Cena 13,00 zł
(VAT 8%)

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

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

PAŹDZIERNIK / OCTOBER
ROK / VOLUME 67

2022 | 10

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

POLSKIE TOWARZYSTWO STATYSTYCZNE
POLISH STATISTICAL ASSOCIATION

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

PAŹDZIERNIK / OCTOBER
ROK / VOLUME 67

2022 | 10 (737)

RADA NAUKOWA / SCIENTIFIC COUNCIL

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

sekretarz/secretary: Paulina Kucharska-Singh, Główny Urząd Statystyczny, Polska

KOLEGIUM REDAKCYJNE / EDITORIAL BOARD

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

ZESPÓŁ REDAKCYJNY / EDITORIAL STAFF

redaktor naczelny / editor-in-chief: Marek Cierpiał-Wolan
zastępca redaktora naczelnego / deputy editor-in-chief: Andrzej Młodak
redaktorzy tematyczni / thematic editors: Małgorzata Tarczyńska-Łuniewska, Agnieszka Zgierska
redaktor merytoryczny / substantive editor: Wioletta Wrzaszcz

sekretarz/secretary: Małgorzata Zygmont, Główny Urząd Statystyczny, Polska

ADRES REDAKCJI / EDITORIAL OFFICE ADDRESS

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

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

Redakcja techniczna, skład i łamanie, opracowanie materiałów graficznych, korekta: Zakład Wydawnictw Statystycznych – zespół pod kierunkiem Wojciecha Szuchty
Technical editing, typesetting, preparation of graphic materials, proofreading: Statistical Publishing Establishment – team supervised by Wojciech Szuchta

Zakład Wydawnictw Statystycznych

Druk i oprawa / Printed and bound by:
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 primary version of the journal, issued in electronic form, is available at ws.stat.gov.pl

© Copyright by Główny Urząd Statystyczny and the authors, some rights reserved. CC BY-SA 4.0 licence

Indeks 381306

Informacje w sprawie sprzedaży czasopisma / Sales of the journal:
Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment
tel./phone +48 22 608 32 10, +48 22 608 38 10
Prenumerata jest prowadzona przez / Subscription is available at RUCH S.A.
Zamówienia na prenumeratę można składać na stronie / Subscriptions can be ordered at
www.prenumerata.ruch.com.pl

Zam. 253/2022 – nakład 285 egz.

SPIS TREŚCI
CONTENTS

OD REDAKCJI

Na łamach październikowego numeru „Wiadomości Statystycznych. The Polish Statistician" publikujemy studium metodologiczne oraz dwie prace przedstawiające wyniki badań przeprowadzonych z zastosowaniem metod statystycznych.

Studium *Estimation risk taking into consideration the effect of forecasting scheme: robust inference about VaR* dr Marty Małeckiej dotyczy problemu ryzyka estymacyjnego przy testowaniu modeli regresyjnych opisujących wartość zagrożoną ryzykiem (Value-at-Risk – VaR). Chodzi o możliwość niespełnienia przez proces w analizie empirycznej standardowych wymogów określających ramy testowe, co skutkuje wysokim prawdopodobieństwem odrzucania przez testy VaR prawidłowych modeli tej klasy ze względu na błędy estymacji popełnione podczas wyznaczania prognoz VaR. Autorka ocenia odporność testów VaR na to ryzyko. Stawia sobie przy tym za cel, po pierwsze, znalezienie metod, które pozwalałyby zniwelować negatywny wpływ ryzyka estymacji na efektywność testów VaR, a po drugie – kompleksowe porównanie metod testowania VaR w odniesieniu do problemu ryzyka estymacyjnego. Przeprowadzone analizy wskazują m.in. na to, że w bardziej realistycznych schematach prognostycznych utrata dokładności spowodowana zastosowaniem rozkładów wynikających z podpróbkowania przewyższa utratę dokładności wynikającą z wystąpienia ryzyka estymacji. Zaproponowane przez autorkę metody są znacznie mniej czasochłonne niż inne i nie wymagają projektowania symulacji w celu przeprowadzenia testu VaR. Okazują się więc optymalne z punktu widzenia realnych działań gospodarczych.

W artykule *Ocena stopnia zagrożenia ubóstwem w ramach podejścia możliwości Sena z zastosowaniem modelu MIMIC* mgr Beata Kraszewska przeprowadza wielowymiarową analizę zagrożenia ubóstwem w ramach podejścia możliwości (ang. *capability approach*) opracowanego przez Amartyę Sena. Autorka określa też determinanty ubóstwa, które wzmacniają lub osłabiają zagrożenie ubóstwem. Do oceny stopnia zagrożenia ubóstwem w ujęciu wielowymiarowym wykorzystuje dane dotyczące Polski (m.in. w układzie regionalnym) uzyskane z Europejskiego Badania Warunków Życia Ludności (EU-SILC) przeprowadzonego przez GUS w 2018 r. Jako narzędzia teoretycznego używa szczególnego wariantu modelu równań strukturalnych (Structural Equations Modeling – SEM), a mianowicie modelu wielu wskaźników i wielu przyczyn (Multiple Indicators and Multiple Causes – MIMIC). Na podstawie analizy, w której posługuje się wskaźnikiem stopnia zagrożenia ubóstwem (Risk of Poverty Index – RPI), autorka stwierdza, że na zmniejszenie stopnia zagrożenia ubóstwem najsilniej wpływa wykształcenie, natomiast na jego zwiększenie – niski status aktywności na rynku pracy (tzn. bycie rencistą lub osobą bezrobotną). Ponadto opisany w pracy przykład empiryczny potwierdza, że ujęcie wielowymiarowe pozwala precyzyjniej niż jednowymiarowe zidentyfikować osoby ubogie i ocenić zagrożenie ubóstwem.

Dr Elżbieta Zalewska i dr Kamila Trzcińska w artykule *Efektywność zajęć zdalnych w czasie pandemii COVID-19* porównują efektywność kształcenia akademickiego w trybie stacjonarnym i zdalnym. Do przeprowadzenia analiz używają danych zaczerpniętych z bazy wyników uzyskanych przez studentów I roku studiów dziennych na kierunku finanse i rachunkowość, prowadzonych na Wydziale Ekonomiczno-Socjologicznym Uniwersytetu Łódzkiego, z egzaminów końcowych z matematyki w latach akademickich 2018/2019 (nauka stacjonarna) i 2020/2021 (nauka zdalna) oraz – jako punktu odniesienia – z egzaminów maturalnych z matematyki na poziomie podstawowym zdanych przez badane osoby. Do oceny efektywności kształcenia wykorzystano krzywą Lorentza i współczynnik Giniego. Wyniki analiz pokazują, że w badanej zbiorowości nie występują

© Główny Urząd Statystyczny
Artykuł udostępniony na licencji CC BY-SA 4.0 / Article available under the CC BY-SA 4.0 licence

istotne różnice pomiędzy efektami nauki zdalnej i stacjonarnej. Można dostrzec nieznaczną przewagę pozytywnych wyników egzaminów końcowych w okresie nauki zdalnej w porównaniu z odpowiednimi rezultatami uzyskanymi podczas nauki stacjonarnej.

W tym numerze wspominamy zmarłego w zeszłym roku prof. dr. hab. Krzysztofa Zadorę, wybitnego śląskiego statystyka i ekonometryka, znakomitego specjalistę z zakresu metod ilościowych i modeli ekonometrycznych, rektora i wieloletniego wykładowcę Akademii Ekonomicznej w Katowicach. Sylwetkę Profesora nakreślił prof. dr hab. Janusz L. Wywiał.

W naszej stałej rubryce „Wydawnictwa GUS" Joanna Sadowy przedstawia listę publikacji GUS wydanych we wrześniu tego roku. Jedną z książek, której poświęca więcej uwagi, jest monografia *Nowoczesne technologie i nowe źródła danych w pomiarze inflacji* pod redakcją naukową Jacka Białka, Mieczysława Kłopotka i Tomasza Panka – t. 70 serii Biblioteka Wiadomości Statystycznych. W opracowaniu zaprezentowano wyniki prac zrealizowanych przez konsorcjum naukowe (składające się z GUS, Szkoły Głównej Handlowej w Warszawie i Instytutu Podstaw Informatyki PAN) w ramach projektu „Budowa zintegrowanego systemu statystyki cen detalicznych – INSTATCENY". Prace te dotyczyły uzyskiwania i przetwarzania danych pochodzących z nowych, alternatywnych źródeł, takich jak dane skanowane, dostarczane przez sieci handlowe, oraz dane skrapowane, pobierane automatycznie ze stron internetowych sklepów.

Życzymy miłej lektury.

FROM THE EDITORIAL TEAM

The October issue of *Wiadomości Statystyczne. The Polish Statistician* comprises a methodological study and two papers presenting the results of research carried out by means of statistical methods.

The study *Estimation risk taking into consideration the effect of forecasting scheme: robust inference about VaR*, by Marta Małecka, PhD, researches the problem of the estimation risk in testing regression Value-at-Risk (VaR) models. More specifically, it refers to the possibility of the VaR violation process not fulfilling the standard requirements defining the testing framework. Such a situation results in a high probability of the VaR tests rejecting correct models of this type due to estimation errors made while assessing VaR. The author examines the robustness of VaR tests to the estimation risk. Her aim is, firstly, to find methods which would make it possible to compensate for the negative influence of the estimation risk on the effectiveness of the VaR tests, and secondly, to comprehensively compare the methods of VaR testing with the estimation risk problem in mind. The analyses performed by the author demonstrate that, for example, in more realistic forecasting schemes, the loss of accuracy caused by the application of the subsampling technique is higher than the loss of accuracy caused by the occurrence of the estimation risk. The methods proposed by the author are much less time-consuming than others, and they do not require simulations in order to perform a VaR test. They therefore appear to be optimal from the point of view of real economic activity.

Beata Kraszewska, MSc, in her paper *Poverty risk estimation on the basis of Sen's capability approach with the application of the MIMIC model* conducts a multidimensional analysis of the poverty threat in the framework of Amartya Sen's capability approach. The author also defines the determinants of poverty which increase or decrease the risk of its occurrence. In order to assess the degree of poverty threat in the framework of multidimensional approach, she uses data concerning

Poland (including regional data) obtained from the European Survey on Living Conditions (EU-SILC) carried out by Statistics Poland in 2018. The author applies a special variant of structural equations modelling (SEM), i.e. multiple indicators and multiple causes model (MIMIC), as the theoretical tool. On the basis of the analysis where the Risk of Poverty Index (RPI) has been used, the author comes to the conclusion that education has the greatest impact on reducing the risk of poverty, while low activity on the job market, i.e. being a pensioner or unemployed, increases the risk of poverty to the largest extent. In addition, the empirical example described in the paper demonstrates that the multidimensional approach allows a more precise identification of poor people and evaluation of the poverty risk than the unidimensional approach.

In their article entitled *Effectiveness of distance learning during the COVID-19 pandemic*, Elżbieta Zalewska, PhD, and Kamila Trzcińska, PhD, compare the effectiveness of academic education in the classroom-learning and distance-learning modes. The authors use data from the database of the results achieved by first-year full-time students majoring in finance and accounting at the Faculty of Economics and Sociology at the University of Lodz in their final exams in mathematics in the years 2018/2019 (classroom learning) and 2020/2021 (distance learning). The scores achieved by these students in the basic-level matura exam in mathematics served as a point of reference. The Lorenz curve and the Gini coefficient were used to assess the effectiveness of learning. The results of the analyses demonstrated there were no significant differences in terms of the effects of distance learning and classroom learning in the given population except for a slight predominance of positive results of final exams during distance learning compared to classroom learning.

In this issue, we are also remembering Krzysztof Zadora, PhD, DSc, ProfTit, who passed away last year. Professor Zadora was a distinguished statistician and econometrician as well as an outstanding specialist in the fields of quantitative methods and econometric models. He served as a Dean and a long-time academic teacher at the University of Economics in Katowice. The life and achievements of Professor Zadora are presented by Janusz L. Wyrwiał, PhD, DSc, ProfTit.

In the *Publications of Statistic Poland* section, Joanna Sadowy presents a list of Statistics Poland's publications issued in September. One of the books in her focus is *Modern technologies and new sources of data in inflation measurement* monograph edited by Jacek Białek, Mieczysław Kłopotek and Tomasz Panek, published as Volume 70 of the BWS series. This study features the results of the joint work performed by a scientific consortium consisting of Statistics Poland, SGH Warsaw School of Economics and the Institute of Computer Science of the Polish Academy of Sciences in the framework of the 'Construction of integrated system of retail prices' (Pol. Budowa zintegrowanego systemu statystyki cen detalicznych – INSTATCENY). Their work related to obtaining and processing data from new, alternative sources, such as scanner data provided by retail chains, and scraped data downloaded automatically from shop websites.

We wish you pleasant reading.

Wiadomości Statystyczne. The Polish Statistician, 2022, vol. 67, 10, VII–VIII
Zaproszenie do nadsyłania artykułów / Call for papers

Zaproszenie do nadsyłania artykułów do działu Spisy powszechne – problemy i wyzwania

Szanowni Państwo,
z okazji przypadających na początek dekady 2020 spisów powszechnych w „Wiadomościach Statystycznych. The Polish Statistician" został utworzony dział Spisy powszechne – problemy i wyzwania. Powstał on jako miejsce twórczej wymiany doświadczeń statystyków i badaczy podejmujących tę tematykę, jedną z najważniejszych w statystyce publicznej. W związku z tym mamy przyjemność zaprosić Państwa do nadsyłania artykułów naukowych, które będą przedstawiać propozycje efektywnych rozwiązań dotyczących spisów – zarówno organizacyjnych, jak i metodologicznych – a także zawierać analizę danych statystycznych uzyskanych w wyniku spisów, prowadzoną z zastosowaniem nowoczesnych narzędzi teoretycznych i informatycznych. Mile widziane są opracowania skupiające się na praktycznych aspektach związanych z przeprowadzaniem spisów, w tym dotyczących obciążenia odpowiedzi i ochrony tajemnicy statystycznej.

