications. The following measures are discussed: the Pearson correlation coefficient, the multivariate correlation coefficient, coefficient of determination, partial correlation coefficient and semi-partial correlation coefficient. A homogeneous approach is applied to the measures presented. Each is defined as a linear correlation coefficient of relevant vectors derived from regression equations. Additionally, mutual relations between the coefficients are described. Bordered matrices have been applied to the calculations, which significantly simplified the process, while the Statistica 13.3 PL program was used to verify the correctness of the calculations. The issue is illustrated in the model of regression of salary growth in Poland in the years 2001–2019 with four covariates, estimated using the least squares method.

**Keywords:** correlation coefficient, multivariate correlation coefficient, coefficient of determination, partial correlation coefficient, semi-partial correlation coefficient, bordered matrices

---

^a Uniwersytet Technologiczno-Humanistyczny im. Kazimierza Pułaskiego w Radomiu, Wydział Ekonomii i Finansów. ORCID: <https://orcid.org/0000-0002-8388-351X>.

## 1. Wprowadzenie

Podstawową miarą współzależności liniowej cech ilościowych jest współczynnik korelacji liniowej Pearsona. W badaniach ekonometrycznych wykorzystuje się go m.in. do doboru zmiennych, sprawdzania przydatności zmiennych objaśniających w modelu, badania współliniowości, koincydencji i jakości modelu. Jego różne zastosowania wynikają z obliczania współczynników korelacji różnych wektorów.

Celem artykułu jest przedstawienie podstawowych miar współzależności cech ilościowych wykorzystywanych w modelowaniu ekonometrycznym oraz ich wybranych zastosowań. Skoncentrowano się na zdefiniowaniu tych miar jako współczynników korelacji liniowej różnych wektorów, danych lub otrzymanych w trakcie modelowania. Podano wzory macierzowe służące do ich wyznaczania. Takie podejście wydaje się uzasadnione, gdyż ułatwia interpretację poszczególnych współczynników i ich wzajemnych zależności. Jednocześnie różni się od powszechnie stosowanego w większości opracowań (np. Dziechciarz, 2015; Ezekiel i Fox, 1959; Zeliaś, 2000). W szczególności dotyczy to współczynników korelacji cząstkowej. W literaturze najczęściej podawane są wzory analityczne jedynie dla modelu regresji liniowej z dwiema lub trzema zmiennymi niezależnymi albo wzory przedstawiające dany współczynnik korelacji jako iloraz wyznaczników pewnych macierzy. Ponadto w polskiej literaturze współczynnik semikorelacji cząstkowej jest w zasadzie pomijany. Przedstawienie mało znanych współczynników korelacji cząstkowej i semikorelacji cząstkowej nadaje pracy dodatkowy walor edukacyjny. Wykorzystanie macierzy brzegowych do obliczeń wymienionych współczynników ułatwia proces obliczeniowy i pozwala zauważyć wzajemne zależności.

## 2. Współczynnik korelacji liniowej

W celu zbadania współzależności liniowej dwóch cech ilościowych  $X$  oraz  $Y$  zwykle obliczamy współczynnik korelacji liniowej. Dla  $n$  wartości tych cech (odpowiednio  $x_t$  oraz  $y_t$ ,  $t = 1, 2, \dots, n$ ) współczynnik ten będzie oznaczany  $r(X, Y)$  i obliczany zgodnie ze wzorem:

$$r(X, Y) = \frac{\sum_{t=1}^n (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^n (x_t - \bar{x})^2 \sum_{t=1}^n (y_t - \bar{y})^2}} \quad (1)$$

gdzie  $\bar{x}$  oraz  $\bar{y}$  – odpowiednio średnie arytmetyczne wartości cechy  $X$  oraz  $Y$ .

Jak wiadomo, współczynnik korelacji liniowej jest symetryczny, przyjmuje wartości z przedziału  $[-1, 1]$ , jego bezwzględna wartość świadczy o sile związku liniowego, a znak – o kierunku tego związku.

W przypadku zależności nieliniowych należy posługiwać się współczynnikiem korelacji nieliniowej  $R(Y, X)$  zmiennej  $Y$  względem zmiennej  $X$ , zgodnie ze wzorem (Zeliaś, 2000):

$$R(Y, X) = \sqrt{1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2}} \quad (2)$$

gdzie  $\hat{y}_t$  – wartość teoretyczna zmiennej  $Y$  w okresie  $t$  oszacowana na podstawie odpowiedniego modelu.

Warto nadmienić, że współczynnik korelacji nieliniowej oraz współczynnik korelacji wielowymiarowej są pierwiastkami ze współczynnika determinacji (na ogół oznaczanego  $R^2$ ) w odpowiednim modelu regresyjnym. Model liniowy można traktować jako szczególny przypadek modelu nieliniowego lub modelu liniowego wielowymiarowego (regresja wielokrotna).

Współczynnik korelacji nieliniowej nie jest symetryczny, przyjmuje wartości z przedziału  $[0, 1]$ , a im bardziej jego wartość zbliża się do 1, tym związek korelacyjny między rozpatrywanymi zmiennymi jest mocniejszy. W celu zilustrowania tych zależności obliczono współczynnik korelacji liniowej  $r(X, Y)$  dla danych pięciu wartości cech  $X$  i  $Y$ . Niezbędne obliczenia przedstawiono w tabl. 1.

**Tabl. 1.** Wyznaczanie współczynnika korelacji linowej cech  $X$  oraz  $Y$

$t$	$x_t$	$y_t$	$y_t - \bar{y}$	$(x_t - \bar{x})(y_t - \bar{y})$	$(x_t - \bar{x})^2$	$(y_t - \bar{y})^2$
1 .....	-2	4	2	-4	4	4
2 .....	-1	1	-1	1	1	1
3 .....	0	0	-2	0	0	4
4 .....	1	1	-1	-1	1	1
5 .....	2	4	2	4	4	4
$\Sigma$ .....	0	10	0	0	10	14

Źródło: obliczenia własne na podstawie danych umownych.

Stąd na podstawie (1) otrzymujemy:

$$r(X, Y) = \frac{\sum_{t=1}^5 (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^5 (x_t - \bar{x})^2 \sum_{t=1}^5 (y_t - \bar{y})^2}} = \frac{0}{\sqrt{10 \cdot 14}} = 0 \quad (3)$$

Można zauważyć, że wartości zmiennej  $Y$  są kwadratami wartości zmiennej  $X$ . Współczynnik korelacji nieliniowej obliczony na podstawie wzoru (2) jest równy 1.



Inne współczynniki wykorzystywane w modelowaniu ekonometrycznym: korelacji wielowymiarowej, korelacji cząstkowej i semikorelacji cząstkowej – które określono jako współczynniki korelacji liniowej odpowiednich wektorów – mają takie same zalety i wady co omówiony współczynnik korelacji liniowej. Należy pamiętać, że są one miarą zależności liniowej cech ilościowych, więc ich niska wartość bezwzględna nie musi świadczyć o braku innej zależności (co pokazały zaprezentowane wcześniej obliczenia). Niewzięcie tego pod uwagę jest najczęstszym błędem przy interpretacji wartości tych współczynników.

### 3. Współczynnik korelacji wielowymiarowej

W dalszej części artykułu przyjęto, że dany jest model postaci:

$$Y = \alpha_1 Z_1 + \alpha_2 Z_2 + \dots + \alpha_k Z_k + \xi \quad (4)$$

Założmy, że mamy  $n$  obserwacji wartości zmiennych i że zmienne występujące w modelu (4) są standaryzowane (wartości średnie obserwacji są równe 0, a odchylenia standardowe wynoszą 1). Otrzymujemy zatem:

$$\mathbf{Z} = [z_{tj}]_{n \times k} \quad \text{oraz} \quad \mathbf{y} = [y_t]_{n \times 1} \quad (5)$$

Oznaczmy przez  $(\mathbf{R}(k) \mathbf{R}_0(k))$  parę korelacyjną określającą model (4). Jak wiadomo (zob. np. Kolupa i Śleszyński, 2010):

$$\mathbf{R}(k) = \frac{1}{n} \mathbf{Z}^T \mathbf{Z} \quad (6)$$

$$\mathbf{R}_0(k) = \frac{1}{n} \mathbf{Z}^T \mathbf{y}$$

Jeśli oznaczymy:

$$\mathbf{A}^T = [a_1 \ a_2 \ \dots \ a_k]_{k \times 1} \quad (7)$$

wektor oszacowań parametrów strukturalnych modelu (4) uzyskany za pomocą metody najmniejszych kwadratów, to:

$$\mathbf{A} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{y} = \mathbf{R}^{-1} \mathbf{R}_0 \quad (8)$$

Wektor wartości teoretycznych zmiennej objaśnianej  $Y$  ma postać:

$$\mathbf{y}^* = \mathbf{Z} \cdot \mathbf{A} = \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{y} = \mathbf{Z} \cdot \mathbf{R}^{-1} \mathbf{R}_0 \quad (9)$$

Z kolei wektor reszt  $\mathbf{u}$  wyznaczamy zgodnie ze wzorem:

$$\mathbf{u} = \mathbf{y} - \mathbf{y}^* = \mathbf{y} - \mathbf{Z} \cdot \mathbf{A} = \mathbf{y} - \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{y} = \mathbf{y} - \mathbf{Z} \cdot \mathbf{R}^{-1} \mathbf{R}_0 \quad (10)$$

Oznaczmy przez  $R$  współczynnik korelacji liniowej wyznaczony dla wektora wartości empirycznych  $\mathbf{y}$  oraz wektora wartości teoretycznych  $\mathbf{y}^*$  (Charemza i Deadman, 1997). Jest to współczynnik korelacji wielowymiarowej (niekiedy nazywany współczynnikiem korelacji wielokrotnej lub wielorakiej). Przyjmuje on wartości z przedziału  $[0, 1]$  (Theil, 1979) i określa siłę oddziaływania zmiennych objaśniających modelu (4) na jego zmienną objaśnianą. Warto dodać, że współczynnik korelacji nieliniowej  $R(Y, X)$  dany wzorem (2), obliczony dla modelu (4), jest współczynnikiem korelacji wielowymiarowej dla tego modelu. Ponieważ dalsze rozważania dotyczą modelu (4), dla tych dwóch współczynników przyjęto identyczne oznaczenia.