Ostatnie spisy powszechne stały się szczególnym wyzwaniem ze względu na niespodziewane problemy organizacyjno-metodologiczne, w ogromnej mierze wynikające z wybuchu pandemii COVID-19 oraz rosnących niepokojów społecznych i gospodarczych obserwowanych na całym świecie. W rezultacie spisy nie tylko dostarczyły konkretnych danych statystycznych, lecz także ujawniły potrzebę zastosowania nadzwyczajnych środków organizacyjnych i technicznych podczas zbierania i przetwarzania danych, co powinno być uwzględnione w ocenie jakości wyników. Paradoksalnie zdarzenia losowe, które wymagają opracowania alternatywnych scenariuszy działań i przygotowania na ryzyko w dalszych badaniach statystycznych, mogą się przyczynić do postępu w ich realizacji. Wymienione wyżej zagadnienia wymagają kompleksowych analiz, które za pośrednictwem „Wiadomości Statystycznych. The Polish Statistician" mają szansę trafić do szerokiego grona odbiorców.

Zachęcamy do zapoznania się z informacjami na temat czasopisma i wskazówkami dla autorów, dostępnymi na stronie ws.stat.gov.pl. Artykuły w języku polskim lub angielskim należy przesyłać na adres: redakcja.ws@stat.gov.pl.

dr Dominik Rozkrut
Prezes Głównego Urzędu Statystycznego
Przewodniczący Rady Naukowej
„Wiadomości Statystycznych.
The Polish Statistician"

dr Marek Cierpiał-Wolan
Redaktor Naczelny
„Wiadomości Statystycznych.
The Polish Statistician"

Issues and challenges in census taking – call for papers

Dear Readers,

In response to the censuses carried out at the beginning of the current decade, a new section entitled *Issues and challenges in census taking* has been introduced to the *Wiadomości Statystyczne. The Polish Statistician* journal. The section is meant as a platform for the creative exchange of experiences among statisticians and researchers dealing with national censuses, which are one of the most important undertakings of official statistics. In this connection, we have the pleasure of inviting you to send in scientific papers proposing effective census-related solutions, both organisational and methodological, as well as works featuring analyses of data obtained through censuses, performed with the use of innovative theoretical and IT tools. We also welcome papers focusing on the practical aspects related to carrying out censuses, including response burden and the protection of statistical confidentiality.

Censuses have recently become especially challenging – this mainly being the effect of unforeseen organisational and methodological problems related to the outbreak of the COVID-19 pandemic and intensifying social and economic instabilities observed throughout the globe. The most recent censuses have therefore not only provided concrete statistical data, but also necessitated the application of extraordinary organisational and technical measures in collecting and processing data, which should be taken into account while assessing the quality of the results. Paradoxically, random occurrences which require devising alternative procedures and preparation for risk in the further research activity are likely to contribute to the progress of research. All the above-mentioned issues, as we can see, require comprehensive analyses, which now, with the help of *Wiadomości Statystyczne. The Polish Statistician*, stand a chance of reaching large audiences.

We would like to invite you to learn more about our journal and the guidelines for authors, which are available on our website: ws.stat.gov.pl. Articles written in Polish or English should be sent to: redakcja.ws@stat.gov.pl.

Dr Dominik Rozkrut

President of Statistics Poland
Chairman of *Wiadomości Statystyczne. The Polish Statistician* Scientific Council

Dr Marek Cierpiał-Wolan

Editor-in-Chief
of Wiadomości Statystyczne. The Polish Statistician

Wiadomości Statystyczne. The Polish Statistician, 2022, vol. 67, 10, 1–27 DOI: 10.5604/01.3001.0016.0657
Studia metodologiczne / Methodological studies
Zgłoszony/submitted: 20.12.2021, poprawiony/revised: 26.09.2022, zaakceptowany/accepted: 4.10.2022

Estimation risk taking into consideration the effect of forecasting scheme: robust inference about VaR[1]

Marta Małecka[a]

Abstract. The paper addresses the issue of estimation risk in VaR testing. The occurrence of estimation risk (also called parameter uncertainty) implies that the observed VaR violation process may not fulfil the standard requirements that underpin the testing framework. As a result, VaR tests may reject correct VaR models due to estimation errors committed when predicting the VaR. The paper examines the robustness of VaR tests to estimation risk. The research is based on an observation indicating that certain elements of a forecasting scheme have a significant influence on estimation risk. Thus, the article extends the previous studies to include several more realistic forecasting schemes than those based solely on a fixed window.

The aim of the research is twofold: firstly, to find methods of mitigating the negative impact of estimation risk on VaR tests, and secondly, to provide a comprehensive comparison of VaR testing methods with reference to the issue of estimation risk. The conducted analyses demonstrate that a proper adjustment of the forecasting scheme yields better results in terms of the accuracy of the tests than correcting estimation errors by means of the subsampling technique.

Keywords: VaR tests, estimation risk, parameter uncertainty
JEL: C12, C52, C53, G17

Ryzyko estymacyjne uwzględniające schemat prognozowania – wnioskowanie o VaR za pomocą metod odpornych

Streszczenie. Artykuł dotyczy problemu ryzyka estymacyjnego przy testowaniu VaR. Występowanie ryzyka estymacyjnego (zwanego również niepewnością parametrów) oznacza, że obserwowany proces przekroczeń VaR może nie spełniać standardowych wymogów określających ramy testowe. W konsekwencji testy VaR mogą odrzucać prawidłowe modele VaR ze względu na błędy estymacji popełnione podczas wyznaczania prognoz VaR. W badaniu omawianym w artykule oceniana jest odporność testów VaR na ryzyko estymacyjne. U podstaw badania leży spostrzeżenie, że ryzyko estymacyjne w istotny sposób zależy od elementów schematu prognozowania. Z tego powodu w badaniu uwzględniono schematy prognozowania bardziej realistyczne niż schemat oparty na ustalonym oknie, co stanowi rozszerzenie w stosunku do wcześniej prowadzonych badań.

Cel badania jest dwojaki: znalezienie metod, które pozwalałyby zniwelować negatywny wpływ ryzyka estymacji na testy VaR, oraz kompleksowe porównanie metod testowania VaR

[1] Artykuł został opracowany na podstawie referatu wygłoszonego na konferencji Multivariate Statistical Analysis MSA 2021, która odbyła się w dniach 8–10 listopada 2021 r. w Łodzi. / The article is based on a paper delivered at the Multivariate Statistical Analysis MSA 2021 conference, held on 8–10 November 2021 in Łódź.

[a] Uniwersytet Łódzki, Katedra Metod Statystycznych, Polska / University of Lodz, Department of Statistical Methods, Poland. ORCID: https://orcid.org/0000-0003-4465-9811. E-mail: marta.malecka@uni.lodz.pl.

© Główny Urząd Statystyczny

w odniesieniu do problemu ryzyka estymacyjnego. Przeprowadzone analizy wskazują m.in. na to, że odpowiednie dostosowanie schematu prognozowania daje lepsze wyniki pod względem dokładności testów niż korygowanie błędów estymacji techniką podpróbkowania.
Słowa kluczowe: testy VaR, ryzyko estymacyjne, niepewność parametrów

1. Introduction

According to the Basel III and Basel IV accords,[2] the current global banking regulations recommend including VaR (Value-at-Risk) tests in banks' internal risk management systems. As a result, the VaR testing framework continues to be an important issue discussed in the financial and statistical literature. Recent studies on this subject have revealed a new methodological problem – the inaccuracy of testing methods due to estimation risk. For example, Escanciano and Olmo (2010) found that estimation risk changes the size of VaR tests. The size, together with the power, are two fundamental properties of a statistical test. They ensure that proper models are not overrejected (i.e. rejected more often than indicated by the chosen significance) and incorrect models are efficiently detected.

The influence of estimation risk on the size of VaR tests results from distortions of the VaR violation process, underpinning the backtesting procedures. These distortions occur because the violation process is based on the estimated, not the actual VaR. On the other hand, test statistic distributions typically rely on the violation process based on the actual VaR. As a consequence of this inconsistency, the VaR models may be rejected not only due to their incorrectness, but also as a result of estimation errors inherent in these models. Intuitively, the scale of estimation risk heavily depends on the forecasting scheme applied to produce VaR forecasts. The components of these forecasting schemes, such as the window choice or frequency of parameter correction, may potentially reduce the estimation errors and their negative impact on subsequent statistical inference. This issue, however, has not yet been studied. The existing proposals as to how to deal with the problem of size distortions caused by estimation risk (Escanciano & Olmo, 2010, 2011) do not investigate forecasting schemes. Instead, they are based on a fixed forecasting scheme and suggest solving this problem by using resampled distributions. According to these proposals, resampled distributions should replace asymptotic distributions. They should be obtained in a way that takes into account estimation risk. Our approach to this problem does not resort to resampling, but focuses on forecasting schemes instead.

In order to overcome the negative impact of estimation risk, we study this problem in the context of various realistic forecasting schemes. We design a Monte Carlo (MC) study that includes the rolling and recursive scheme along with the fixed one. This implies the need to carry out nested simulations, which substantially

[2] These accords contain supervisory recommendations issued by the Basel Committee on Banking Supervision (BCBS), which is the international committee setting prudential regulations of banks. It has 45 members including central banks and bank supervisors from 28 jurisdictions.

increase the computational demands of the study; nevertheless it is essential for drawing practical conclusions. We demonstrate that when estimation risk occurs, the careful selection of the forecasting scheme is of particular importance, since it has critical influence on the properties of the test. The results of our research lead to different conclusions than those presented in previous studies, as we provide evidence that adjusting the forecasting scheme yields better results in terms of test accuracy than correcting the estimation error by means of resampling techniques. Thus, we argue that the tests may be effectively applied with the use of asymptotic distributions instead of time-consuming and computationally-intensive simulation methods. Our proposal, however, should involve performing parameter corrections in VaR models at a suitable frequency. These conclusions are of great practical importance considering the supervisory recommendation to incorporate VaR tests in the internal procedures of banks.

Our approach is also novel because it encompasses VaR tests belonging to various classes[3] in one study. By doing so, it expands the range of the earlier comparisons of VaR testing methods provided by Berkowitz et al. (2011) and Pajhede (2017),[4] who did not address the problem of estimation risk. Our study examines the estimation risk effects for VaR tests belonging to six distinct classes. The first class encompasses unconditional coverage tests, represented by the pioneering procedure by Kupiec (1995). This test is important to the banking industry, as it establishes a procedure which underpins international regulatory rules (BCBS, 2017). The remaining five distinct classes of conditional coverage tests are: the Markovian, correlation-based, regression-based, duration-based and spectral tests. In the first group of the above-mentioned tests, the early Markov-chain Christoffersen's one was the leading reference from 1998 until 2017, when it was generalised by Pajhede (2017). Our study, therefore, is based on the generalised Pajhede version of this test. Out of the group of the correlation-based procedures, we focus on the test of the Ljung-Box-type, which draws directly upon the autocorrelation function of violations. This approach is particularly important, as, since Berkowitz et al. (2011) proposed using it for the evaluation of VaR models, it has been commonly treated as a benchmark in empirical studies. Out of the group of regression-based VaR tests, we chose the binary-choice procedure developed by Dumitrescu et al. (2012), who generalised and extended the early test by Engle and Manganelli (2004). To represent the class of the duration-based VaR tests, we selected the geometric-VaR test by Pelletier and Wei (2016). This method generalises the geometric approach proposed by Christoffersen

[3] We limit the scope of our study to univariate VaR tests. Previous research also proposed to employ the multivariate approach, i.e. to use multiple VaR levels (Colletaz et al., 2013; Gordy & McNeil, 2018; Hurlin & Tokpavi, 2006; Kratz et al., 2018; Leccadito et al., 2014; Wied et al., 2016) as a natural extension to these methods.

[4] The recent overview of VaR testing methods provided by Pajhede (2017) encompassed the Markov-chain class tests, autocorrelation tests, regression-based tests and duration-based tests; however, neither did it include the generalisation of the early regression-based test of Dumitrescu et al. (2012) nor the generalisation of the duration-based test of Pelletier and Wei (2016). It also did not include any of the spectral methods.

and Pelletier (2004) and follows a number of other VaR testing procedures related to the geometric distribution (Berkowitz et al., 2011; Candelon et al., 2011; Engle & Russel, 1998; Haas, 2005; Krämer & Wied, 2015; Ziggel et al., 2014). The spectral methods, following research by Małecka (2016), are represented in our study by the Anderson-Darling VaR test, which draws upon the spectral approach proposed by Berkowitz et al. (2011).

The aim of our study is twofold: firstly, to find methods of mitigating the negative impact of the estimation risk on VaR tests, and secondly, to provide a comprehensive comparison of VaR testing methods with reference to the issue of estimation risk. The key element of the study, which is an extension of previous research on estimation risk in VaR tests, is the inclusion of various realistic forecasting schemes.