Kwadrat współczynnika korelacji wielowymiarowej nazywamy współczynnikiem determinacji. Wyraża on procent zmienności zmiennej zależnej wytłumaczony przez model. Dla modeli szacowanych na podstawie danych w postaci szeregów czasowych przyjmuje zwykle wysoką wartość (większą od 0,9), natomiast w przypadku równań szacowanych przy wykorzystaniu danych przekrojowych często akceptowane są modele o znacznie niższej wartości współczynnika determinacji (większej od 0,3) (Gajda, 2004).

Mając parę korelacyjną, współczynnik determinacji można wyznaczyć zgodnie ze wzorem:

$$R^2 = \mathbf{R}_0^T \mathbf{R}^{-1} \mathbf{R}_0 \quad (11)$$

W celu wyznaczenia wartości  $R^2$  można posłużyć się macierzą brzegową  $\mathbf{K}$  postaci:

$$\mathbf{K} = \begin{bmatrix} \mathbf{R} & \mathbf{R}_0 \\ -\mathbf{R}_0^T & 0 \end{bmatrix} \quad (12)$$

Jeśli na macierzy  $\mathbf{K}$  danej wzorem (12) dokonamy przekształceń elementarnych, w wyniku których w miejscu macierzy wewnętrznej  $\mathbf{R}$  otrzymamy macierz górną trójkątną z jedynkami na głównej przekątnej  $\mathbf{R}^*$ , a w miejscu wektora  $-\mathbf{R}_0^T$  otrzymamy wektor zerowy, to w miejscu liczby 0 otrzymamy wartość współczynnika determinacji  $R^2$ . W efekcie macierz  $\mathbf{K}$  zostanie przekształcona w macierz  $\mathbf{K}^*$  postaci:

$$\mathbf{K}^* = \begin{bmatrix} \mathbf{R}^* & \mathbf{R}_0^* \\ 0, \dots, 0 & R^2 \end{bmatrix} \quad (13)$$

Liczbę  $1 - R^2$  nazywamy współczynnikiem zbieżności (niekiedy tolerancją) i oznaczamy przez  $\varphi^2$ . Określa on, jaki procent zmienności zmiennej zależnej nie jest wyjaśniony przez zmienne objaśniające modelu, zgodnie ze wzorem:

$$\varphi^2 = 1 - \mathbf{R}_0^T \mathbf{R}^{-1} \mathbf{R}_0 \quad (14)$$

Jeśli mając daną macierz  $\mathbf{P}$  postaci:

$$\mathbf{P} = \begin{bmatrix} \mathbf{R} & \mathbf{R}_0 \\ \mathbf{R}_0^T & 1 \end{bmatrix} \quad (15)$$

dokonamy przekształceń elementarnych, w wyniku których w miejscu macierzy wewnętrznej  $\mathbf{R}$  otrzymamy macierz górną trójkątną z jedynkami na głównej przekątnej  $\mathbf{R}^*$ , a w miejscu wektora  $\mathbf{R}_0^T$  otrzymamy wektor zerowy, to w miejscu liczby 1 otrzymamy wartość współczynnika zbieżności  $\varphi^2$ . W efekcie macierz  $\mathbf{P}$  zostanie przekształcona w macierz  $\mathbf{P}^*$  postaci:

$$\mathbf{P}^* = \begin{bmatrix} \mathbf{R}^* & \mathbf{R}_0^* \\ 0, \dots, 0 & \varphi^2 \end{bmatrix} \quad (16)$$

#### 4. Współczynnik korelacji cząstkowej

Aby wprowadzić definicję współczynnika korelacji cząstkowej  $r_i^*$  zmiennej objaśnianej  $Y$  ze zmienną objaśniającą  $Z_i$  w modelu (4), nazwijmy model (4) modelem pierwotnym, a model powstały z niego przez usunięcie  $i$ -tej zmiennej objaśnianej – modelem uciętym. Jeśli model pierwotny jest określony przez parę korelacyjną  $(\mathbf{R}(k), \mathbf{R}_0(k))$  daną wzorem (6), to model ucięty jest określony przez parę  $(\mathbf{R}_{ii}, \mathbf{R}_{0i})$ , gdzie macierz  $\mathbf{R}_{ii}$  powstaje z macierzy  $\mathbf{R}(k)$  przez skreślenie wiersza i kolumny o numerze  $i$ , a wektor  $\mathbf{R}_{0i}$  powstaje z wektora  $\mathbf{R}_0(k)$  przez skreślenie  $i$ -tej składowej.

Oznaczmy jako  $\mathbf{u}_i$  wektor reszt dla modelu uciętego. Przez analogię do wzoru (10) mamy:

$$\mathbf{u}_i = \mathbf{y} - \mathbf{Z}(-i) \mathbf{R}_{ii}^{-1} \mathbf{R}_{0i} \quad (17)$$

gdzie macierz  $\mathbf{Z}(-i)$  powstaje z macierzy obserwacji  $\mathbf{Z}$  danej wzorem (5) przez odzucenie  $i$ -tej kolumny.

Modelem wewnętrznym dla modelu pierwotnego (6) nazywamy model, w którym zmienną objaśnianą jest  $i$ -ta zmienna objaśniająca modelu pierwotnego, natomiast zbiór zmiennych objaśniających jest identyczny jak w modelu uciętym. Model wewnętrzny określa para korelacyjna  $(\mathbf{R}_{ii}, \boldsymbol{\rho}_i)$ , gdzie  $\boldsymbol{\rho}_i$  oznacza  $i$ -tą kolumnę macierzy  $\mathbf{R}(k)$  bez  $i$ -tego elementu. Oznaczmy jako  $\mathbf{v}_i$  wektor reszt dla modelu wewnętrznego. Przez analogię do (17) otrzymujemy:

$$\mathbf{v}_i = \mathbf{z}_i - \mathbf{Z}(-i)\mathbf{R}_{ii}^{-1}\boldsymbol{\rho}_i \quad (18)$$

gdzie  $\mathbf{z}_i$  –  $i$ -ta kolumna macierzy  $\mathbf{Z}$  danej wzorem (5).

Współczynnikiem korelacji cząstkowej  $r_i^*$  zmiennej objaśnianej  $Y$  ze zmienną objaśniającą  $Z_i$  w modelu (4) nazywamy współczynnik korelacji liniowej pomiędzy wektorami reszt modeli wewnętrznego i uciętego, a zatem:

$$r_i^* = r(\mathbf{u}_i, \mathbf{v}_i) \quad (19)$$

Stąd:

$$r_i^* = \frac{\mathbf{u}_i^T \mathbf{v}_i}{\sqrt{\mathbf{u}_i^T \mathbf{u}_i} \sqrt{\mathbf{v}_i^T \mathbf{v}_i}} \quad (20)$$

Ponieważ spełnione są równości (Kolupa, Pacholewicz i Śleszyński, 1992):

$$\begin{aligned} \mathbf{u}_i^T \mathbf{v}_i &= n(r_i - \boldsymbol{\rho}_i^T \mathbf{R}_{ii}^{-1} \mathbf{R}_{0i}) \\ \mathbf{u}_i^T \mathbf{u}_i &= n(1 - \mathbf{R}_{0i}^T \mathbf{R}_{ii}^{-1} \mathbf{R}_{0i}) \\ \mathbf{v}_i^T \mathbf{v}_i &= n(1 - \boldsymbol{\rho}_i^T \mathbf{R}_{ii}^{-1} \boldsymbol{\rho}_i) \end{aligned} \quad (21)$$

gdzie:

$r_i$  – współczynnik korelacji liniowej zmiennej objaśnianej  $Y$  ze zmienną objaśniającą  $Z_i$ ,  
 $n$  – liczba obserwacji,

to wzór (20) przyjmuje postać:

$$r_i^* = \frac{r_i - \boldsymbol{\rho}_i^T \mathbf{R}_{ii}^{-1} \mathbf{R}_{0i}}{\sqrt{(1 - \mathbf{R}_{0i}^T \mathbf{R}_{ii}^{-1} \mathbf{R}_{0i})(1 - \boldsymbol{\rho}_i^T \mathbf{R}_{ii}^{-1} \boldsymbol{\rho}_i)}} \quad (22)$$

Wartości współczynnika korelacji cząstkowej zawierają się w przedziale  $[-1, 1]$ . Jest on miarą korelacji pomiędzy daną zmienną objaśniającą  $Z_i$ , z uwzględnieniem jej skorelowania z pozostałymi zmiennymi objaśniającymi modelu, a zmienną objaśnianą  $Y$ , z uwzględnieniem jej skorelowania z pozostałymi zmiennymi objaśniającymi.

Kwadrat współczynnika korelacji cząstkowej nazywamy współczynnikiem determinacji cząstkowej.

Zauważmy, że jeśli we wzorze (6) przyjmiemy  $k = 2$ , to otrzymamy wyniki analogiczne do podanych przez Zeliasia (2000). Mamy bowiem:

$$\mathbf{R}(2) = \begin{bmatrix} 1 & r_{12} \\ r_{12} & 1 \end{bmatrix}, \mathbf{R}_0(2) \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \quad (23)$$

Wtedy przykładowo dla  $i = 1$ :

$$\boldsymbol{\rho}_1 = [r_{12}], \mathbf{R}_{11} = [1], \mathbf{R}_{01} = [r_2] \quad (24)$$

$$r_i^* = \frac{r_1 - \boldsymbol{\rho}_1^T \mathbf{R}_{11}^{-1} \mathbf{R}_{01}}{\sqrt{(1 - \mathbf{R}_{01}^T \mathbf{R}_{11}^{-1} \mathbf{R}_{01})(1 - \boldsymbol{\rho}_1^T \mathbf{R}_{11}^{-1} \boldsymbol{\rho}_1)}} = \frac{r_1 - r_{12}r_2}{\sqrt{(1 - r_2^2)(1 - r_{12}^2)}} \quad (25)$$

Z kolei dla  $i = 2$  otrzymujemy:

$$r_2^* = \frac{r_2 - \boldsymbol{\rho}_2^T \mathbf{R}_{22}^{-1} \mathbf{R}_{02}}{\sqrt{(1 - \mathbf{R}_{02}^T \mathbf{R}_{22}^{-1} \mathbf{R}_{02})(1 - \boldsymbol{\rho}_2^T \mathbf{R}_{22}^{-1} \boldsymbol{\rho}_2)}} = \frac{r_2 - r_{12}r_1}{\sqrt{(1 - r_1^2)(1 - r_{12}^2)}} \quad (26)$$

Warto nadmienić, że aby wyznaczyć współczynnik korelacji cząstkowej  $r_i^*$  zgodnie ze wzorem (22), wygodnie jest posłużyć się podwójną macierzą brzegową  $\Delta_i$  postaci:

$$\Delta_i = \begin{bmatrix} \mathbf{R}_{ii} & \boldsymbol{\rho}_i & \mathbf{R}_{0i} \\ \boldsymbol{\rho}_i^T & 1 & r_i \\ \mathbf{R}_{0i}^T & r_i & 1 \end{bmatrix} \quad (27)$$

Jeśli na elementach macierzy  $\Delta_i$  wykonamy przekształcenia elementarne, otrzymamy macierz  $(\Delta_i)^*$ , w której w miejsce macierzy wewnętrznej  $\mathbf{R}_{ii}$  będziemy mieć macierz górną trójkątną z jedynkami na głównej przekątnej (oznaczymy ją  $(\mathbf{R}_{ii})^*$ , natomiast w miejsce wektorów  $\boldsymbol{\rho}_i^T$  oraz  $\mathbf{R}_{0i}^T$  – wektory zerowe, czyli:

$$\Delta_i = \begin{bmatrix} \mathbf{R}_{ii} & \boldsymbol{\rho}_i & \mathbf{R}_{0i} \\ \boldsymbol{\rho}_i^T & 1 & r_i \\ \mathbf{R}_{0i}^T & r_i & 1 \end{bmatrix} \sim (\Delta_i)^* = \begin{bmatrix} (\mathbf{R}_{ii})^* & (\boldsymbol{\rho}_i)^* & (\mathbf{R}_{0i})^* \\ 0, \dots, 0 & a_i & b_i \\ 0, \dots, 0 & b_i & c_i \end{bmatrix} \quad (28)$$

Wówczas na podstawie (22) otrzymamy:

$$r_i^* = \frac{b_i}{\sqrt{a_i \cdot c_i}} \quad (29)$$

Warto zwrócić uwagę, że wykorzystując macierz (28), możemy dla dowolnego  $i = 1, 2, \dots, k$  wyznaczyć współczynnik zbieżności  $\varphi^2$  dla modelu pierwotnego (4). Spełniona jest bowiem równość:

$$\varphi^2 = c_i - \frac{b_i^2}{a_i} \quad (30)$$

Spełnione są także następujące zależności (Theil, 1979):

$$\varphi^2 = \varphi_{(-i)}^2 (1 - (r_i^*)^2) \quad (31)$$

gdzie  $\varphi_{(-i)}^2$  – współczynnik zbieżności dla modelu uciętego.

Jeśli przez  $R_{(-i)}^2$  oznaczymy współczynnik determinacji dla modelu uciętego, to równość (31) można zapisać w nieco innej postaci:

$$R^2 - R_{(-i)}^2 = (r_i^*)^2 (1 - R_{(-i)}^2) \quad (32)$$

Z równości (32) wynika, że dodatkowy udział  $i$ -tej zmiennej objaśniającej jest równy iloczynowi kwadratu współczynnika korelacji cząstkowej tej zmiennej ze zmienną objaśnianą w danym modelu i współczynnika zbieżności modelu uciętego. Jeśli współczynnik zbieżności modelu uciętego jest mały, to nie ma już miejsca na kolejną zmienną objaśniającą, dlatego jej dodatkowy udział w wyjaśnianiu zmiennej zależnej będzie niewielki.

## 5. Współczynnik semikorelacji cząstkowej

Podobny do współczynnika korelacji cząstkowej jest współczynnik semikorelacji cząstkowej  $r_i^{(*)}$ , będący miarą korelacji pomiędzy daną zmienną niezależną  $Z_i$ , z uwzględnieniem jej skorelowania z pozostałymi zmiennymi objaśniającymi modelu,

a zmienną objaśnianą  $Y$ , bez uwzględnienia jej skorelowania z pozostałymi zmiennymi objaśniającymi modelu. Można zatem powiedzieć, że jest to współczynnik korelacji liniowej pomiędzy wektorem reszt modelu wewnętrznego a wektorem wartości zmiennej objaśnianej  $Y$ . Wobec tego:

$$r_i^{(*)} = r(\mathbf{v}_i, \mathbf{y}) \quad (33)$$

Ponieważ zmienne w modelu pierwotnym (a więc i w wewnętrznym) są standaryzowane, możemy zapisać:

$$r_i^{(*)} = \frac{\mathbf{v}_i^T \mathbf{y}}{\sqrt{\mathbf{y}^T \mathbf{y}} \sqrt{\mathbf{v}_i^T \mathbf{v}_i}} \quad (34)$$

Wykorzystując (18), otrzymujemy:

$$\begin{aligned} \mathbf{v}_i^T \mathbf{y} &= (z_i - \mathbf{Z}(-i) \mathbf{R}_{ii}^{-1} \boldsymbol{\rho}_i)^T \mathbf{y} = \\ &= nr_i - \boldsymbol{\rho}_i^T \mathbf{R}_{ii}^{-1} (\mathbf{Z}(-i))^T \mathbf{y} = \\ &= nr_i - n \boldsymbol{\rho}_i^T \mathbf{R}_{ii}^{-1} \mathbf{R}_{0i} \end{aligned} \quad (35)$$

Na podstawie (21) – ponieważ zmienna zależna  $Y$  jest standaryzowana – wzór (34) przyjmuje postać:

$$r_i^{(*)} = \frac{\mathbf{v}_i^T \mathbf{y}}{\sqrt{\mathbf{y}^T \mathbf{y}} \sqrt{\mathbf{v}_i^T \mathbf{v}_i}} = \frac{r_i - \boldsymbol{\rho}_i^T \mathbf{R}_{ii}^{-1} \mathbf{R}_{0i}}{\sqrt{1 - \boldsymbol{\rho}_i^T \mathbf{R}_{ii}^{-1} \boldsymbol{\rho}_i}} \quad (36)$$

Warto zauważyć, że w celu obliczenia współczynnika semikorelacji cząstkowej możemy również posłużyć się podwójną macierzą brzegową (27). Na podstawie (28) otrzymujemy:

$$r_i^{(*)} = \frac{b_i}{\sqrt{a_i}} \quad (37)$$

Współczynnik semikorelacji cząstkowej jest liczbą z przedziału  $[-1, 1]$ .

Kwadrat współczynnika semikorelacji cząstkowej – współczynnik determinacji semicząstkowej – oznacza stosunek wyłącznej zmienności danej zmiennej objaśniającej  $Z_i$  do całkowitej zmienności zmiennej objaśnianej  $Y$ . Przyjmuje wartości z przedziału  $[0, 1]$ . Podobnie jak w wypadku współczynnika determinacji cząstkowej jego wartość bliska 0 oznacza, że zmienna

objaśniająca niesie bezużyteczną informację o zmiennej objaśnianej w danym modelu, a zatem jest nadmiarowa (StatSoft, 2006).