2. Literature review

The notion of estimation risk in the VaR context (also called the model risk) was defined by Escanciano and Olmo (2010). Prior to their work, studies on testing VaR centred around establishing criteria for distinguishing correct VaR models and proposals of how to verify these criteria. The main achievements in defining these economically relevant, testable criteria are attributed to Kupiec (1995), who proposed the verification of the VaR violation probability, and Christoffersen (1998), who argued that the probability of violating VaR should not only coincide with the chosen VaR coverage level (unconditional coverage criterion – VaR level) but also be constant in time (conditional coverage criterion). The criteria put forward by Kupiec and Christoffersen were accompanied by proposals of tests embedded in the likelihood ratio and first-order Markov-chain framework. These pioneering tests were followed by numerous other proposals of how to verify the unconditional coverage and independence criteria. Among them, there were proposals to use standard approaches like the correlation-based Ljung-Box test or the regression-based tests developed by Dumitrescu et al. (2012) and Engle and Manganelli (2004). The Markov-chain approach with higher-order dependencies was advocated by Pajhede (2017). A large group of the methods proposed as the follow-up to the pioneering tests mentioned before can be classified as duration-based tests. They use the transformation of VaR violations into durations and include tests proposed by Berkowitz et al. (2011), Candelon et al. (2011), Christoffersen and Pelletier (2004), Engle and Russel (1998), Haas (2005), Krämer and Wied (2015), Pelletier and Wei (2016) and Ziggel et al. (2014).

A different perspective was adopted towards spectral tests, which are based on the Fourier transformation of the autocorrelation function. Such an approach was used to test VaR by Berkowitz et al. (2011) and Gordy and McNeil (2018). An important development of these methods were multi-level and multivariate tests by Berkowitz (2001), Colletaz et al. (2013), Hurlin and Tokpavi (2006), Kratz et al. (2018),

Leccadito et al. (2014) and Wied et al. (2016). These competing VaR testing procedures were assessed on the basis of two fundamental statistical properties – the size and the power, which is the standard convention.

Very few papers have taken into account the problem of size distortions caused by estimation risk reported by Escanciano and Olmo (2010). The suggestions of how to handle this problem evolved around simulation methods based on resampling, as initiated by the study of Escanciano and Olmo (2011). They suggested simulation methods are 'natural simpler alternatives' to the asymptotic theory. The authors argued that asymptotic distributions incorporating estimation risk for VaR tests encounter technical problems which could be overcome through robust resampling techniques: the block bootstrap and subsampling. According to their research, these techniques may be used to approximate the actual sampling distributions of VaR test statistics. Their study showed that resampled distributions have an advantage over theoretical ones; however, the research was limited to the assumptions of the fixed forecasting scheme only. Since then, several novel proposals of tests have been presented with regard to estimation risk. Candelon et al. (2011) investigated the influence of estimation risk on their GMM-based VaR test and the possibility of reducing it by implementing subsampling. Dumitrescu et al. (2012) included the bootstrapping technique instead of subsampling as a remedy to the problem of estimation risk occurring in the dynamic-binary-choice VaR test. Pelletier and Wei (2016), who recently proposed an extension to the geometric duration-based VaR test, indicated that the issue of estimation risk should be subject to further studies. The general conclusion from these studies is that the effects of estimation risk drastically decrease the accuracy of tests (where accuracy is understood as a test size compliant with the selected significance level). The subsampling and the block-bootstrap methods reduce this negative influence. However, even when corrected with the use of these methods, the true test size is clearly distinct from the chosen significance level.

The studies above employed subsampling or block-bootstrapping as a means to addressing the problem of estimation risk, as recommended by Escanciano and Olmo (2011). That research was based on the fixed forecasting scheme. The fixed forecasting scheme assumes that parameters are estimated only once from a fixed number of initial observations. We broaden the previous studies by waiving this assumption and including several commonly used forecasting schemes. We show that the proper adjustment of the analysed components produces better effects in dealing with estimation risk than the previously proposed methods.

3. Estimation errors in VaR tests

The criteria developed for backtesting VaR models are based on the properties of the violation process, which compares the forecasted VaR with the realised returns. To formally define these criteria, let us denote R_t as the continuous return from the

portfolio at time t, Ω_t is the set of all information at time t, and $\mathcal{F}_t$ defines the σ-algebra generated by the subsets of Ω_t. The p%-VaR $q_p(R_t)$ is the conditional p-quantile of R_t, i.e. $\mathbb{P}(R_t < q_p(R_t)|\mathcal{F}_t) = p$. Let us assume that VaR forecasts are produced from the parametric model $\mathcal{M} = \{VaR_{t|t-1,p}(\theta): \theta \in \Theta \subset \mathbb{R}^n\}$. The violation process is defined by:

$$I_{t,\theta}(p) = \mathbb{I}\{R_t < VaR_{t|t-1,p}(\theta)\}, \tag{1}$$

where $\mathbb{I}_A(x) = 1$ if $x \in A$, and $\mathbb{I}_A(x) = 0$ if $x \notin A$.

Equation (1) clearly demonstrates that the violation process and its properties which give rise to the VaR testing criteria are heavily dependent on parameter vector θ. The nuisance parameters in θ introduce estimation risk to the backtesting framework, which tends to be neglected in standard backtesting procedures and thus requires verification. As noted by Escanciano and Olmo (2010), these procedures are typically based on the simplifying assumption that such $\theta^\star \in \Theta$ exists that $VaR_{t|t-1,p}(\theta^\star) = q_p(R_t)$. Moreover, it is usually assumed that $\theta^\star$ is known. These assumptions may be jointly written as follows:

$$VaR_{t|t-1,p}(\theta) = VaR_{t|t-1,p}(\theta^\star) = q_p(R_t). \tag{2}$$

According to (2) and provided that the VaR model is correct, the underlying violation process shown in (1) reduces to:

$$I_t(p) = \mathbb{I}\{R_t < q_p(R_t)\}. \tag{3}$$

Based on the process described in (3), the VaR testing criteria are formulated as

$$\mathbb{P}(I_{t,\theta}(p) = 1) = p \tag{4}$$

and

$$\mathbb{P}(I_{t,\theta}(p) = 1|\mathcal{F}_{t-1}) = \mathbb{P}(I_{t,\theta}(p) = 1), \tag{5}$$

which are called the unconditional coverage and independence criterion, respectively. Jointly, these criteria require the conditional probability of violation to be constant and equal to p, i.e.

$$\mathbb{P}(I_{t,\theta}(p) = 1|\mathcal{F}_{t-1}) = p, \tag{6}$$

which is called the conditional coverage criterion.

Conditions (4)–(6) hold true only when (3) is the true violation process. In a realistic setting, however, process (3) is unobservable, while the observed VaR violations are the realisations of (1). These realisations correspond to the out-of-sample VaR forecasts for the observed return series. Assuming series of a length of T, i.e. $\{R_1, \dots, R_T\}$, the beginning R observations are used only to estimate model $\mathcal{M}$ and produce forecasts for the remaining period. The forecasted VaR series $\{VaR_{R+1|R,p}(\theta), \dots, VaR_{T|T-1,p}(\theta)\}$ of the $P = T - R$ length, compared with the P out-of-sample returns $\{R_{R+1}, \dots, R_T\}$, forms observed violation series $\{I_{R+1,\theta}(p), \dots, I_{T,\theta}(p)\}$, where criteria (4)–(6) do not hold true. However, following the standard convention, the asymptotic distributions of the VaR test statistics are based on the assumption that this violation sample is the series of realisations of (3) instead of (1). This may cause the results of the backtesting to be misleading. Also, the typical studies that evaluate the fundamental statistical properties of test statistics – the size and the power – are based on simulations that erroneously assume that, under the correct model, (3) is the violation process.

Two ways of dealing with the problem of estimation risk were proposed by Escanciano and Olmo (2011). The first method involved the development of an asymptotic theory that incorporated the influence of the estimation risk on the evaluation of risk models. Based on this theory, the correct asymptotic distributions of two VaR tests were derived: one dedicated to the unconditional coverage and the other to the conditional coverage hypothesis. However, the authors noted that such distributions may be difficult to derive analytically for other available testing procedures. Moreover, the quick inflow of new, generalised and extended tests makes such an analytical approach infeasible. In order to address these issues, Escanciano and Olmo suggested that the problem of estimation risk be resolved by the employment of simulation methods. They proposed two resampling techniques: subsampling and the block bootstrap to approximate the true sampling distribution of the test statistics.

The idea of the subsampling approach consists in using a subsampling approximation of the actual statistic distribution. This approximation is formed in a way that mimics the actual processes of estimating, forecasting VaR and testing. The approximate distribution is based on values of the test statistic computed on smaller subsets of data – subsamples. The s-th subsample from return data $\{R_1, R_2, \dots, R_T\}$ is defined as $\{R_s, R_{s+1}, \dots, R_{s+u-1}\}$, where $s \in \{1, \dots, T - u + 1\}$ and u is the subsample size. Then these subsamples are divided into two parts that correspond to the division of the initial sample into the in-sample part of length R and the out-of-sample part of length P. Consequently, each subsample has the in-sample part of size R_u and the out-of-sample part of size P_u, where $\frac{P_u}{R_u} = \frac{P}{R}$. The

in-sample parts are used for estimating model $\mathcal{M}$ and the out-of-sample parts to compute the realisations of the test statistic. This technique was adopted by Candelon et al. (2011), who studied the effects of the estimation risk on their GMM VaR tests.

As regards the bootstrap approach, the sampling distribution of the test statistic derives from estimating model $\mathcal{M}$ on the basis of resamples of size R drawn with the replacement from the standardised residual series. These residuals are obtained with the use of the parameter estimates from the in-sample parts $\{R_1, R_2, \dots, R_R\}$. Following the resampling phase, the residuals are transformed into bootstrap samples $\{R_1^B, R_2^B, \dots, R_R^B\}$, which produce bootstrap parameter estimates and out-of-sample bootstrap VaR forecasts. In effect, a new violation series is formed and used for testing. A sufficient number of repetitions yields series of realisations of the test statistic, which approximates its distribution. Such a technique was applied in the study by Dumitrescu et al. (2012), where the dynamic binary choice VaR test was developed.

The studies by Candelon et al. (2011) and Dumitrescu et at. (2012) led to two main conclusions: test accuracy drastically deteriorates when estimation errors occur, and resampling methods reduce the effects of the estimation errors. However, even corrected with the help of these methods, the true test size was clearly distinct from the chosen significance level. An important limitation to these studies from the practical point of view was the simplifying assumption that VaR forecasts were obtained from the fixed forecasting scheme. This scheme takes for granted that parameters are estimated only once from a fixed number of beginning observations. We ignore this assumption and investigate the influence of estimation errors under several commonly used forecasting schemes.

The forecasting schemes we analyse include three main types: fixed, rolling and recursive. While the fixed scheme relies on the same parameter estimates throughout the whole sample, the other schemes involve parameter corrections, which are made at a given frequency. To reflect these components in the violation process, we adopt the following notation:

- $\{I_{t,\theta^\star}(p)\}$ – the violation process based on the true VaR;
- $\{I_{t,\theta_f}(p)\}$ – the violation process based on VaR predictions made by a fixed forecasting scheme;
- $\{I_{t,\theta_{rol}^d}(p)\}$ – the violation process based on VaR predictions made by a rolling forecasting scheme with a parameter correction made every d days;
- $\{I_{t,\theta_{rec}^d}(p)\}$ – the violation process based on VaR predictions made by a recursive forecasting scheme with a parameter correction made every d days.

The analysis of the above forecasting schemes is directly related to the first aim of our study – to find ways of coping with the problem of estimation risk. As opposed

to the previous studies, our research examines the possibilities of reducing the effects of estimation risk by adjusting the components of the forecasting scheme. In order to indicate such possibilities, we study the influence of the forecasting scheme components on the accuracy of the VaR test. We compare our outcomes with the results of the estimation errors correction through the previously proposed resampling methods.

Our research extends the previous studies on the undertaken subject also in the sense that it includes VaR tests belonging to six distinct classes, which allows for a broad comparison of these methods. In addition to the unconditional coverage test that underpins the international regulatory framework, we examine joint conditional coverage tests. Within this group, we first select the representatives of the testing procedures that directly refer to the correlation structure of the violation process. We consider methods developed in the standard Markov-chain convention or with the use of the autocorrelation function estimates. We then include indirect methods using the regression-based or the duration-based approach. Finally, we examine methods that were developed with the use of the spectral theory. Whenever available, we use the generalised versions of the test statistics.

4. Backtesting framework

Depending on the class, VaR backtests either refer to conditions (4)–(6) directly or operate on some implications of these conditions. They are most often designed in a way that allows tests of joint hypothesis (6). This hypothesis is typically reduced to a simplified criterion:

$$\mathbb{P}\left(I_{t,\theta}(p) = 1 \middle| I_{t-1,\theta}(p), I_{t-2,\theta}(p), \ldots\right) = p, \tag{7}$$

which is a special case of (6). This criterion is equivalent to the requirement that the $\{I_{t,\theta}(p)\}$ process is i.i.d.[5] Bernoulli with parameter p. While recent VaR tests generally correspond with the conditional coverage criterion in either of the two forms: (6) or (7), the pioneer VaR test by Kupiec (1995) was dedicated to checking unconditional coverage (4) exclusively. This test is still important for two reasons: firstly, it continues to be used by the international regulator and secondly, it serves to complement more complex procedures that verify VaR violation independence but have low power against incorrect unconditional coverage. This test checks (4) through the likelihood ratio statistic:

$$LR_{uc}^{K} = -2(P_1\log(p) + (P - P_1)\log(1 - p) - P_1\log(\hat{p}) - (P - P_1)\log(1 - \hat{p})), \tag{8}$$

[5] Independent and identically distributed.

where P_1 is the number of VaR violations in the out-of-sample P returns $\{R_{R+1}, R_2, \dots, R_T\}$ and $\hat{p}$ is the ML estimate of the violation probability given by $\hat{p} = \frac{P_1}{P}$.