Zauważmy, że:

$$r_i^* = \frac{r_i^{(*)}}{\sqrt{1 - \mathbf{R}_{0i}^T \mathbf{R}_{ii}^{-1} \mathbf{R}_{0i}}} = \frac{r_i^{(*)}}{\sqrt{\varphi_{(-i)}^2}} \quad (38)$$

Na podstawie (38) stwierdzamy, że współczynnik semikorelacji cząstkowej jest zawsze co najwyżej równy współczynnikowi korelacji cząstkowej oraz że oba mają identyczne znaki.

Równość (38) można zapisać też w innej, równoważnej postaci:

$$\left(r_i^{(*)}\right)^2 = (r_i^*)^2 \varphi_{(-i)}^2 \quad (39)$$

gdzie  $\varphi_{(-i)}^2$  – współczynnik zbieżności (tolerancja) odpowiedniego modelu uciętego.

## 6. Porównanie omówionych współczynników korelacji liniowej

Zanim przejdziemy do przykładu prezentującego wyznaczenie i interpretację omówionych w artykule współczynników korelacji liniowej, podsumujemy ich wybrane własności.

Współczynnik korelacji wielowymiarowej wyznacza się dla wektorów wartości empirycznej i teoretycznej zmiennej objaśnianej modelu. Jest zawsze nieujemny, a jego wartość jest nie mniejsza od największej wartości bezwzględnej współczynnika korelacji liniowej zmiennej objaśnianej i objaśniającej. Jego kwadrat to współczynnik determinacji, wyrażający procent zmienności zmiennej zależnej wyjaśniony przez model.

Współczynnik korelacji cząstkowej danej zmiennej objaśniającej ze zmienną objaśnianą modelu, czyli współczynnik korelacji liniowej wektorów reszt modeli uciętego i wewnętrznego, jest miarą korelacji liniowej pomiędzy zmienną objaśnianą i objaśniającą z uwzględnieniem ich skorelowania z pozostałymi zmiennymi objaśniającymi. Jego wartość może być mniejsza lub większa od wartości odpowiedniego współczynnika korelacji liniowej  $r_i$ , a znak jest zgodny ze znakiem oszacowania parametru strukturalnego przy danej zmiennej objaśniającej modelu. Iloczyn kwadratu współczynnika korelacji cząstkowej i współczynnika zbieżności modelu uciętego wyraża dodatkowy udział danej zmiennej objaśniającej w modelu mierzony wzrostem wartości współczynnika determinacji w stosunku do współczynnika determinacji modelu uciętego. Mały dodatkowy udział świadczy o tym, że zmienna jest zbędna w modelu.



Współczynnik semikorelacji cząstkowej jest współczynnikiem korelacji liniowej pomiędzy wektorem reszt modelu wewnętrznego a wektorem wartości empirycznych zmiennej objaśnianej  $Y$ . Stanowi on miarę korelacji liniowej pomiędzy daną zmienną niezależną  $Z_i$ , z uwzględnieniem jej skorelowania z pozostałymi zmiennymi objaśniającymi modelu, a zmienną objaśnianą  $Y$ , bez uwzględnienia jej skorelowania z pozostałymi zmiennymi objaśniającymi modelu. Jest zawsze co najwyżej równy współczynnikowi korelacji cząstkowej (oba zawsze mają identyczne znaki) i jednocześnie nie większy od odpowiedniego współczynnika korelacji liniowej  $r_i$ .

Współczynnik determinacji semicząstkowej (czyli kwadrat współczynnika semikorelacji cząstkowej) jest iloczynem współczynnika zbieżności modelu uciętego oraz kwadratu współczynnika korelacji cząstkowej. Jego duża wartość bezwzględna świadczy o tym, że zmienna objaśniająca, po wyeliminowaniu z niej wpływu pozostałych zmiennych objaśniających modelu, ma istotny wpływ na zmienną zależną modelu.

Każdy z omówionych współczynników korelacji eksponuje inne własności zmiennych w modelu, dlatego wskazane jest najpierw wyznaczenie wartości ich wszystkich i dopiero na tej podstawie wyciąganie wniosków o jakości zmiennej czy modelu.

## 7. Przykład praktyczny

Na zakończenie prezentujemy przykład wyznaczania omówionych współczynników korelacji na podstawie modelu wzrostu płac w Polsce, szacowanego metodą najmniejszych kwadratów, z wykorzystaniem danych za lata 2001–2019 zaczerpniętych z Banku Danych Makroekonomicznych GUS. Wykonane obliczenia mają wyłącznie charakter ilustrujący wcześniejsze rozważania.

Rozpatrzmy następujący model:

$$y_t = \alpha_0 z_{0t} + \alpha_1 z_{1t} + \alpha_2 z_{2t} + \alpha_3 z_{3t} + \alpha_4 z_{4t} + \xi_t \quad (40)$$

gdzie:

$y$  – wzrost przeciętnego rocznego wynagrodzenia brutto w %,

$z_0$  – zmienna tożsamościowo równa 1,

$z_1$  – zmienna czasowa w latach,

$z_2$  – zmiana cen towarów i usług konsumpcyjnych w %,

$z_3$  – zmiana PKB brutto *per capita* w %,

$z_4$  – stopa bezrobocia rejestrowanego w %.

W modelu wykorzystano dane empiryczne za 19 lat, a zatem liczba obserwacji  $n = 19$ . Dane zamieszczono w tabl. 2.

**Tabl. 2.** Wyniki obserwacji zmiennych modelu (40)

$t$	$y_t$	$z_{0t}$	$z_{1t}$	$z_{2t}$	$z_{3t}$	$z_{4t}$
2001 .....	2,5	1	1	5,5	4,44	17,5
2002 .....	0,7	1	2	1,9	3,99	19,0
2003 .....	3,4	1	3	0,8	4,43	20,0
2004 .....	0,7	1	4	3,5	10,35	19,0
2005 .....	1,8	1	5	2,1	6,20	17,6
2006 .....	4,0	1	6	1,0	8,10	14,8
2007 .....	5,5	1	7	2,5	11,04	11,2
2008 .....	5,9	1	8	4,2	8,23	9,5
2009 .....	2,0	1	9	3,5	6,61	12,1
2010 .....	1,4	1	10	2,6	4,34	12,4
2011 .....	1,4	1	11	4,3	8,38	12,5
2012 .....	0,1	1	12	3,7	4,01	13,4
2013 .....	2,8	1	13	0,9	1,79	13,4
2014 .....	3,2	1	14	0,0	3,88	11,4
2015 .....	4,5	1	15	-0,9	4,75	9,7
2016 .....	4,3	1	16	-0,6	3,39	8,2
2017 .....	3,7	1	17	2,0	6,89	6,6
2018 .....	5,4	1	18	1,6	6,39	5,8
2019 .....	4,8	1	19	2,3	7,13	5,2

Źródło: Bank Danych Makroekonomicznych GUS.

Na podstawie danych z tabl. 2, po dokonaniu standaryzacji zmiennych, zgodnie ze wzorem (6) wyznaczono parę korelacyjną określającą model (40):

$$\mathbf{R}(4) = \begin{bmatrix} 1 & -0,411 & -0,160 & -0,900 \\ -0,411 & 1 & 0,395 & 0,215 \\ -0,160 & 0,395 & 1 & -0,083 \\ -0,900 & 0,215 & -0,083 & 1 \end{bmatrix}, \quad \mathbf{R}_0(4) = \begin{bmatrix} 0,403 \\ -0,339 \\ 0,237 \\ -0,631 \end{bmatrix} \quad (41)$$

Z analizy uzyskanych wartości współczynników korelacji liniowej zmiennej endogenicznej z poszczególnymi zmiennymi objaśniającymi modelu (40) wynika, że w badanym okresie z roku na rok następował wzrost przeciętnego wynagrodzenia, ale nie nadążał on za wzrostem cen (współczynnik korelacji liniowej tych cech wynosił  $-0,339$ ). Analogicznie zwiększeniu wzrostu wydajności pracy (mierzonej wzrostem PKB *per capita*) towarzyszył wzrost wynagrodzeń. Wzrost bezrobocia powodował zaś spadek wartości zmiennej objaśnianej.

Za pomocą metody najmniejszych kwadratów wyznaczono oszacowania parametrów modelu o zmiennych standaryzowanych, określonego przez parę korelacyjną ( $\mathbf{R}(4), \mathbf{R}_0(4)$ ) daną wzorem (41)¹:

¹ W celu uproszczenia zapisu zmienne i parametry modelu o zmiennych standaryzowanych (41) oraz modelu wyjściowego (39) oznaczono identycznie.

$$y_t = \alpha_1 z_{1t} + \alpha_2 z_{2t} + \alpha_3 z_{3t} + \alpha_4 z_{4t} + \xi_t \quad (42)$$

Wykorzystując (8), otrzymano:

$$A = R^{-1}R_0 = \begin{bmatrix} -1,408 \\ -0,574 \\ 0,093 \\ -1,767 \end{bmatrix} \quad (43)$$

Warto zwrócić uwagę na rozbieżność znaków oszacowania parametru stojącego przy pierwszej zmiennej objaśniającej modelu (czas) oraz współczynnika korelacji liniowej tej zmiennej ze zmienną objaśnianą, a zatem brak koincydencji zmiennej czasowej. Można stąd wysnuć wniosek, że analizowany model wskazuje na zbyt wolne tempo wzrostu płać.

Do wyznaczenia współczynnika korelacji cząstkowej  $r_1^*$  zmiennej objaśnianej  $Y$  ze zmienną objaśniającą  $Z_1$  w modelu (40), zgodnie ze wzorem (22), wykorzystano podwójną macierz brzegową (27). Stąd:

$$\Delta_1 = \begin{bmatrix} R_{11} & \rho_1 & R_{01} \\ \rho_1^T & 1 & r_1 \\ R_{01}^T & r_1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0,395 & 0,215 & -0,411 & -0,339 \\ 0,395 & 1 & -0,083 & -0,160 & 0,237 \\ 0,215 & -0,083 & 1 & 0,900 & -0,631 \\ -0,411 & -0,160 & 0,900 & 1 & 0,403 \\ -0,339 & 0,237 & -0,631 & 0,403 & 1 \end{bmatrix} \quad (44)$$

W wyniku przekształceń elementarnych wykonanych na macierzy (44) zgodnie ze wzorem (28) otrzymano:

$$(\Delta_1)^* = \begin{bmatrix} (R_{11})^* & (\rho_1)^* & (R_{01})^* \\ 0, \dots, 0 & \alpha_1 & b_1 \\ 0, \dots, 0 & b_1 & c_1 \end{bmatrix} = \begin{bmatrix} 1 & -0,395 & 0,215 & -0,411 & -0,339 \\ 0 & 1 & -0,199 & 0,003 & 0,439 \\ 0 & 0 & 1 & -0,881 & -0,526 \\ 0 & 0 & 0 & 0,117 & -0,164 \\ 0 & 0 & 0 & -0,164 & 0,467 \end{bmatrix} \quad (45)$$

Na podstawie (29), wykorzystując (45), można już wyznaczyć współczynnik korelacji cząstkowej  $r_1^*$ :

$$r_1^* = \frac{b_1}{\sqrt{a_1 \cdot c_1}} = \frac{-0,164}{\sqrt{0,117 \cdot 0,467}} = -0,704 \quad (46)$$

Współczynnik  $r_1^*$  ma wartość ujemną – znak współczynnika korelacji cząstkowej jest zawsze zgodny ze znakiem oszacowania parametru stojącego przy danej zmiennej. Wysoka wartość bezwzględna tego współczynnika świadczy o tym, że pierwsza zmienna objaśniająca modelu ma istotny wpływ na część zmiennej zależnej niewyjaśnioną przez pozostałe zmienne objaśniające. Zgodnie ze wzorem (30) współczynnik zbieżności dla modelu pierwotnego (42) jest równy:

$$\varphi^2 = c_1 - \frac{b_1^2}{a_1} = 0,467 = \frac{(-0,164)^2}{0,117} = 0,236 \quad (47)$$

Współczynnik zbieżności  $\varphi_{(-1)}^2$  dla modelu uciętego wynosi:

$$\varphi_{(-1)}^2 = c_1 = 0,467 \quad (48)$$

Następnie wyznaczono dodatkowe udziały zmiennej  $Z_1$  modelu pierwotnego. Zgodnie ze wzorem (32) otrzymano:

$$\varphi_{(-1)}^2 - \varphi^2 = (r_1^*)^2(1 - R_{(-1)}^2) = (r_1^*)^2\varphi_{(-1)}^2 \quad (49)$$

$$\varphi_{(-1)}^2 - \varphi^2 = (-0,704)^2 \cdot 0,467 = 0,232$$

Dodatkowy udział zmiennej  $Z_1$  jest wysoki. Zmienna ta w sposób istotny podnosi jakość modelu (40), mierzoną wartością współczynnika zbieżności.

Współczynnik semikorelacji cząstkowej wyznaczono na podstawie (37), wykorzystując (45). Otrzymano:

$$r_1^{(*)} = \frac{b_1}{\sqrt{a_1}} = \frac{-0,164}{\sqrt{0,117}} = -0,481 \quad (50)$$

Duża wartość bezwzględna tego współczynnika świadczy o przydatności zmiennej  $Z_1$  w modelu. Pierwsza zmienna objaśniająca modelu, po wyeliminowaniu z niej wpływu pozostałych zmiennych niezależnych, istotnie wpływa na zmienną zależną.

Na zakończenie rozważań dotyczących pierwszej zmiennej zauważmy, że liczba  $a_1$ , otrzymana w macierzy (45), jest współczynnikiem zbieżności dla modelu wewnętrznego (w którym zmienną objaśnianą jest zmienna  $Z_1$ , a zbiór zmiennych objaśniających jest identyczny ze zbiorem w modelu uciętym). Jeśli więc przez  $R_{(w1)}^2$  oznaczymy współczynnik determinacji dla tego modelu wewnętrznego, to:

$$R_{(w1)}^2 = 1 - a_1 = 1 - 0,117 = 0,883 \quad (51)$$

Dla zmiennej  $Z_2$  otrzymano:

$$\Delta_2 = \begin{bmatrix} \mathbf{R}_{22} & \boldsymbol{\rho}_{12} & \mathbf{R}_{02} \\ \boldsymbol{\rho}_2^T & 1 & r_2 \\ \mathbf{R}_{02}^T & r_2 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -0,160 & -0,900 & -0,411 & 0,403 \\ -0,160 & 1 & -0,083 & 0,395 & 0,237 \\ -0,900 & -0,083 & 1 & 0,215 & -0,631 \\ -0,411 & 0,395 & 0,215 & 1 & -0,339 \\ 0,403 & 0,237 & -0,631 & -0,339 & 1 \end{bmatrix} \quad (52)$$

W wyniku przekształceń elementarnych wykonanych na macierzy (52), zgodnie ze wzorem (28) mamy:

$$(\Delta_2)^* = \begin{bmatrix} (\mathbf{R}_{22})^* & (\boldsymbol{\rho}_2)^* & (\mathbf{R}_{02})^* \\ 0, \dots, 0 & \alpha_2 & b_2 \\ 0, \dots, 0 & b_2 & c_2 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -0,160 & -0,900 & -0,411 & 0,403 \\ 0 & 1 & -0,232 & 0,338 & 0,309 \\ 0 & 0 & 1 & -0,565 & -1,443 \\ 0 & 0 & 0 & 0,676 & -0,388 \\ 0 & 0 & 0 & -0,388 & 0,458 \end{bmatrix} \quad (53)$$

$$r_2^* = \frac{b_2}{\sqrt{a_2 \cdot c_2}} = \frac{-0,388}{\sqrt{0,676 \cdot 0,458}} = -0,697 \quad (54)$$

Druga zmienna objaśniająca modelu, po wyeliminowaniu z niej wpływu pozostałych trzech zmiennych niezależnych, istotnie wpływa na zmienną zależną:

$$\varphi^2 = c_2 - \frac{b_2^2}{a_2} = 0,458 - \frac{(-0,388)^2}{0,676} = 0,236 \quad (55)$$

$$\varphi_{(-2)}^2 = c_2 = 0,458 \quad (56)$$

Dodatkowe udziały zmiennej  $Z_2$  modelu pierwotnego wynoszą:

$$\varphi_{(-2)}^2 - \varphi^2 = (r_2^*)^2 (1 - R_{(-2)}^2) = (r_2^*)^2 \varphi_{(-2)}^2 \quad (57)$$

$$\varphi_{(-2)}^2 - \varphi^2 = (-0,697)^2 \cdot 0,458 = 0,223$$

Jak widać, dodatkowy udział zmiennej  $Z_2$  również jest wysoki i zmienna ta jest przydatna w modelu.

Współczynnik semikorelacji cząstkowej wynosi:

$$r_2^{(*)} = \frac{b_2}{\sqrt{a_2}} = \frac{-0,388}{\sqrt{0,676}} = -0,472 \quad (58)$$

Jego duża wartość bezwzględna świadczy o tym, że zmienna  $Z_2$  wnosi istotne informacje do modelu. Druga zmienna objaśniająca modelu, po wyeliminowaniu z niej wpływu pozostałych zmiennych niezależnych, istotnie wpływa na zmienną zależną.

Współczynnik determinacji dla odpowiedniego modelu wewnętrznego jest równy:

$$R_{(w2)}^2 = 1 - a_2 = 1 - 0,676 = 0,324 \quad (59)$$

Analogiczne obliczenia należałoby przeprowadzić dla dwóch ostatnich zmiennych modelu (42). W tym miejscu ograniczymy się do podania w postaci tabelarycznej wyników końcowych, które zostały uzyskane z wykorzystaniem programu Statistica 13.3 PL (tabl. 3). Zauważmy, że dla dwóch pierwszych zmiennych są one identyczne z wynikami uzyskanymi w wyniku zastosowania macierzy brzegowej (27), co potwierdza poprawność przeprowadzonych obliczeń.

**Tabl. 3.** Udziały zmiennych wyróżnionych w modelu (40)

Zmienne	Oszacowania parametrów	Współczynnik			
		korelacji cząstkowej	semikorelacji cząstkowej	zbieżności modeli wewnętrznych $\varphi_{(wi)}^2$ (tolerancja)	determinacji modeli wewnętrznych $R_{(wi)}^2$
$Z_1$ .....	-1,408	-0,704	-0,481	0,117	0,883
$Z_2$ .....	-0,574	-0,697	-0,472	0,676	0,324
$Z_3$ .....	0,093	0,155	0,076	0,675	0,325
$Z_4$ .....	-1,767	-0,795	-0,635	0,129	0,871

Źródło: obliczenia własne z wykorzystaniem programu Statistica 13.3 PL.