The Markov-chain framework proposed by Christoffersen (1998), which extends the LR_{uc}^K approach to a joint test of the conditional coverage hypothesis, operates on first-order dependencies. To improve its power, it was generalised by Pajhede (2017) into a generalised Markov-chain VaR test. This test is based on two types of generalised probabilities: the excited p_E and the steady p_S. They describe the probability of VaR violation in $I_{t,\theta}(p)$ on condition that the previous violation has occurred (p_E) or has not occurred (p_S) in $\{I_{t-1,\theta}(p), I_{t-2,\theta}(p), \dots, I_{t-h,\theta}(p)\}$. The ML estimates of these probabilities are $\hat{p}_E = T_{11}/(T_{10} + T_{11})$ and $\hat{p}_S = T_{01}/(T_{00} + T_{01})$, where $P_{11} = \sum_{t=R+1}^{T} I_t J_{t-1}$, $P_{01} = \sum_{t=R+1}^{T} I_t(1 - J_{t-1})$, $P_{10} = \sum_{t=R+1}^{T} (1 - I_t)J_{t-1}$, $P_{00} = \sum_{t=R+1}^{T} (1 - I_t)(1 - J_{t-1})$, and $J_{t-1} = = 1\{\sum_{i=1}^{h} I_{t-i} > 0\}$.[6] The estimates of the excited and the steady probabilities are used to build the generalised Markov-chain likelihood ratio test statistic:

$$LR_{cc}^{M,gen} = -2\big((T_{01} + T_{11})\log(p) + (T_{00} + T_{10})\log(1 - p) + -T_{11}\log(\hat{p}_E) - T_{10}\log(1 - \hat{p}_E) - T_{01}\log(\hat{p}_S) - T_{00}\log(1 - \hat{p}_S)\big). \tag{9}$$

Another direct way to conduct VaR backtesting on a custom number of dependencies between violations is through sample autocorrelations:

$$\rho(k) = \frac{1}{P - k} \sum_{t=R+k+1}^{T} \frac{(I_{t,\theta}(p) - p)(I_{t-k,\theta}(p) - p)}{p(1 - p)}. \tag{10}$$

This idea was utilised for example by Berkowitz et al. (2011), who proposed a VaR correlation test in the spirit of Ljung and Box (1978). This test uses the statistic in the following form:

$$LB_{cc} = P(P + 2) \sum_{k=1}^{m} \frac{(\rho(k))^2}{P - k}, \tag{11}$$

where m is a chosen lag order.

[6] This corrects formula (18) from Pajhede (2017), where $J_{t-1} = 1\{\sum_{i=1}^{k} I_{t-1} > 0\}$, which in our notation takes the form of $J_{t-1} = 1\{\sum_{i=1}^{h} I_{t-i} > 0\}$.

The regression-based approach, which constitutes another class of tests, is built on the observation that under (6) the conditional expectation of demeaned process $I_{t,\theta}(p) - p$ is equal to zero, i.e.

$$\mathbb{E}(I_{t,\theta}(p) - p|\mathcal{F}_{t-1}) = 0. \tag{12}$$

This expectation can be modelled by linear regression. In that case restriction (12) can be verified by means of testing the significance of the regression parameters. However, the binary character of the dependent variable implies that residuals from such a regression are heteroscedastic. Thus, the early proposition based on simple linear regression is replaced with tests that use the binary choice models. In the generalised test, proposed by Dumitrescu et al. (2012), the dynamic binary response model

$$\mathbb{P}(I_{t,\theta}(p) = 1|\mathcal{F}_t) = E(I_{t,\theta}(p)|\mathcal{F}_t) = F(\pi_t) \tag{13}$$

uses c.d.f.[7] F (usually Gaussian or exponential) and index π_t, which is given by the autoregressive equation. The general form of its representation is as follows:

$$\begin{aligned}\pi_t = \eta + \sum_{i=1}^{q_1} \lambda_i \pi_{t-i} + \sum_{i=1}^{q_2} \phi_i I_{t-i,\theta}(p) + \sum_{i=1}^{q_3} \psi_i l(x_{t-i}) + \\ + \sum_{i=1}^{q_4} \gamma_i l(x_{t-i}) I_{t-i,\theta}(p),\end{aligned} \tag{14}$$

where l links the variables from Ω_t with π_t and q_1, q_2, q_3 and q_4 are chosen lag orders. The restrictions on the parameters, corresponding to (12) are tested through the likelihood ratio of the form:

$$\begin{aligned}LR_{cc}^{db} = -2(\log L\left(\{I_{R+1,\theta}(p), \dots, I_{T,\theta}(p)\}\right|\hat{\eta}, \hat{\lambda}, \hat{\phi}, \hat{\psi}, \hat{\gamma}) + \\ - \log L(\{I_{R+1,\theta}(p), \dots, I_{T,\theta}(p)\}|\eta = F^{-1}(p), \lambda = \phi = \psi = \gamma = 0)),\end{aligned} \tag{15}$$

where $\log L(\{I_{R+1,\theta}(p), \dots, I_{T,\theta}(p)\}|\hat{\eta}, \hat{\lambda}, \hat{\delta}, \hat{\psi}, \hat{\gamma}) = \sum_{t=R+1}^{T} (I_{t,\theta}(p) \log F(\pi_t) + - (1 - I_{t,\theta}(p)) \log (1 - F(\pi_t)))$, $\lambda = (\lambda_1, \dots, \lambda_{q_1})$, $\phi = (\phi_1, \dots, \phi_{q_2})$, $\psi = (\psi_1, \dots, \psi_{q_3})$, $\gamma = (\gamma_1, \dots, \gamma_{q_4})$.

Another way of verifying criterion (6) is through the duration-based framework. It transforms violation process $\{I_{t,\theta}(p)\}$ into a $\{D_{j,\theta}(p)\}$ process of durations and requires that the latter process follows the geometric distribution with parameter p. The transformation into durations is defined as:

[7] Cumulative distribution function.

$$D_{j,\theta}(p) = t_j - t_{j-1}, \tag{16}$$

where t_j denotes the time of the j-th VaR violation in process $\{I_{t,\theta}(p)\}$. The generalised geometric-VaR test by Pelletier and Wei (2016) confronts the hazard function of the geometric distribution with the following general form of a hazard function:

$$\lambda_d^j = ad^{b-1}e^{-cVaR_{t_{j-1}+d}}, \tag{17}$$

where $0 \leq a < 1$, $0 \leq b \leq 1$, $c \geq 0$. The relevant restrictions on its parameters are verified through the likelihood ratio:

$$\begin{aligned} LR_{cc}^{geom,VaR} = -2(&\log L\left(\{D_{1,\theta}(p), \dots, D_{N,\theta}(p)\}\middle|\hat{a}, \hat{b}, \hat{c}\right) + \\ &- \log L(\{D_{1,\theta}(p), \dots, D_{N,\theta}(p)\}|a = p, b = 1, c = 0)), \end{aligned} \tag{18}$$

where
$\log L\left(\{D_{1,\theta}(p), \dots, D_{N,\theta}(p)\}|a, b\right) = C_1 \log S\left(D_{1,\theta}(p)\right) + (1 - C_1)\log f\left(D_{1,\theta}(p)\right) +$
$+ \sum_{j=2}^{N-1} \log f(D_{j,\theta}(p)) + C_N \log S(D_{N,\theta}(p)) + (1 - C_N) \log f(D_{N,\theta}(p)), \quad f(d) =$
$= (1 - \lambda_1^j)(1 - \lambda_2^j) \dots (1 - \lambda_{d-1}^j)\lambda_d^j, \quad S(d) = 1 - (1 - \lambda_1^j)(1 - \lambda_2^j) \dots (1 - \lambda_{d-1}^j)$,
$\{D_{1,\theta}(p), \dots, D_{N,\theta}(p)\}$ is the sample of durations and $\{C_1, \dots, C_N\}$ indicates censoring.

A different transformation of the $\{I_{t,\theta}(p)\}$ process is used within the spectral VaR testing framework. These methods rely on spectral density – the spectral transform of the autocovariance function of the violation process. Since under the i.i.d. Bernoulli violations, this spectral transform is a flat function, the tests consist in comparing the observed spectral density with the theoretical flat shape. The Anderson-Darling test statistic proposed by Małecka (2016) is given by:

$$SD_{cc}^{AD} = \int_0^1 \frac{U(t)^2}{t(1-t)} dt, \tag{19}$$

where $U(t)$ is the function that measures the distance between the estimated and theoretical spectral density. It is given by:

$$U_P(t) = \sqrt{2P} \int_0^{\pi t} \left(\frac{\mathcal{P}_P(\omega)}{\sigma(0)} - \frac{1}{2\pi} \right) d\omega = \frac{\sqrt{2}}{\pi} \sum_{k=1}^{P-1} \sqrt{P}\rho(k) \frac{\sin k\pi t}{k}, \quad t \in [0, 1], \tag{20}$$

where $\mathcal{P}_P(\omega) = \frac{1}{2\pi}\sum_{k=-(P-1)}^{P-1} \sigma(k)e^{-ik\omega}$ is the periodogram estimate of the spectral density, $\sigma(k)$ and $\rho(k)$, respectively, are the estimates of the k-th order autocovariances and autocorrelations of the violation process.

Under standard conditions, the likelihood ratio statistics are asymptotically χ^2 distributed with the number of degrees of freedom corresponding to the number of tested restrictions. Therefore, the Kupiec LR_{uc}^{K}, the Markov-chain $LR_{cc}^{M,gen}$, the Ljung-Box LB_{cc} and the dynamic binary LR_{cc}^{db} tests use χ_1^2, χ_2^2, χ_k^2 and $\chi_{q_1+q_2+q_3+q_4+1}^2$ distributions, respectively. The theoretical distribution of the duration-based statistic diverges from the standard asymptotic likelihood ratio distribution due to the specific boundary case. It has been shown to be a mixture distribution of form $0.25\chi_1^2 + 0.5\chi_2^2 + 0.25\chi_3^2$ (Małecka, 2016). The SD_{cc}^{AD} test is based on the Anderson-Darling distribution (Anderson & Darling, 1952).

5. Robust inference

5.1. Size of VaR tests

The standard VaR backtesting procedures are based on the theoretical distributions given in Section 4. Whether such tests guarantee a reliable classification of VaR models depends on several aspects. Firstly, tests based on these distributions need to be well-sized. Here, by 'well-sized' we understand well-sized under the absence of estimation risk. This means that, under the absence of estimation risk, the frequency of rejecting correct models should coincide with the chosen significance level. Since all the considered distributions are based on limit theorems, this property may be violated as a result of applying asymptotic properties to finite samples. Secondly, the tests need to be robust to estimation risk. As argued in Section 3, the convergence of the test statistics to their theoretical distributions may be affected by this additional factor. The estimation risk results from the noise in the violation process that follows from the procedure of estimating and forecasting VaR with the effect that proper risk models may be overrejected due to estimation errors. Thirdly, the tests should be effective at detecting incorrect models. In particular, they should be powerful both against incorrect unconditional and conditional coverage. Out of these three aspects, we focus on the influence of estimation risk on test accuracy. However, to properly combine all the elements of the testing procedure, the evaluation of the effects of the estimation risk is preceded by a size study that assumes the absence of such a risk. This study serves two purposes: it indicates the tests which have the potential to be well-sized in finite samples when based on asymptotic distributions, and it provides benchmark size estimates for a further robustness check. The tests

that do not meet the requirements of this preliminary size assessment are excluded from the further phases of the procedure. What then follows is the core part of the process which introduces estimation risk and finally, both parts are complemented by the power evaluation.

The finite sample test sizes are evaluated by means of the MC method based on rejection frequencies observed in simulations which assume correct conditional coverage. To ensure the comparability with previous studies (Candelon et al., 2011; Dumitrescu et al., 2012; Escanciano & Olmo, 2010, 2011), our simulations are based on the t-GARCH process of the $R_t = \sqrt{h_t}\epsilon_t$ form, where ϵ_t follows Student t distribution t_v and the variance is represented by:

$$h_t = \omega_1 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}^2, \tag{21}$$

with the following parameter values: $\omega_1^* = 0.0001$, $\alpha_1^* = 0.1$, $\beta_1^* = 0.85$ and $v^* = 10$. Therefore, model $\mathcal{M}$ in our study is $\mathcal{M} = \{VaR_{t|t-1,p}(\omega_1, \alpha_1, \beta_1, v)$: $\omega_1, \alpha_1, \beta_1 > 0, \alpha_1 + \beta_1 < 1\}$. The returns generated by the process above are compared with the VaR estimates set to $VaR_{t|t-1,p}(\omega_1^*, \alpha_1^*, \beta_1^*) = \sqrt{h_t} F_{t_{v^*}}^{-1}(p)$, where F_{t_v} denotes the c.d.f. of Student t distribution t_v. The resulting violation process – $\{I_{1,\theta^*}(p), \dots, I_{P,\theta^*}(p)\}$ – is then used to conduct testing procedures. The sample sizes are as follows: $P = 250, 500, 750, 1{,}000, 1{,}250, 1{,}500$ and the standard significance levels: $\alpha = 0.01, 0.05, 0.1$. The VaR coverage level is first set to the typical $p = 5\%$ and subsequently to $p = 1\%$, which corresponds to the current trends in regulatory standards. The rejection frequencies are computed over 10,000 MC repetitions.

The study combines VaR backtesting procedures from various classes. The first procedure is the LR_{uc}^K unconditional coverage test by Kupiec (1995), which continues to be the underlying procedure of the Basel standards (BCBS, 2017). This test is followed by conditional coverage tests belonging to five classes. Wherever available, we use the generalised versions of the test statistics. Thus, from the Markov-chain framework, we apply the generalised $LR_{cc}^{M,gen}$ test by Pajhede (2017). Relying on the primary study of its properties, we select parameters p_E and p_S to be the violation probabilities on condition that the previous violation occurred in the last five observations. The same lag order is chosen in the LB_{cc} Ljung-Box-type test, which is based on Berkowitz et al. (2011). This procedure represents the correlation class of VaR tests. In the LR_{cc}^{db} dynamic binary test, representing the class of regression-based methods, we adopt the following representation of the π_t index: $\pi_t = \eta + \lambda_1 \pi_{t-1}$. This choice follows from the accuracy results presented by Dumitrescu et al. (2012), who developed this test. The broad class of the duration-

based methods is represented by the $LR_{cc}^{geom,VaR}$ generalised statistic proposed by Pelletier and Wei (2016), which uses time dependencies in the violation process as well as lagged VaR forecasts as explanatory variables. To represent the spectral methods, we apply the SD_{cc}^{AD} test based on the Anderson-Darling statistic, as advocated by Małecka (2016).

The size results under the absence of estimation risk, presented in Table 1, show substantial differences in test accuracy both among the specific procedures and between the VaR coverage levels. There is clearly less accuracy when operating on the lower, 1% VaR level. For 1% VaR, the only tests that may be treated as satisfactory in terms of size are the LR_{uc}^{K} Kupiec test and the SD_{cc}^{AD} spectral test. In the latter case, this is additionally limited to the highest significance level of 0.1. One more restriction that holds for such a low VaR level is that the sample size should consist of at least 750 observations for both tests. As regards the 5% VaR level, the accuracy of the test improves greatly. In this case, most considered procedures seem to be well-sized at all significance levels. These are the LR_{uc}^{K}, $LR_{cc}^{M,gen}$, LB_{cc}, $LR_{cc}^{geom,VaR}$, and SD_{cc}^{AD} tests. The recommendable sample sizes depend on the type of procedure, but for most tests, they start with 250 or 500 observations. Only the LB_{cc} Ljung-Box and SD_{cc}^{AD} spectral tests seem to require longer samples of e.g. 1,000 or more observations. One clear exception emerges in this general description of the convergence of the true test sizes to nominal levels – the LR_{cc}^{db} dynamic binary test. This procedure appears to be systematically undersized. Its true size most often proves to be below half of the nominal significance level. Due to such a large level of inaccuracy, this test is considered not fit for implementation based on the theoretical distribution. Thus, we exclude it from the further parts of our study. Systematic discrepancies are also observed for the Ljung-Box test. This test, in contrast to the LR_{cc}^{db}, seems to be oversized. However, in this case, the inaccuracies are on a much smaller scale, so we decide to include this test in our research.

As the LR_{uc}^{K}, $LR_{cc}^{M,gen}$, LB_{cc}, $LR_{cc}^{geom,VaR}$, and SD_{cc}^{AD} procedures appear to satisfy the correct-size property under the absence of estimation risk, we treat them as having the potential to be well-sized also under the presence of such a risk. In these circumstances, they may be recommended for use with asymptotic distributions. However, their robustness to estimation errors needs to be checked in relevant simulation experiments, which we carry out in the next part of our study.

Table 1. Size estimates for the VaR tests under the absence of estimation risk by significance levels

Series length	5% VaR					
	LR_{uc}^{K}	$LR_{cc}^{M,gen}$	LB_{cc}	LR_{cc}^{db}	$LR_{cc}^{geom,VaR}$	SD_{cc}^{AD}
0.01						
250	0.0084	0.0101	0.0482	0.0049	0.0097	0.0162
500	0.0122	0.0165	0.0342	0.0050	0.0096	0.0124
750	0.0133	0.0097	0.0207	0.0050	0.0114	0.0095
1,000	0.0098	0.0100	0.0190	0.0046	0.0113	0.0112
1,250	0.0106	0.0101	0.0170	0.0040	0.0122	0.0124
1,500	0.0099	0.0114	0.0168	0.0060	0.0099	0.0106
0.05						
250	0.0582	0.0529	0.1160	0.0309	0.0447	0.0416
500	0.0536	0.0518	0.0810	0.0274	0.0504	0.0403
750	0.0559	0.0538	0.0666	0.0230	0.0574	0.0419
1,000	0.0525	0.0484	0.0664	0.0230	0.0545	0.0538
1,250	0.0433	0.0518	0.0635	0.0230	0.0549	0.0497
1,500	0.0496	0.0534	0.0619	0.0280	0.0471	0.0489
0.1						
250	0.1121	0.1483	0.1446	0.0459	0.0928	0.0692
500	0.1018	0.1180	0.1237	0.0488	0.1067	0.0729
750	0.1146	0.1019	0.1110	0.0570	0.1096	0.0763
1,000	0.1092	0.1028	0.1132	0.0400	0.1033	0.0896
1,250	0.1048	0.0978	0.1107	0.0510	0.1086	0.0947
1,500	0.0924	0.1050	0.1063	0.0490	0.0967	0.1027

(cont.)

Series length	1% VaR					
	LR_{uc}^{K}	$LR_{cc}^{M,gen}$	LB_{cc}	LR_{cc}^{db}	$LR_{cc}^{geom,VaR}$	SD_{cc}^{AD}
0.01						
250	0.0039	0.0037	0.1173	0.0030	0.0010	0.0355
500	0.0042	0.0031	0.1978	0.0089	0.0034	0.0416
750	0.0076	0.0042	0.0599	0.0030	0.0060	0.0449
1,000	0.0120	0.0074	0.0691	0.0070	0.0099	0.0447
1,250	0.0102	0.0068	0.0513	0.0020	0.0091	0.0394
1,500	0.0120	0.0068	0.0593	0.0010	0.0070	0.0368
0.05						
250	0.0142	0.0174	0.1173	0.0088	0.0111	0.0438
500	0.0627	0.0244	0.1978	0.0176	0.0218	0.0589
750	0.0362	0.0259	0.2984	0.0124	0.0383	0.0768
1,000	0.0556	0.0298	0.0883	0.0270	0.0423	0.0800
1,250	0.0663	0.0436	0.1087	0.0150	0.0364	0.0785
1,500	0.0563	0.0374	0.0885	0.0180	0.0372	0.0810
0.1						
250	0.0422	0.0432	0.1173	0.1071	0.0281	0.0611
500	0.0627	0.0761	0.1978	0.0590	0.1025	0.0718
750	0.0974	0.0644	0.2998	0.0382	0.0891	0.0987
1,000	0.1155	0.0864	0.1462	0.0510	0.0715	0.1033
1,250	0.1193	0.0716	0.1259	0.0430	0.0714	0.1091
1,500	0.1248	0.0720	0.1401	0.0410	0.0743	0.1081

Source: author's calculations.

5.2. Estimation risk

Our study of the robustness to estimation risk focuses on the components of the forecasting scheme, i.e. the type of the scheme and the frequency of the parameter correction. To the best of our knowledge, the influence of the forecasting scheme on the VaR testing framework has not yet been examined. The available studies are limited to fixed forecasting schemes only. These papers show the dramatic effects that estimation errors have on test accuracy and advocate the use of the subsampling (Candelon et al., 2011) or bootstrap technique (Dumitrescu et al., 2012) to correct these errors. In interpreting their results, however, two factors are of practical importance. Firstly, such results cannot be generalised, since, by definition, the fixed forecasting scheme creates the largest estimation risk. Secondly, the industry standard requires the adoption of other schemes and the correction of parameters with a predefined frequency. Therefore, our study includes several realistic forecasting schemes.

The MC study of the influence of estimation errors on test accuracy requires generating a T-element sample of the returns from model $\mathcal{M}$, which serves to provide both the in-sample data used for estimating the parameters of $\mathcal{M}$, and the out-of-sample data used to forecast the violation process and conduct the testing. The rejection frequency observed over a sufficiently large number of repetitions during the procedure of estimation and testing gives the approximate test size. Due to the inclusion of the estimation process, such a proxy for the test size involves estimation risk.

In order to include the forecasting scheme effects, we start with a fixed forecasting scheme which assumes that parameters are estimated once, from the R beginning observations. These estimates serve to produce VaR forecasts throughout the remaining P-observation period, where $P = T - R$. Based on the obtained forecasts, we generate violation process $\{I_{R+1,\theta_f}(p), \dots, I_{T,\theta_f}(p)\}$. Then, we proceed to the rolling scheme, in which the estimation window of length P is moved over the sample and used to produce one-day-ahead forecasts. We apply this scheme at different frequencies of parameter correction: every day, every five days and every 10 days. We obtain the following violation processes: $\{I_{R+1,\theta^1_{rol}}(p), \dots, I_{T,\theta^1_{rol}}(p)\}$, $\{I_{R+1,\theta^5_{rol}}(p), \dots, I_{T,\theta^5_{rol}}(p)\}$ and $\{I_{R+1,\theta^{10}_{rol}}(p), \dots, I_{T,\theta^{10}_{rol}}(p)\}$, respectively, from these procedures. For the recursive scheme, the procedures are analogous, except that the initial estimation window of length P is not moved but extended to include more observations. Everyday, five-day and 10-day parameter corrections serve to generate the following violation processes: $\{I_{R+1,\theta^1_{rec}}(p), \dots, I_{T,\theta^1_{rec}}(p)\}$, $\{I_{R+1,\theta^5_{rec}}(p), \dots, I_{T,\theta^5_{rec}}(p)\}$ and $\{I_{R+1,\theta^{10}_{rec}}(p), \dots, I_{T,\theta^{10}_{rec}}(p)\}$, respectively. Each of these violation processes involves a different level of estimation risk. We assess the

impact of this risk on the test size by applying all of the considered VaR tests to all of the generated violation processes during each MC repetition. The size estimates produced from these procedures are called uncorrected size estimates. Despite the fact that the procedures involve parameter corrections, the word 'uncorrected' signifies consistency with the previous studies and refers to the use of standard asymptotic distributions instead of the resampled ones. The uncorrected sizes corresponding to the various forecasting schemes are computed over 5,000 MC repetitions.

A natural extension to our robustness check is the incorporation of the forecasting scheme components which we study as a means of controlling the estimation risk. For this purpose, we compare the scale of the reduction in the estimation risk achieved with the help of the most suitable forecasting scheme with the performance of the subsampling technique,[8] proposed previously to address the issue of estimation risk. The sizes corrected by the subsampling method are obtained by using subsampled approximates of the test statistic distributions. The subsampled distrubutions are generated at each MC repetition. Generating them requires that the subsets of length u, of form $\{R_s, R_{s+1}, \dots, R_{s+u-1}\}$, $s = 1, \dots, T-u+1$ are divided into the in-sample parts of length R_u and the out-of-sample parts of length P_u, so that $\frac{P_u}{R_u} = \frac{P}{R}$ and $R_u + P_u = u$. These subsets are moved across the data, which gives $N_u = T-u+1$ repetitions. For each $s = 1, \dots, T-u+1$, the following substeps are conducted:[9]

1. estimation of the parameters of model $\mathcal{M}$ from the in-sample part: $\{R_s, R_{s+1}, \dots, R_{s+R_u-1}\}$;
2. producing a VaR violation process: $\{I_{s+R_u}(p), \dots, I_{s+u-1}(p)\}$;
3. computing the test statistic from: $\{I_{s+R_u}(p), \dots, I_{s+u-1}(p)\}$.

The steps above are repeated 1,000 times to receive the subsampling-corrected distribution. With this distribution, we compute the p-value for the test statistics and make test decisions. From a large number of repetitions which involve a whole simulation procedure and together with the nested simulations from substeps 1–3, we can approximate the test sizes into what we call subsampling-corrected sizes. Since such nested simulations are very time-consuming, we set the number of MC repetitions to 1,000. Together with 1,000 repetitions of substeps 1–3, it gives $1{,}000 \cdot 1{,}000 = 1{,}000{,}000$ repetitions.

The study of the effects of estimation risk includes all the previously considered tests except for the LR_{cc}^{db} procedure. This test is rejected as it generates the largest

[8] The choice of the subsampling technique as a representative of the resampling methods was determined by our preliminary study, where this method proved more accurate than the bootstrap one.

[9] A detailed description of the subsampling technique for correcting estimation risk can be found e.g. in Candelon et al. (2011).

size distortions, which show the irrelevance of using its asymptotic distribution in finite samples. Therefore, it does not fulfil our aim of indicating asymptotic tests that are both accurate and robust to estimation risk. We provide the results for a 5% VaR and a 0.05 significance level.

The estimates of the uncorrected and corrected sizes are presented in Table 2. The results for the fixed forecasting scheme confirm the conclusions from Candelon et al. (2011) about the large influence of estimation risk on the accuracy of tests, i.e. estimation errors cause the test accuracy to decrease dramatically. The uncorrected test sizes reach the level of over 0.1 instead of the nominal 0.01 level, 0.3 instead of 0.05 and 0.4 instead of 0.1. These results almost exactly match the scale of influence reported by Candelon et al. (2011).