Analizując wyniki zawarte w tabl. 3, a także wektor korelacji dany wzorem (41), należy stwierdzić, że mając model regresji liniowej, warto wyznaczyć współczynniki korelacji liniowej i cząstkowej oraz semikorelacji cząstkowej i dopiero na tej podstawie wyciągać wnioski, a nie opierać się tylko na jednym z tych współczynników, każdy bowiem dostarcza innych informacji. Dopiero niska wartość bezwzględna wszystkich trzech współczynników może świadczyć o nieprzydatności danej zmiennej w modelu (w omawianym przykładzie taką zmienną była  $Z_3$  – zmiana PKB brut-

to *per capita*, która miała przybliżać zmianę wydajności pracy). Z kolei zmienna  $Z_2$  (indeks cen towarów i usług) ma niską wartość bezwzględną współczynnika korelacji liniowej ze zmienną objaśnianą (0,335), ale wysokie wartości bezwzględne współczynników korelacji cząstkowej i semikorelacji cząstkowej. Świadczy to o wyjaśnieniu znacznej części zmiennej objaśnianej, która nie została wyjaśniona przez inne zmienne egzogeniczne modelu, i przydatności tej zmiennej. Podobnie duży zasób nowych informacji o zmiennej endogenicznej wnoszą dwie pozostałe zmienne i dlatego powinny pozostać w omawianym modelu.

## 8. Podsumowanie

Przedstawiona w artykule definicja współczynników korelacji wielowymiarowej, korelacji cząstkowej i semikorelacji cząstkowej jako współczynników korelacji liniowej Pearsona odpowiednich wektorów pozwala na ujednoczenie ich procesu obliczeniowego oraz ich interpretacji. Z kolei zastosowanie podwójnej macierzy brzegowej znacznie ułatwia wyznaczanie tych współczynników, a obliczenia są dzięki temu przejrzystsze. Zaprezentowany przez autora sposób postępowania odbiega od powszechnie spotykanego w literaturze, który w wielu przypadkach polega na podawaniu wzorów na poszczególne współczynniki korelacji wykorzystujących wyznaczniki odpowiednich macierzy.

Zaproponowane podejście ma znaczenie praktyczne i edukacyjne. Zastosowanie metody obliczania z wykorzystaniem macierzy brzegowych nawet w przypadku stosunkowo licznej próby nie nastręcza większych trudności – wystarczy w tym celu napisać kilka formuł w arkuszu kalkulacyjnym. Obliczenia są przejrzyste i zrozumiałe, dzięki temu bardziej intuicyjne, a jednocześnie cechuje je duża dokładność. Jest to warte polecenia, szczególnie dla początkujących badaczy oraz studentów przed zapoznaniem się z pakietami statystycznymi. Należy przy tym podkreślić, że stosując przedstawione w artykule podejście, definicje i wzory, można wyznaczyć współczynniki korelacji wielowymiarowej, cząstkowej i semicząstkowej za pomocą arkusza kalkulacyjnego i funkcji współczynnika korelacji liniowej odpowiednich wektorów, nawet bez posługiwania się macierzami brzegowymi czy pakietem statystycznym.

## Bibliografia

- Charemza, W. W., Deadman, D. F. (1997). *Nowa ekonometria*. Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Dziechciarz, J. (red.). (2015). *Ekonometria: metody, przykłady, zadania*. Wrocław: Wydawnictwo Uniwersytetu Ekonomicznego.
- Ezekiel, M., Fox, K. A. (1959). *Methods of Correlation and Regression Analysis: Linear and Curvilinear*. New York, London: Wiley & Sons.

Gajda, J. B. (2004). *Ekonometria*. Warszawa: Wydawnictwo C.H. Beck.

Kolupa, M., Pacholewicz, E., Śleszyński, Z. (1992). *Metody algebry liniowej w ekonometrii*. Radom: Wydawnictwo Wyższej Szkoły Inżynierskiej.

Kolupa, M., Śleszyński, Z. (2010). *Metody ekonometryczne*. Radom: Wydawnictwo Politechniki Radomskiej.

StatSoft. (2006) *Elektroniczny Podręcznik Statystyki PL*. Pobrane z: <https://www.statsoft.pl/textbook/stathome.html>.

Theil, H. (1979). *Zasady ekonometrii*. Warszawa: Państwowe Wydawnictwo Naukowe.

Zeliaś, A. (2000). *Metody statystyczne*. Warszawa: Polskie Wydawnictwo Ekonomiczne.



## WYDAWNICTWA GUS. MAJ 2020 PUBLICATIONS OF STATISTICS POLAND. MAY 2020

W ofercie wydawniczej Głównego Urzędu Statystycznego z ubiegłego miesiąca warto zwrócić uwagę na następujące publikacje:

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



**Tytuł:** *Portrety polskich regionów 2020*

**Title:** *Portraits of Polish regions 2020*

**Język:** polski, angielski

**Language:** Polish, English

**Dodatkowe informacje:** opracowanie dostępne w wersji elektronicznej

**Additional information:** publication available in the electronic version

Publikacja, wydawana w cyklu pięcioletnim, jest cennym źródłem informacji o rozwoju województw. Zawiera opis zachodzących w nich zjawisk społecznych, gospodarczych i środowiskowych. Część informacji przedstawiono w układzie regionów i podregionów, co obrazuje ich wewnętrzne zróżnicowanie. Każde województwo scharakteryzowano w możliwie ujednolicony sposób, z uwzględnieniem jego specyfiki i cech charakterystycznych. Komentarze analityczne bazują na wynikach badań statystycznych z zakresu demografii, rynku pracy, gospodarki oraz stanu i ochrony środowiska. W celu zilustrowania zmian w czasie niektóre zagadnienia zostały przedstawione w ujęciu dynamicznym, a także w relacji do przeciętnych wielkości dla kraju.



**Tytuł:** *Zielone Płuca Polski w 2018 r.*

**Title:** *Green Lungs of Poland in 2018*

**Język:** polski (przedmowa, spis treści i 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

Opracowanie, ukazujące się co trzy lata, stanowi kompendium wiedzy o Zielonych Płucach Polski – czyli północno-wschodniej części kraju wyodrębnionej ze względu na unikatowość przyrodniczą i kulturową – w świetle danych statystycznych.

nych. Przedstawiono w nim m.in. charakterystykę warunków naturalnych, stanu i ochrony środowiska, sytuacji demograficznej i społecznej, infrastruktury komunalnej i mieszkaniowej oraz kultury i turystyki na tym obszarze. Najnowszą edycję wzbogacono o komentarz analityczny oraz informacje o pomocy społecznej i odpadach komunalnych zbieranych selektywnie. Dane ujęto w przekrojach wojewódzkich i na tle kraju w 2018 r., a niektóre informacje zaprezentowano także na poziomie powiatów i gmin. W uwagach metodycznych podano definicje podstawowych pojęć, niezbędne do właściwej interpretacji danych.

W maju br. ukazały się ponadto:

- „Biuletyn statystyczny” nr 4/2020;
- *Ceny robót budowlano-montażowych i obiektów budowlanych (marzec 2020 r.);*
- *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach 2000–2020 (maj 2020). Z pogłębioną prezentacją wyników dla sekcji zakwaterowanie i gastronomia;*
- *Polska w liczbach 2020* [folder];
- *Produkcja ważniejszych wyrobów przemysłowych w kwietniu 2020 r.;*
- *Rocznik Statystyczny Rolnictwa 2019;*
- *Skup i ceny produktów rolnych w 2019 r.;*
- *Sytuacja społeczno-gospodarcza kraju w kwietniu 2020 r.;*
- „Wiadomości Statystyczne. The Polish Statistician” nr 4/2020;
- „Wiadomości Statystyczne. The Polish Statistician” nr 5/2020;
- *Wyniki finansowe przedsiębiorstw niefinansowych I–XII 2019;*
- *Zeszyt metodologiczny. Ochrona przyrody;*
- *Zeszyt metodologiczny. Ochrona środowiska;*
- *Zeszyt metodologiczny. Świadczenia na rzecz rodziny i ich beneficjenci.*

Wersje elektroniczne wszystkich publikacji GUS są dostępne na stronie [stat.gov.pl/publikacje/publikacje-a-z](http://stat.gov.pl/publikacje/publikacje-a-z).

Electronic versions of all the publications by Statistics Poland are available at [stat.gov.pl/en/publications](http://stat.gov.pl/en/publications).

**Justyna Gustyn** (Główny Urząd Statystyczny, Departament Opracowań Statystycznych)

## DLA AUTORÓW FOR THE AUTHORS

(for the English translation of the information given below, please visit [ws.stat.gov.pl/ForAuthors](http://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 Ministerstwa Nauki i Szkolnictwa Wyższego. Zgodnie z komunikatem MNiSW z dnia 31 lipca 2019 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych wraz z przypisaną liczbą punktów „WS” otrzymały 20 punktów.

„Wiadomości Statystyczne. The Polish Statistician” są udostępniane w następujących bazach indeksacyjnych i repozytoriach: Agro, BazEkon, Central and Eastern European Online Library (CEEOL), Central European Journal of Social Sciences and Humanities (CEJSH), 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](mailto:redakcja.ws@stat.gov.pl).

Artykuł powinien być utrzymany w formie bezosobowej i zawierać streszczenie, słowa kluczowe oraz kod/kody JEL. Tytuł, streszczenie i słowa kluczowe powinny być podane w języku polskim i angielskim. Jeżeli w pracy występują tablice, wykresy lub mapy, powinny być umieszczone w treści artykułu. W osobnym pliku należy podać dane do wykresów. **Prosimy o niestosowanie stylów i ograniczenie formatowania do wymogów redakcyjnych.** Więcej informacji w podrozdziale *Wymogi redakcyjne* i następnych podrozdziałach.

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

Redakcja rozpoczyna postępowanie kwalifikujące artykuł do opublikowania po przedłożeniu przez autora oświadczenia o przeniesieniu praw majątkowych do artykułu.

Zgłoszony artykuł jest oceniany i opracowywany w trzyetapowym 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 artykułu, 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, do której afiliowana jest zgłoszona praca; w przypadku artykułu 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 dostarczają redakcji zmodyfikowaną wersję pracy. Jeśli pojawi się różnica zdań dotycząca zasadności proponowanych zmian, autorzy są zobligowani do uzasadnienia swojego stanowiska.

3. **Ocena dopuszczająca do publikacji**, dokonywana przez Kolegium Redakcyjne (KR) 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. Kolegium Redakcyjne ocenia artykuł pod względem poprawności i spójności merytorycznej oraz zaleca autorowi wprowadzenie poprawek, jeśli są one konieczne, aby praca spełniała wymogi czasopisma. W przypadku podjęcia decyzji o niepublikowaniu artykułu autorowi przysługuje prawo do odwołania. W tym celu 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. W przypadku nieuwzględnienia danej uwagi autor jest zobligowany do uzasadnienia swojego stanowiska.**

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” dokłada wszelkich starań, aby utrzymać najwyższe standardy etyczne, zgodnie z wytycznymi Komitetu ds. Etyki Publikacyjnej (COPE), dostępnymi na stronie internetowej [www.publicationethics.org](http://www.publicationethics.org), oraz wykorzystuje wszystkie możliwe środki mające na celu zapobieżenie nadużyciom i nierzetelności autorskiej. Przyjęte zasady postępowania obowiązują autorów, redakcję, 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 zastosowanej w nich metodologii. W przypadku nowatorskich metod analizy pożądane 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 opracowaniu 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

symposium naukowe, 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 publikacji.
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 poświadczeniem na piśmie 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 o tym niezwłocznie poinformować redakcję w celu dokonania korekty, wycofania tekstu lub zamieszczenia odpowiedniego sprostowania. W przypadku korekty artykułu już opublikowanego jego nowa wersja jest zamieszczana na stronie internetowej „WS” wraz ze stosownym wyjaśnieniem.

### **3.2. Odpowiedzialność redakcji**

1. Redakcja „WS” odpowiada za zorganizowanie i sprawny przebieg procesu wydawniczego, na który składają się: wstępna ocena zgłoszonego maszynopisu, ocena recenzentów (w przypadku artykułów naukowych), ocena KR, redakcja językowa, redakcja techniczna, skład i łamanie oraz korekta.
2. Redakcja ustala zasady obowiązujące w procesie wydawniczym, informuje jego uczestników o konieczności ich przestrzegania i egzekwuje je na każdym z jego etapów oraz dba o stałą aktualizację informacji na temat przyjętych zasad na stronie internetowej i na łamach czasopisma.
3. Redakcja nie może pozostawać w jakimkolwiek konflikcie interesów w odniesieniu do przyjmowanych artykułów. Przez konflikt interesów należy rozumieć sytuację, w której wszelkie interesy lub związki (służbowe, finansowe lub inne) mogą mieć wpływ na obiektywną ocenę zgłoszonego maszynopisu lub decyzję o jego publikacji.
4. W celu przeciwdziałania nierzetelności naukowej redakcja wymaga od autorów złożenia oświadczenia, w którym deklarują oni, że zgłaszany artykuł nie narusza praw autorskich osób trzecich, nie był dotychczas publikowany i że jest ich oryginalnym dziełem, oraz określają swój wkład w opracowanie artykułu.
5. Redaktorzy weryfikują zgłoszony maszynopis pod względem zgodności z celem i zakresem tematycznym czasopisma oraz spełniania wymogów redakcyjnych „WS”, a także ewentualnych przejawów nierzetelności naukowej i możliwości wystąpienia konfliktu interesów. Obiektywną ocenę oryginalności tekstu zapewnia system antyplagiatowy stosowany przez redakcję.

6. Redakcja jest odpowiedzialna za ustalenie spójnych kryteriów oceny artykułu oraz wybór niezależnych recenzentów, którzy są zobligowani do złożenia oświadczenia o przestrzeganiu zasad etyki recenzowania COPE ([publicationethics.org/resources/guidelines-new/cope-ethical-guidelines-peer-reviewers](http://publicationethics.org/resources/guidelines-new/cope-ethical-guidelines-peer-reviewers)) i niewystępowaniu konfliktu interesów. Informacje dotyczące maszynopisu mogą być przekazywane przez redakcję wyłącznie autorem, recenzentem, wydawcy lub innym doradcom redakcyjnym.
7. W przypadku podejrzenia nadużyć redakcja postępuje zgodnie z procedurami COPE.
8. Redakcja zapewnia, że zmiany dokonane w tekście na etapie prac redakcyjnych nie naruszają zasadniczej myśli autorów.
9. Kolegium Redakcyjne, podejmując decyzję o publikacji artykułu, kieruje się wyłącznie wynikiem dyskusji dotyczącej zgłoszonego artykułu, w której uwzględniane są oceny recenzentów oraz opinie redaktorów tematycznego i merytorycznego. Rezultat ten zależy od merytorycznej oceny wartości artykułu, jego oryginalności i jasności przekazu, a także od ścisłego związku z celem i zakresem tematycznym miesięcznika.
10. W przypadku podjęcia decyzji o niepublikowaniu przesłanego materiału redakcja nie może go w żaden sposób wykorzystać bez pisemnej zgody autora.

### 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. W uzasadnionych przypadkach 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 jakiejkolwiek 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 na zasadach 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.

**Zachęcamy do przygotowania pracy z wykorzystaniem szablonu artykułu „WS” – do pobrania ze strony [ws.stat.gov.pl/ForAuthors](http://ws.stat.gov.pl/ForAuthors).**

### 4.1. Struktura i zawartość artykułu

Wymagane elementy artykułu:

1. Tytuł, autor.
2. Streszczenie (objętość do 1200 znaków ze spacjami, forma bezosobowa). W przypadku artykułu opisującego badanie empiryczne powinno zawierać: cel, przedmiot, okres i metodę badania, źródła danych, najważniejsze wnioski z badania. W przypadku artykułów o innym charakterze należy podać co najmniej cel artykułu, 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.**

3. Słowa kluczowe – najistotniejsze użyte w pracy pojęcia lub wyrażenia (nie mniej niż trzy). Słowa kluczowe powinny być zawarte w streszczeniu i/lub tytule.
4. Kod/kody z klasyfikacji Journal of Economic Literature (JEL).
5. 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).
6. W przypadku artykułu opisującego badanie empiryczne wymagane są następujące części:
  - wprowadzenie, zawierające: cel badania, uzasadnienie podjętego problemu badawczego i odniesienie do literatury przedmiotu, chyba że przegląd literatury stanowi odrębną część artykułu;
  - metoda badania, zawierająca: przedmiot i okres badania, źródła danych i zastosowane metody badawcze;
  - wyniki badania;
  - podsumowanie, które powinno być zwięzłe i odzwierciedlać istotę problemu badawczego przedstawionego w artykule, bez podawania danych liczbowych; wnioski powinny odnosić się do treści artykułu, a w szczególności do celu badania.Wszystkie części powinny być opatrzone numerami.
7. Bibliografia, zawierająca pełny wykaz prac i materiałów przywołanych w artykule, przygotowana zgodnie z wymogami czasopisma.

### 4.2. Przygotowanie artykułu

1. Tekst należy zapisać alfabetem łacińskim. Nazwy własne, tytuły itp. oryginalnie zapisane innym alfabetem powinny być poddane transliteracji.
2. Nie należy stosować stylów; formatowanie należy ograniczyć do wymogów redakcyjnych.