Table 2. Size estimates of VaR tests under the presence of estimation risk by significance levels

Assumption	LR_{uc}^{K}	$LR_{cc}^{M,gen}$	LB_{cc}	$LR_{cc}^{geom,VaR}$	SD_{cc}^{AD}
0.01					
Absence of estimation risk	0.0098	0.0100	0.0190	0.0113	0.0112
Fixed scheme: uncorrected asymptotic distribution	0.1594	0.1476	0.1052	0.1833	0.0469
subsampling corrected distribution	0.0266	0.0280	0.0300	0.0280	0.0250
Rolling scheme: 10-day parameter correction	0.0037	0.0036	0.0334	0.0244	0.0148
5-day parameter correction	0.0031	0.0033	0.0294	0.0234	0.0133
everyday parameter correction	0.0035	0.0039	0.0240	0.0211	0.0091
Recursive scheme: 10-day parameter correction	0.0150	0.0125	0.0271	0.0173	0.0142
5-day parameter correction	0.0141	0.0114	0.0261	0.0179	0.0131
everyday parameter correction	0.0148	0.0107	0.0240	0.0169	0.0114
0.05					
Absence of estimation risk	0.0525	0.0484	0.0664	0.0545	0.0538
Fixed scheme: uncorrected asymptotic distribution	0.2667	0.2664	0.2036	0.3498	0.1261
subsampling corrected distribution	0.0680	0.0650	0.0540	0.1160	0.0620
Rolling scheme: 10-day parameter correction	0.0213	0.0319	0.1206	0.1046	0.0679
5-day parameter correction	0.0201	0.0293	0.1126	0.1035	0.0626
everyday parameter correction	0.0199	0.0284	0.1007	0.0997	0.0544
Recursive scheme: 10-day parameter correction	0.0607	0.0585	0.0951	0.0847	0.0597
5-day parameter correction	0.0594	0.0561	0.0922	0.0846	0.0566
everyday parameter correction	0.0584	0.0561	0.0880	0.0836	0.0547
0.1					
Absence of estimation risk	0.1092	0.1028	0.1132	0.1033	0.0896
Fixed scheme: uncorrected asymptotic distribution	0.3393	0.3564	0.2683	0.4610	0.1932
subsampling corrected distribution	0.0833	0.1100	0.0800	0.2161	0.0762
Rolling scheme: 10-day parameter correction	0.0591	0.0716	0.1957	0.1961	0.1251
5-day parameter correction	0.0578	0.0681	0.1841	0.1959	0.1165
everyday parameter correction	0.0585	0.0645	0.1719	0.1925	0.1049
Recursive scheme: 10-day parameter correction	0.1106	0.1111	0.1580	0.1649	0.1127
5-day parameter correction	0.1113	0.1104	0.1550	0.1640	0.1104
everyday parameter correction	0.1115	0.1077	0.1475	0.1659	0.1059

Note. All size results in the table assume testing a 5% VaR. The results indicating the absence of estimation risk have been provided for the purpose of comparison.

Source: author's calculations.

The results also confirm that the subsampling technique reduces the effects of the estimation errors. The match between the subsampling-corrected sizes and the nominal levels depends on the adopted procedure. However, looking at all the procedures, the corrected sizes fit in a 0.02–0.03 interval for the nominal level of 0.01, in 0.05–0.11 for 0.05, and in 0.08–0.21 for 0.1. These size results are closer to the desired values than the uncorrected ones, but still, the quality of the match may be questioned during practical applications.

From the point of view of our aims, the key conclusions refer to the effects that various forecasting schemes have on estimation risk. The inclusion of the rolling and recursive schemes into the studied processes significantly changes the influence of estimation risk on the test sizes. This effect is large enough to alter the conclusions presented in the previous studies, advocating for the application of resampling methods. Our research demonstrates that in most cases, the employment of a properly selected forecasting scheme results in better-sized tests than the use of a subsampling technique. In two cases – for $LR_{cc}^{geom,VaR}$ and SD_{cc}^{AD} – any forecasting scheme with a parameter correction, regardless of whether it is rolling or recursive, yields better results in terms of the test size than the subsampling method. For LR_{uc}^{K} and $LR_{cc}^{M,gen}$, the recursive scheme is recommended. This scheme takes full advantage of the available sample size for estimating the parameters, so the sample length is crucial in these cases. For these two tests, the results obtained with the help of the recursive scheme at any frequency of parameter correction are more accurate than the ones coming from the subsampled distributions. As regards the improvements obtained in various tests resulting from the changes of the forecasting scheme, the largest gains in test accuracy occur along a shift from a fixed forecasting scheme to a rolling or recursive scheme. The differences then between daily, five-day or 10-day corrections are systematic although rather minor.

The only test where subsampling outperforms testing based on asymptotic distribution and where the parameter corrections in any of the schemes do not reduce this effect is the Ljung-Box LB_{cc} test. This confirms our results from Section 5.1 of our study of the test size, where it was examined assuming the absence of estimation risk and assuming the test being systematically oversized. After changing the forecasting scheme, this effect is still present. Therefore, we conclude that the use of asymptotic distribution is not recommended for this test.

Our conclusions about the influence of the parameter corrections on reducing the negative effects of estimation risk are important in terms of the practical applications of VaR tests. We demonstrate that most of the tests may be effectively implemented with asymptotic distributions without the need to apply simulation methods. This translates into the time needed for implementing the testing procedures. The time

needed by subsampling for computational purposes is incomparably longer than the time needed for parameter corrections. What is more, the complexity associated with introducing parameter corrections into the estimation procedure is negligible compared to the complexity connected with the implementation of the subsampling technique.

5.3. Power of VaR tests

The power evaluation complements our study of test properties and serves to choose the procedures that are most effective in detecting incorrect VaR models. This part of our study focuses on the tests classified as well-sized and robust to estimation risk under a suitable forecasting scheme. These are: the Kupiec LR_{uc}^{K}, the generalised Markov $LR_{cc}^{M,gen}$, the geometric-VaR $LR_{cc}^{geom,VaR}$, and the spectral SD_{cc}^{AD} tests.

The power study we conduct involves various forms of distortions from the null which correspond to the construction of the considered tests. Since we examine both unconditional and conditional coverage tests, we need two types of alternatives. Moreover, in the class of the conditional coverage tests, some procedures (e.g. the spectral test) are designed mainly with the aim of testing the independence rather than the joint conditional coverage hypothesis. Their classification as conditional coverage tests follows from the inclusion of parameter p in their statistics; however, intuitively, they exhibit low power against the incorrect unconditional coverage. To take into account all these cases, we evaluate the test power in a two-stage simulation study. The first set of experiments involves various violations of the unconditional coverage postulate, while the experiments carried out during the second stage are designed to violate the independence property exclusively (the unconditional coverage and independence properties jointly form the conditional coverage property).

Each stage of the power study consists of several variants characterised by different distances from the null. For experiments corresponding to the unconditional coverage hypothesis, these variants are implemented in a straightforward way through the manipulation of parameter p. This parameter is expressed as $p = \delta p^*$, where p^* denotes the VaR coverage assumed under the null. The considered levels of δ are: $\delta = 0.5,\ 0.75,\ 1.25,\ 1.5$. We report the results for $p^* = 5\%$.

Measuring the distance from the null in experiments that violate the independence property requires us to change the representation of the return process from the t-GARCH to the $Normal$-GARCH. The latter model allows for the explicit measurement of the scale of the violation of the independence property by the scale of the volatility clustering. The volatility clustering, in turn, is quantified by

the autocorrelation of the squared returns. Under the normality assumption and the additional assumption of the existence of the fourth moment, the first-order autocorrelation of the squared returns in the GARCH(1,1) process is expressed analytically by $\rho_1 = \alpha_1 + \alpha_1^2\beta_1/(1 - 2\alpha_1\beta_1 - \beta_1^2)$. The existence of the fourth moment can be verified by criterion $(\alpha_1 + \beta_1)^2 + 2\alpha_1^2 < 1$. If this criterion does not hold, the first-order autocorrelation behaves similarly to $\rho_1 = \alpha_1 + \beta_1/3$ (Ding & Granger, 1996), as has already been shown. Using the exact (when available) or approximate expressions for ρ_1, we design the experiments to target the ρ_1 levels: 0.1, 0.2, 0.3 and 0.4. These levels are obtained by adjusting parameter α_1 with ω_1 and β_1 fixed at the same levels as in the size study.[10]

The results of the first stage of the power study which examines the test effectiveness in discovering the incorrect unconditional VaR level are presented in Table 3. They clearly show that all of the tests were effective in completing the abovementioned task apart from one which proves unsuitable for this purpose, i.e. the spectral SD_{cc}^{AD} test. Its poor performance can be explained by the way it is constructed. It is specifically designed for the purpose of testing the independence property, and based on the autocorrelation function. Only the inclusion of parameter p allows this test to be classified as a conditional coverage test. However, it exhibits a feature common to all correlation-based tests, i.e. insensitivity to violations of unconditional moments.[11]

Out of the three tests performing well, one is the unconditional coverage test and the two others the conditional coverage tests. Thus, the two latter tests can potentially handle well both the violations of the unconditional VaR level and the violations of the independence property. According to our results, all three tests exhibit similar efficiency in detecting incorrect VaR levels. All of them show that the true VaR level amounting to 0.75 or 1.25 of the tested level cannot be effectively detected, while all tests become efficient if the true VaR reaches 0.5 or 1.5 of the tested value. Usually samples of 500 observations are sufficiently long, but the sample length of 1,000 gives the power of at least 0.8 in all the cases.

[10] Such experiments enable us to study the power as a function of the distribution parameter. In this way we improve the existing studies, which are based on arbitrarily chosen GARCH-model-based alternatives (single or multiple), without any information about the distance from the null. Although various alternatives that have more complex representations might perform better at describing real financial processes, we believe they are less suitable for the purpose of the power study.

[11] This feature is also exhibited by the Ljung-Box LB_{cc} test; however, we do not discuss it here – we excluded it from this part of the study for other reasons.

Table 3. Estimates of the power of VaR tests against incorrect unconditional coverage

Series length	LR_{uc}^{K}				$LR_{cc}^{M,gen}$			
	distance from the null measured by δ							
	0.5	0.75	1.25	1.5	0.5	0.75	1.25	1.5
250	0.566	0.173	0.163	0.404	0.444	0.119	0.110	0.297
500	0.875	0.307	0.210	0.624	0.736	0.195	0.183	0.547
750	0.959	0.391	0.336	0.833	0.913	0.299	0.253	0.734
1,000	0.992	0.514	0.392	0.896	0.971	0.363	0.320	0.839
1,250	0.998	0.608	0.476	0.951	0.992	0.447	0.382	0.918
1,500	1.000	0.670	0.545	0.978	0.998	0.541	0.460	0.957

(cont.)

Series length	$LR_{cc}^{geom,VaR}$				SD_{cc}^{AD}			
	distance from the null measured by δ							
	0.5	0.75	1.25	1.5	0.5	0.75	1.25	1.5
250	0.508	0.155	0.087	0.254	0.073	0.059	0.048	0.050
500	0.784	0.233	0.148	0.483	0.067	0.056	0.055	0.059
750	0.931	0.345	0.211	0.685	0.061	0.052	0.055	0.059
1,000	0.978	0.413	0.291	0.808	0.055	0.051	0.057	0.055
1,250	0.994	0.504	0.357	0.904	0.052	0.044	0.050	0.052
1,500	0.998	0.572	0.414	0.943	0.047	0.044	0.055	0.057

Note. All power results in the table assume testing a 5% VaR.
Source: author's calculations.

The results of examining the power against violations of the independence property, presented in Table 4, show that all tests admitted to this part of the study seem to cope well with this issue. For example, if the sample size is at least 500, all the tests can detect, with the power of over 0.7, the violations of the independence corresponding to the volatility clustering at the level of 0.2. This level corresponds to the level of 0.2 of the autocorrelation coefficient of the squared returns in the violation process. The Markov $LR_{cc}^{M,gen}$ and the geometric-VaR $LR_{cc}^{geom,VaR}$ tests exhibit outstanding performance with short samples. In their cases, the power of over 0.7 is attainable with the shortest examined samples (250 observations). The spectral SD_{cc}^{AD} test seems to require slightly longer samples, but in this case, the power of 0.7 is attainable starting from 500 observations.

Table 4. Estimates of the power of VaR tests against violations of independence property

Series length	LR_{uc}^{K}				$LR_{cc}^{geom,VaR}$				SD_{cc}^{AD}			
	distance from the null measured by ρ_1											
	0.1	0.2	0.3	0.4	0.5	0.75	1.25	1.5	0.5	0.75	1.25	1.5
250	0.285	0.551	0.715	0.799	0.287	0.594	0.764	0.849	0.282	0.476	0.555	0.579
500	0.410	0.766	0.907	0.959	0.433	0.826	0.948	0.985	0.447	0.726	0.810	0.823
750	0.548	0.902	0.978	0.994	0.584	0.943	0.993	0.999	0.579	0.868	0.922	0.927
1,000	0.624	0.951	0.994	0.999	0.679	0.980	0.999	1.000	0.655	0.928	0.966	0.967
1,250	0.718	0.981	0.999	1.000	0.780	0.994	1.000	1.000	0.741	0.965	0.985	0.987
1,500	0.791	0.992	1.000	1.000	0.836	0.999	1.000	1.000	0.793	0.981	0.993	0.994

Note. All power results in the table assume testing a 5% VaR.

Source: author's calculations.

6. Conclusions

VaR testing procedures are a part of the global banking supervisory system. Originally, these procedures were developed with the use of theoretical asymptotic distributions. However, the approach based on theoretical distributions has been questioned since 2010, when estimation risk was first considered in the context of the VaR testing framework. It was argued that estimation risk disturbed the convergence of the distributions of test statistics to their standard asymptotic distributions. Thus, the replacement of the asymptotic distributions with those obtained by means of resampling methods was proposed as a remedy to the aforementioned issue.

While agreeing with the argument that estimation risk disturbs the convergence to the asymptotic distributions, we raised the question of whether the proposed solution was the best possible. We recalled the results from previous studies which have also been confirmed by our research, demonstrating that the accuracy of the tests corrected with resampling methods was still not satisfactory. We also indicated that all previous studies were based on a simplified assumption that the VaR estimates were obtained by means of the fixed forecasting scheme. This assumption had a large impact on the results, as, by definition, this type of forecasting schemes generate the largest estimation risk. Therefore, we argued that these results could not be generalised.

To examine the ways of dealing with estimation risk, we waived the assumption of the fixed forecasting scheme and studied more realistic ones – the rolling and the recursive schemes. Our results altered the conclusions from the previous studies. We showed that under the more realistic forecasting schemes, the loss of accuracy due to the use of resampled distributions exceeded the loss of accuracy resulting from the occurrence of estimation risk. Therefore, for the sake of accuracy in VaR testing, it is

better to rely on asymptotic distributions and minimise the influence of estimation errors by a suitable forecasting scheme than to use resampling methods.

Our approach has two more advantages, which seem important in the practical applications of VaR testing. The methods we proposed are much more time-efficient and do not require designing simulations to conduct a VaR test. They are, therefore, optimal from the point of view of real business operations.

Apart from the above general conclusions about dealing with the issue of estimation risk, our contribution to the undertaken subject also provides a broad comparison of the VaR testing methods, belonging to six distinct classes. To the best of our knowledge, this paper is the first to present such a comparison involving estimation risk. We selected the VaR tests that fulfil three optimality postulates: the tests are well-sized in finite samples when based on asymptotic distributions, they are robust to estimation risk, and the most efficient in detecting incorrect VaR models. Two conditional coverage tests turned out to be most satisfactory in the light of the postulates listed above – the generalised Markov $LR_{cc}^{M,gen}$ and the geometric-VaR $LR_{cc}^{geom,VaR}$ tests. They are both accurately sized and exhibit similar, outstanding results in terms of their power. The $LR_{cc}^{M,gen}$ test prevailed in robustness to estimation risk, thus this procedure proved best.

Both the above-mentioned tests, being of the conditional coverage type, allow for the joint testing of the unconditional coverage and independence postulates. Another procedure appearing to perform exceptionally well in all the aspects was the spectral SD_{cc}^{AD} test. However, it is dedicated to testing the independence property solely and has no power against incorrect unconditional coverage. Therefore, it may be used only together with an unconditional coverage test. An unconditional coverage test that fulfils all our postulates is the LR_{uc}^{K} Kupiec test. Thus, the approach based on the Kupiec and spectral tests seems equivalent to the one based on the generalised Markov procedure. All these procedures may be effectively implemented with the use of asymptotic distributions, provided that VaR forecasting is conducted with a suitable forecasting scheme.

References

Anderson, T. W., & Darling, D. A. (1952). Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics*, *23*(2), 193–212. https://doi.org/10.1214/aoms/1177729437.

Basel Committee on Banking Supervision. (2017). *High-level summary of Basel III reforms*. Bank for International Settlements. https://www.bis.org/bcbs/publ/d424_hlsummary.pdf.

Berkowitz, J. (2001). Testing Density Forecasts, With Applications to Risk Management. *Journal of Business & Economic Statistics, 19*(4), 465–474. https://dx.doi.org/10.1198/07350010152596718.

Berkowitz, J., Christoffersen, P., & Pelletier, D. (2011). Evaluating Value-at-Risk Models with Desk-Level Data. *Management Science*, *57*(12), 2213–2227. https://doi.org/10.1287/mnsc.1080.0964.

Candelon, B., Colletaz, G., Hurlin, C., & Tokpavi, S. (2011). Backtesting Value-at-Risk: a GMM duration-based test. *Journal of Financial Econometrics*, *9*(2), 314–343. https://doi.org/10.1093/jjfinec/nbq025.

Christoffersen, P. (1998). Evaluating Interval Forecasts. *International Economic Review*, *39*(4), 841–862. https://doi.org/10.2307/2527341.

Christoffersen, P., & Pelletier, D. (2004). Backtesting Value-at-Risk: A Duration-Based Approach. *Journal of Financial Econometrics*, *2*(1), 84–108. https://doi.org/10.1093/jjfinec/nbh004.

Colletaz, G., Hurlin, C., & Pérignon, C. (2013). The Risk Map: A new tool for validating risk models. *Journal of Banking & Finance*, *37*(10), 3843–3854. https://doi.org/10.1016/j.jbankfin.2013.06.006.

Ding, Z., & Granger, C. W. J. (1996). Modeling Volatility Persistence of Speculative Returns: A New Approach. *Journal of Econometrics*, *73*(1), 185–215. http://dx.doi.org/10.1016/0304-4076(95)01737-2.

Dumitrescu, E.-I., Hurlin, Ch., & Pham, V. (2012). Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests (HAL Working Papers No. halshs-00671658). https://halshs.archives-ouvertes.fr/halshs-00671658/document.

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Jornual of Business & Economic Statisitics*, *22*(4), 367–381. https://doi.org/10.1198/073500104000000370.

Engle, R. F., & Russel, J. R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica*, *66*(5), 1127–1162. https://doi.org/10.2307/2999632.

Escanciano, J. C., & Olmo, J. (2010). Backtesting Parametric Value-at-Risk With Estimation Risk. *Journal of Business and Economic Statistics*, *28*(1), 36–51. https://doi.org/10.1198/jbes.2009.07063.

Escanciano, J. C., & Olmo, J. (2011). Robust Backtesting Tests for Value-at-risk Models. *Journal of Financial Econometrics*, *9*(1), 132–161. https://doi.org/10.1093/jjfinec/nbq021.

Gordy, M. B., & McNeil, A. J. (2018). *Spectral backtests of forecast distributions with application to risk management* (Finance and Economics Discussion Series No. 2018-021). https://doi.org/10.17016/FEDS.2018.021.

Haas, M. (2005). Improved duration-based backtesting of value-at-risk. *Journal of Risk*, *8*(2), 17–38. https://doi.org/10.21314/JOR.2006.128.

Hurlin, Ch., & Tokpavi, S. (2006). Backtesting value-at-risk accuracy: a simple new test. *Journal of Risk*, *9*(2), 19–37. https://doi.org/10.21314/JOR.2007.148.

Kratz, M., Lok, Y. H., & McNeil, A. J. (2018). Multinomial VaR backtests: A simple implicit approach to backtesting expected shortfall. *Journal of Banking & Finance*, *88*, 393–407. https://doi.org/10.1016/j.jbankfin.2018.01.002.

Krämer, W., & Wied, D. (2015). A simple and focused backtest of value at risk. *Economics Letters*, *137*, 29–31, https://doi.org/10.1016/j.econlet.2015.10.028.

Kupiec, P. H. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal of Derivatives, 3*(2), 73–84. https://doi.org/10.3905/jod.1995.407942.

Leccadito, A., Boffelli, S., & Urga, G. (2014). Evaluating the Accuracy of Value-at-Risk Forecasts: New Multilevel Tests. *International Journal of Forecasting*, *30*(2), 206–216. https://doi.org/10.1016/j.ijforecast.2013.07.014.

Ljung, G. M., & Box, G. E. P. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika, 65*(2), 297–303. https://doi.org/10.1093/biomet/65.2.297.

Małecka, M. (2016). Spectral VaR test statistical properties. In M. Papież & S. Śmiech (Eds.), *The 10th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings* (pp. 102–109). Foundation of the Cracow University of Economics.

Pajhede, T. (2017). Backtesting Value-at-Risk: A Generalized Markov Test. *Journal of Forecasting*, *36*(5), 597–613. https://doi.org/10.1002/for.2456.

Pelletier, D., & Wei, W. (2016). The geometric-VaR backtesting method. *Journal of Financial Econometrics*, *14*(4), 725–745. https://doi.org/10.1093/jjfinec/nbv015.

Wied, D., Weiß, G. N. F., & Ziggel, D. (2016). Evaluating Value-at-Risk forecasts: a new set of multivariate backtests. *Journal of Banking & Finance*, *72*, 121–132. https://doi.org/10.1016/j.jbankfin.2016.07.014.

Ziggel, D., Berens, T., Weiß, G. N. F., & Wied, D. (2014). A new set of improved value-at-risk backtests. *Journal of Banking & Finance*, *48*, 29–41. https://doi.org/10.1016/j.jbankfin.2014.07.005.

Wiadomości Statystyczne. The Polish Statistician, 2022, vol. 67, 10, 28–47 DOI: 10.5604/01.3001.0016.0658
Statystyka w praktyce / Statistics in practice
Zgłoszony/submitted: 23.05.2022, poprawiony/revised: 8.10.2022, zaakceptowany/accepted: 14.10.2022

Ocena stopnia zagrożenia ubóstwem w ramach podejścia możliwości Sena z zastosowaniem modelu MIMIC

Beata Kraszewska[a]

Streszczenie. Ubóstwo jest problemem nie tylko dotkniętej nim jednostki, lecz także społeczeństwa. Utrudnia ono lub nawet uniemożliwia wzrost gospodarczy i rozwój społeczny kraju. Dlatego bardzo duże znaczenie mają działania, które mogą spowodować redukcję zasięgu ubóstwa, jak również udzielanie bieżącej pomocy jednostkom (gospodarstwom domowym) żyjącym w sferze ubóstwa. Głównym celem badania omawianego w artykule jest dokonanie wielowymiarowej analizy zagrożenia ubóstwem w ramach podejścia możliwości (ang. *capability approach*) opracowanego przez Amartayę Sena. Dodatkowym celem jest określenie determinant ubóstwa, które wzmacniają lub osłabiają zagrożenie ubóstwem, oraz zestawienie wyników oceny stopnia zagrożenia ubóstwem uzyskanych w ujęciu wielowymiarowym i jednowymiarowym. Badanie przeprowadzono dla Polski (m.in. w układzie regionalnym), na podstawie danych z Europejskiego Badania Warunków Życia Ludności (EU-SILC) zrealizowanego przez GUS w 2018 r. Operacjonalizacji pomiaru dokonano przy użyciu szczególnego wariantu modelu równań strukturalnych (Structural Equations Modeling – SEM) – modelu wielu wskaźników i wielu przyczyn (Multiple Indicators and Multiple Causes – MIMIC). Wykorzystano w nim symptomy deprywacji odzwierciedlające stopień, w jakim badane jednostki nie były w stanie zaspokoić swoich potrzeb, i uwzględniono czynniki zwiększające lub zmniejszające ryzyko ubóstwa.

Wyniki przeprowadzonych szacunków uwidoczniły, że na zmniejszenie stopnia zagrożenia ubóstwem najsilniej wpływa poziom wykształcenia, natomiast za jego zwiększenie odpowiada przede wszystkim niski status aktywności na rynku pracy – bycie rencistą lub osobą bezrobotną. Najbardziej zagrożone ubóstwem w ujęciu wielowymiarowym są: makroregion wschodni, obszary wiejskie, gospodarstwa domowe rencistów i gospodarstwa nierodzinne wieloosobowe. Wyniki wielowymiarowej oceny stopnia zagrożenia ubóstwem porównano z pomiarem ubóstwa w ujęciu jednowymiarowym. Wskazano na różnice pomiędzy rankingami zagrożenia ubóstwem według makroregionów, wielkości miejscowości zamieszkania, źródła utrzymania gospodarstw domowych oraz ich typów w ujęciu wielowymiarowym i jednowymiarowym. Porównanie to pokazuje, że ujęcie jednowymiarowe jest niewystarczające do identyfikacji osób ubogich.

Słowa kluczowe: zagrożenie ubóstwem, podejście możliwości, model MIMIC

JEL: I32, C3

[a] Szkoła Główna Handlowa w Warszawie, Kolegium Analiz Ekonomicznych; Urząd Statystyczny w Poznaniu, Oddział w Kaliszu, Polska / SGH Warsaw School od Economics, Collegium of Economic Analysis; Statistical Office in Poznan, Branch Office in Kalisz, Poland. ORCID: https://orcid.org/0000-0002-2410-4133.

© Beata Kraszewska
Artykuł udostępniony na licencji CC BY-SA 4.0 / Article available under the CC BY-SA 4.0 licence

Poverty risk estimation on the basis of Sen's capability approach with the application of the MIMIC model

Abstract. Poverty is a problem which concerns not only the affected individual, but also a whole society. Poverty tends to hinder or even prevent the economic growth and social development of a country. Therefore, undertaking action which could reduce poverty is of great importance, as is the provision of ongoing help to impoverished individuals (households). The main aim of the research presented in the paper is a multidimensional analysis of poverty risk in the framework of capability approach developed by Amartya Sen. The article also aims at defining the determinants of poverty which increase or decrease the risk of its occurrence, and comparing the poverty risk estimations obtained through the adoption of the multidimensional and unidimensional approach. The research concerning Poland (e.g. by regions) used data from the European Union Statistics on Income and Living Conditions (EU-SILC) survey, carried out by Statistics Poland in 2018. The measurement was operationalised by means of a special variant of structural equations modelling (SEM) – the multiple indicators and multiple causes model (MIMIC). The MIMIC model used deprivation symptoms that reflected the degree to which respondents were unable to satisfy their needs. The model also took into account the factors that increase or reduce the risk of poverty.

The results of the estimations demonstrate that the level of education has the most significant impact on reducing the risk of poverty, while a low activity on the job market, i.e. being a pensioner or unemployed, increases the odds of becoming poor to the greatest extent. The regions most prone to poverty in multidimensional terms are: the Eastern macro-region, rural areas, pensioners' households and non-family multi-person households. The aforementioned results have been compared with the measurement of poverty according to the unidimensional approach. The study shows the differences between the rankings of poverty risk by macro-region, size of the place of residence, source of household income and by household type from obtained through multidimensional and unidimensional approaches. This comparison demonstrates that the unidimensional approach is insufficient for the identification of the poor.