3. Objętość artykułu łącznie ze streszczeniem, słowami kluczowymi, bibliografią, tablicami, wykresami i innymi materiałami graficznymi nie powinna być mniejsza niż 10 stron ani przekraczać 20 stron maszynopisu.
4. Edytor tekstu: Microsoft Word, format *.doc lub *.docx.
5. Krój czcionki:
  - Arial – tytuł, autor, streszczenia, słowa kluczowe, kody JEL, tablice, zestawienia, wykresy, przypisy, śródtytuły;
  - Times New Roman – tekst główny, bibliografia.
6. Wielkość czcionki:
  - 14 pkt – tytuł, autor, tytuły rozdziałów;
  - 12 pkt – tekst główny, tytuły podrozdziałów;
  - 10 pkt – pozostałe elementy.
7. Marginesy – 2,5 cm z każdej strony.
8. Interlinia – 1,5 wiersza; tablice i przypisy – 1 wiersz; przed tytułami rozdziałów i podrozdziałów oraz po nich – pusty wiersz.
9. Wcięcie akapitowe – 0,4 cm; bibliografia – bez wcięcia, wysunięcie 0,4 cm.
10. Przy wycieniach należy posłużyć się listą punktowaną z punktarami w postaci kropek (wysunięcie 0,4 cm, wcięcie 0 cm); wiersze (oprócz ostatniego) zakończone średnikiem.
11. Strony ponumerowane automatycznie.
12. Tablice i elementy graficzne (wykresy, mapy, schematy) muszą być przywołane w tekście.
13. 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.
14. Tablice muszą być edytowalne. Nie należy stosować rastrów, cieniowania, pogrubiania czy też podwójnych linii itp.
15. Wskazówki dotyczące opracowywania map znajdują się w publikacji *Mapy statystyczne. Opracowanie i prezentacja danych*, dostępnej na stronie internetowej GUS [stat.gov.pl/statystyka-regionalna/publikacje-regionalne/podreczniki-atlasy/podreczniki/mapy-statystyczne-opracowanie-i-prezentacja-danych,1,1.html](http://stat.gov.pl/statystyka-regionalna/publikacje-regionalne/podreczniki-atlasy/podreczniki/mapy-statystyczne-opracowanie-i-prezentacja-danych,1,1.html).
16. Pod tablicami, wykresami, schematami i innymi elementami graficznymi należy podać źródło opracowania.
17. Oznaczenia literowe należy zapisywać następująco: liczby i inne wielkości niezłożone – małe lub duże litery, kursywa, bez pogrubienia (np.  $a$ ,  $A$ ,  $y(x)$ ,  $a_i$ ); wektory – małe litery, kursywa, pogrubione (np.  $\mathbf{a}$ ,  $\mathbf{w}$ ,  $\mathbf{y}(x)$ ,  $\mathbf{w}_i$ ); macierze – duże litery, proste, pogrubione (np.  $\mathbf{A}$ ,  $\mathbf{M}$ ,  $\mathbf{Y}(x)$ ,  $\mathbf{M}_i$ ).
18. 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.

19. Stosowane są skróty: tablica – tabl., wykres – wykr.
20. Przypisy rzeczowe, słownikowe lub informacyjne należy umieszczać na dole strony. Przypisy bibliograficzne, zgodnie ze standardem APA (American Psychological Association), należy podawać w tekście głównym.
21. Bibliografię należy przygotować zgodnie ze standardem APA.

### 4.3. Zasady przywoływania publikacji w treści artykułu

1. Jeden autor: bez względu na to, ile razy przywoływana jest praca, zawsze należy podać nazwisko autora i datę publikacji pracy, a w przypadku więcej niż jednej pracy danego autora opublikowanej w tym samym roku należy dodać kolejne litery alfabetu przy dacie (np. 2001a). Przykład zapisu: Jak stwierdza Iksiński (2001)... Badania wskazują, że... (Iksiński, 2001).
2. Dwoch autorów: bez względu na to, ile razy przywoływana jest praca, zawsze należy podać nazwiska obu autorów i datę publikacji pracy, a w przypadku więcej niż jednej pracy tych autorów opublikowanej w tym samym roku należy dodać kolejne litery alfabetu przy dacie. Nazwiska autorów zawsze należy łączyć spójnikiem „i”, nawet w przypadku przywoływania publikacji obcojęzycznej. Przykład zapisu: Jak sugerują Iksiński i Nowak (1999)... Badania wskazują, że... (Iksiński i Nowak, 1999).
3. Od trzech do pięciu autorów: przywołanie po raz pierwszy – należy wymienić nazwiska wszystkich autorów, rozdzielając je przecinkami i stawiając spójnik „i” pomiędzy dwoma ostatnimi nazwiskami. Przy kolejnych wskazaniach tej samej pracy należy zastosować określenie „i współpracownicy” (w przypadku umieszczenia przywołania nazwisk w strukturze zdania) lub „i in.” (gdy nazwiska autorów nie stanowią części struktury zdania). Przykład zapisu z przywołaniem po raz pierwszy: Jak sugerują Nowak, Iksiński i Jankiewicz (2003)... Badania (Nowak, Iksiński i Jankiewicz, 2003) wskazują, że... Przykład zapisu z kolejnymi przywołaniami: Badania Nowaka i współpracowników (2003)... Badania te wskazują, że... (Nowak i in., 2003).
4. Sześciu i więcej autorów: należy wymienić tylko nazwisko pierwszego autora, zarówno gdy praca przywoływana jest po raz pierwszy, jak i w późniejszych przywołaniach, natomiast pozostałych autorów należy zastąpić określeniem „i współpracownicy” (w przypadku umieszczenia przywołania nazwisk w strukturze zdania) lub „i in.” (gdy nazwiska autorów nie stanowią części struktury zdania). W literaturze załącznikowej należy umieścić nazwiska wszystkich autorów pracy. Przykład zapisu: Nowakowski i współpracownicy (1997) twierdzą, że... Pierwsze badania na ten temat sugerują... (Nowakowski i in., 1997).
5. Przywoływanie jednocześnie kilku prac: należy wymienić je alfabetycznie według nazwiska pierwszego autora. Przywołania kolejnych prac muszą być oddzielone średnikami. Lata wydania prac tego samego autora / tych samych autorów muszą być oddzielone przecinkiem. Przykład zapisu: Iksiński (2001); Nowak i Iksiński (1999, 2005); (Iksiński, 1997, 1999, 2004a, 2004b; Nowak i Iksiński, 1999).
6. Przywoływanie pracy za innym autorem: stosuje się w tekście, natomiast w bibliografii należy umieścić jedynie pracę czytaną. Przykład zapisu: Jak wykazał Nowakowski (1990; za: Zieniecka, 2007)... Badania sugerują, że... (Nowakowski, 1990; za: Zieniecka, 2007).

7. Bibliografia powinna być zamieszczona na końcu opracowania. Prace należy zapisać alfabetycznie według nazwiska pierwszego autora. W przypadku dwóch lub więcej prac tego samego autora / tych samych autorów należy je uporządkować według roku publikacji. Jeśli kilka prac tego samego autora / tych samych autorów zostało opublikowanych w tym samym roku, należy uporządkować prace alfabetycznie według tytułu i wstawić litery a, b, c itd. po roku publikacji.

#### 4.4. Przykłady opisu bibliograficznego

1. Artykuł w czasopiśmie, w którym każdy kolejny numer/zeszyt (issue) w ramach jednego rocznika ma osobną numerację stron (w każdym zeszycie pierwsza strona opatrzona jest numerem 1): Nazwisko, X., Nazwisko 2, X. Y., Nazwisko 3, Z. (rok). Tytuł artykułu. *Tytuł czasopisma, rocznik(zeszyt)*, strona początku–strona końca.
2. Artykuł w czasopiśmie, w którym kolejne numery/zeszyty (issues) w ramach jednego rocznika nie mają osobnej numeracji stron (pierwsza strona w kolejnym zeszycie opatrzona jest numerem kolejnym po ostatniej stronie w zeszycie poprzednim): Nazwisko, X., Nazwisko 2, X. Y., Nazwisko 3, Z. (rok). Tytuł artykułu. *Tytuł czasopisma, rocznik*, strona początku–strona końca. Jeśli artykuł ma numer DOI (Digital Object Identifier), należy podać go na końcu opisu bibliograficznego: Nazwisko, X., Nazwisko 2, X. Y. (rok). Tytuł artykułu. *Tytuł czasopisma, rocznik*, strona początku–strona końca. DOI: xxxxx.
3. Książka: Nazwisko, X., Nazwisko 2, X. Y. (rok). *Tytuł książki*. Miejsce wydania: Wydawnictwo.
4. Książka napisana pod redakcją: Nazwisko, X. (red.). (rok). *Tytuł książki*. Miejsce wydania: Wydawnictwo.
5. Rozdział w pracy zbiorowej: Nazwisko, X. (rok). Tytuł rozdziału. W: Y. Nazwisko, B. Nazwisko 2 (red.), *Tytuł książki* (s. strona początku–strona końca). Miejsce wydania: Wydawnictwo.
6. Jeśli dany tekst znajduje się na stronie internetowej i nie jest artykułem w czasopiśmie, książką ani rozdziałem w książce, należy podać autora, datę publikacji (jeśli jest znana), tytuł, a następnie zamieścić informacje o stronie, z której został pobrany, oraz – jeśli są to materiały informacyjne – datę dostępu. Tekst: Nazwisko, X. (rok). *Tytuł tekstu*. Pobrane z: adres strony internetowej (dostęp: 21.03.2019).

**Artykuł przygotowany w sposób niezgodny z powyższymi wskazówkami będzie odesłany z prośbą o dostosowanie jego formy do wymagań redakcji.**

## ZAKRES TEMATYCZNY DZIAŁÓW

### THEMATIC SCOPE OF SECTIONS

(for the English translation of the information given below, please visit [ws.stat.gov.pl/AimScope](http://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; zamieszczone 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.