Keywords: poverty risk, capability approach, MIMIC model

1. Wprowadzenie

Żaden region na świecie nie jest wolny od ubóstwa. Problem ten występuje także w Unii Europejskiej, dlatego jej kraje członkowskie, w tym Polska, zobowiązały się do opracowania politycznych strategii ograniczenia ubóstwa i wykluczenia społecznego. W strategii *Europa 2020* określono, że liczba osób zagrożonych ubóstwem w Polsce powinna zmniejszyć się do końca 2020 r. o 1,5 mln (European Commission [EC], 2010). Także w *Agendzie na rzecz zrównoważonego rozwoju 2030*, sformułowanej przez Zgromadzenie Ogólne Organizacji Narodów Zjednoczonych (ONZ) w 2015 r., a przyjętej jednogłośnie przez wszystkie kraje członkowskie, pierwszym z 17 wyznaczonych celów rozwoju jest wyeliminowanie ubóstwa we wszystkich jego przejawach na całym świecie do 2030 r. (EC, 2019; ONZ, 2015).

Pomiar ubóstwa najczęściej jest wykonywany w ujęciu jednowymiarowym, które skupia się wyłącznie na bieżących dochodach lub wydatkach, co często prowadzi do nie w pełni poprawnej identyfikacji osób ubogich.

Celem badania omawianego w artykule jest dokonanie wielowymiarowej analizy zagrożenia ubóstwem w ramach podejścia możliwości (ang. *capability approach)* opracowanego przez Amartayę Sena. Dodatkowym celem jest określenie determinant ubóstwa, które wzmacniają lub osłabiają zagrożenie ubóstwem, oraz zestawienie wyników oceny zagrożenia ubóstwem uzyskanych w ujęciu wielowymiarowym i jednowymiarowym.

2. Metoda badania

W analizie ubóstwa badaną jednostką jest osoba, a wszystkie miary ubóstwa oblicza się dla populacji osób. Jednak identyfikacja osób ubogich odbywa się na podstawie wskazania gospodarstw domowych, których członkami są te osoby. W sytuacji, w której symptomy (wskaźniki cząstkowe) oraz determinanty ubóstwa nie dotyczą osób, ale ich gospodarstw domowych, każdej osobie w gospodarstwie przypisuje się takie wartości wskaźników, jakie cechują całe gospodarstwo.

Dla potrzeb omawianego badania przyjęto ekonomiczną definicję ubóstwa (Panek i Zwierzchowski, 2021, s. 11). Zgodnie z nią ubóstwo występuje wtedy, gdy gospodarstwo domowe nie dysponuje środkami finansowymi – zarówno pieniężnymi, w postaci dochodów bieżących i dochodów z poprzednich okresów (oszczędności), jak i w formie nagromadzonych zasobów materialnych – pozwalającymi na zaspokojenie jego podstawowych potrzeb na minimalnym, akceptowalnym poziomie w danym kraju.

Wielowymiarową analizę zagrożenia ubóstwem przeprowadzono dla Polski w ujęciu zarówno ogólnokrajowym, jak i regionalnym. Wykonano ją także w przekrojach społeczno-demograficznych gospodarstw domowych, których członkami są osoby ubogie. Wykorzystano dane pochodzące z Europejskiego Badania Warunków Życia Ludności (European Union Statistics on Income and Living Conditions – EU-SILC) zrealizowanego przez Główny Urząd Statystyczny (GUS) w 2018 r. Kategorią dochodów przyjętą w EU-SILC jest roczny ekwiwalentny dochód do dyspozycji gospodarstwa domowego. Otrzymuje się go w wyniku podzielenia kwoty dochodów do dyspozycji gospodarstw domowych przez odpowiadające im skale ekwiwalentności. W badaniu omawianym w artykule zastosowano zmodyfikowaną skalę OECD: pierwszej osobie dorosłej w gospodarstwie przypisuje się skalę równą 1, każdej następnej osobie dorosłej – 0,5, a dziecku – 0,3.

Stopień zagrożenia ubóstwem oceniano za pomocą wskaźnika stopnia zagrożenia ubóstwem (Risk of Poverty Index – RPI)[1], skonstruowanego na podstawie zmody-

[1] Zasady konstrukcji wskaźnika RPI opisano w dalszej części artykułu; patrz wzory (1) i (2) .

fikowanego wskaźnika stopnia zagrożenia wykluczeniem społecznym (Social Exclusion Index – SEI; Panek i Zwierzchowski, 2022). Wskaźnik RPI umożliwia porównanie stopnia zagrożenia ubóstwem według jednostek terytorialnych, wielkości miejscowości zamieszkania oraz grup typologicznych gospodarstw domowych. Pozwala także na porównanie stopnia zagrożenia ubóstwem w czasie.

Do operacjonalizacji pomiaru ubóstwa w ramach podejścia możliwości Sena zastosowano szczególny przypadek modelu równań strukturalnych (Structural Equations Modeling – SEM) – model wielu wskaźników i wielu przyczyn (Multiple Indicators and Multiple Causes – MIMIC; Bollen, 1989; Konarski, 2014; Krishnakumar, 2008; Krishnakumar i Ballon, 2008; Muthen i in., 1994). Model MIMIC został opracowany przez Hausera i Goldbergera (1971), a następnie spopularyzowany przez Jöreskoga i Goldbergera (1975), którzy nadali mu aktualnie funkcjonującą nazwę i przedstawili jego szczegółowe założenia jako szczególnego przypadku modelu SEM. Model MIMIC składa się z podmodeli pomiarowego i strukturalnego (Hair jr. i in., 2014). W równaniach podmodelu pomiarowego rozpatrywana jest relacja między nieobserwowalnymi możliwościami oraz obserwowalnymi symptomami ubóstwa. Natomiast równania podmodelu strukturalnego określają wpływ obserwowalnej determinanty ubóstwa (cech osób i ich gospodarstw domowych) na kształtowanie się nieobserwowalnych możliwości.

Zaproponowana operacjonalizacja pomiaru ubóstwa w ramach podejścia możliwości, z zastosowaniem modelu MIMIC, pozwala na pomiar stopnia zagrożenia ubóstwem z uwzględnieniem zbioru rzeczywistych sposobów funkcjonowania badanych osób, odzwierciedlających deprywację ich możliwości. Ponadto model ten umożliwia ocenę wpływu zewnętrznych determinant ubóstwa – które wzmacniają lub osłabiają zagrożenie ubóstwem –na możliwości badanych osób. Pozwala także na wskazanie, które z symptomów ubóstwa najsilniej odzwierciedlają stopień zagrożenia ubóstwem.

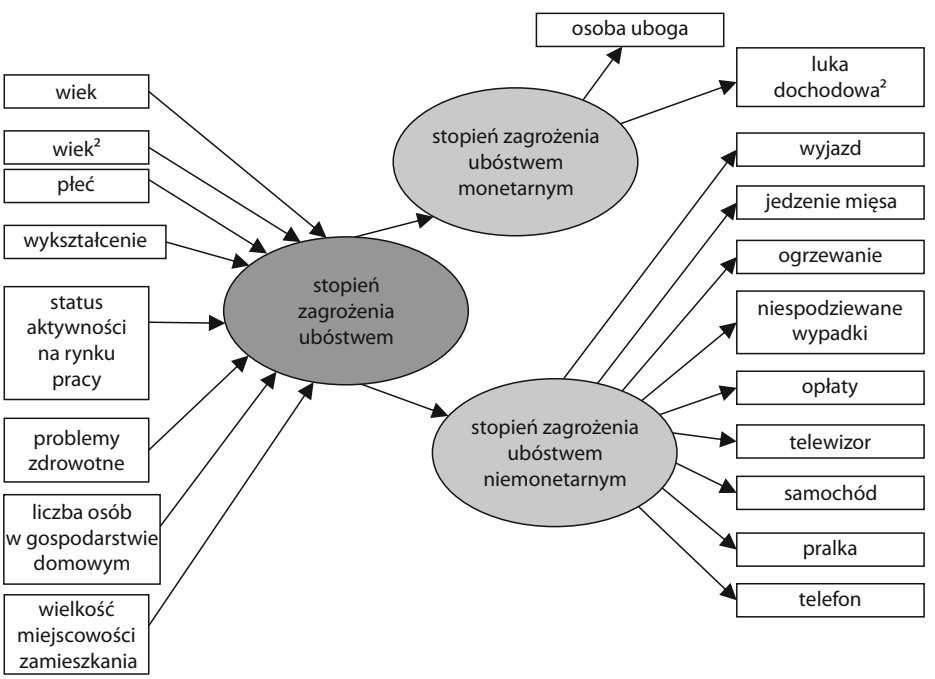
Dopasowanie modeli do analizowanej (empirycznej) macierzy kowariancji oceniano w dwóch krokach. W pierwszym sprawdzono ogólne dopasowanie modeli za pomocą:

- wskaźnika relatywnego dopasowania modelu (Confirmatory Fit Index – CFI); jest on znormalizowany w przedziale [0; 1]: im wyższa wartość wskaźnika, tym lepsze dopasowanie modelu. O bardzo dobrym dopasowaniu modelu świadczą wartości powyżej 0,95;
- wskaźnika absolutnego dopasowania modelu (Root Mean Square Error of Approximation – RMSEA): im niższa wartość wskaźnika, tym lepsze dopasowanie modelu. Jako górną wartość świadczącą o bardzo dobrym dopasowaniu modelu przyjmuje się 0,08.

W drugim kroku sprawdzono lokalne dopasowanie modeli na podstawie oceny istotności poszczególnych parametrów (Konarski, 2014; McDonald i Ho, 2002).

Praca badawcza była realizowana etapowo. Najpierw wyróżniono dwie domeny w kategorii ubóstwa: ubóstwo monetarne i ubóstwo niemonetarne (deprywację materialną) jako punkt wyjścia do budowy modelu MIMIC. Następnie do każdej domeny ubóstwa przypisano symptomy i determinanty ubóstwa. W wymiarze monetarnym za symptomy ubóstwa uznano, zgodnie z zaleceniami UE sformułowanymi w strategii *Europa 2020*, dwa podstawowe wskaźniki ubóstwa: status przynależności do sfery ubóstwa oraz lukę dochodową (GUS, 2015; Panek i Zwierzchowski, 2021). W wymiarze niemonetarnym (deprywacji materialnej) oparto się na rozwiązaniu stosowanym przez UE w analizach sfery ubóstwa, tj. wykorzystaniu dziewięciu symptomów deprywacji materialnej (EC, 2011; Eurostat, b.r.; GUS, 2019). Zależności te zostały przedstawione w postaci schematu modelu MIMIC dla zagrożenia ubóstwem monetarnym, niemonetarnym i w obu wymiarach łącznie (dalej także jako: wielowymiarowe zagrożenie ubóstwem). Definicje i sposób konstrukcji wskaźników cząstkowych – symptomów ubóstwa monetarnego i niemonetarnego – przedstawiono w zestawieniu 1.

Schemat modelu MIMIC dla zagrożenia ubóstwem monetarnym i niemonetarnym oraz łącznie dla wielowymiarowego zagrożenia ubóstwem



Źródło: opracowanie własne.

Zestawienie 1. Cząstkowe wskaźniki ubóstwa monetarnego i niemonetarnego

Nazwa	Definicja
Ubóstwo monetarne	
Status przynależności do sfery ubóstwa (osoba uboga)	Osoba należąca do gospodarstwa domowego o ekwiwalentnym dochodzie do dyspozycji niższym niż granica ubóstwa, wynosząca 60% mediany ekwiwalentnego dochodu do dyspozycji
Kwadrat luki dochodowej (luka dochodowa^2)	Dystans ekwiwalentnego dochodu gospodarstwa domowego badanej osoby od granicy ubóstwa; jest określany w zł^2; luka dochodowa osób nieubogich jest równa 0
Ubóstwo niemonetarne	
Brak możliwości wyjazdu na tygodniowy wypoczynek przynajmniej raz w roku (wyjazd)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych nie ma możliwości opłacenia tygodniowego wyjazdu wszystkich członków gospodarstwa na wypoczynek raz w roku
Brak środków finansowych na zakup mięsa albo jego wegetariańskiego odpowiednika co drugi dzień (jedzenie mięsa)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych nie ma możliwości jedzenia mięsa, w tym ryb (albo ich wegetariańskiego odpowiednika), co drugi dzień
Brak środków finansowych na ogrzewanie mieszkania (ogrzewanie)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych nie ma możliwości ogrzewania mieszkania stosownie do potrzeb
Brak środków finansowych na pokrycie niespodziewanego wydatku (niespodziewane wydatki)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych nie ma możliwości pokrycia z własnych środków niespodziewanego wydatku w wysokości 810 zł
Brak środków finansowych na terminowe uiszczanie opłat związanych z mieszkaniem (opłaty)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych ma w ostatnim roku zaległości w opłatach związanych z mieszkaniem i spłatach kredytu na zakup lub wykup mieszkania
Brak telewizora z powodów finansowych (telewizor)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych nie posiada telewizora do odbioru w kolorze
Brak samochodu z powodów finansowych (samochód)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych nie posiada samochodu
Brak pralki z powodów finansowych (pralka)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych nie posiada pralki automatycznej ani wirnikowej
Brak telefonu z powodów finansowych (telefon)	Osoba należąca do gospodarstwa domowego, które ze względów finansowych nie posiada telefonu stacjonarnego ani komórkowego

Źródło: opracowanie własne na podstawie: EC (2011); Eurostat (b.r.); GUS (2015, 2019); Panek i Zwierzchowski (2021).

Jako determinanty ryzyka ubóstwa w wymiarze zarówno monetarnym, jak i niemonetarnym przyjęto dwie grupy zmiennych: cechy członków gospodarstwa domowego oraz cechy całego gospodarstwa (GUS i Urząd Statystyczny w Łodzi, 2013; Jalan i Ravallion, 1998; Kurowska, 2008; Nawawi i in., 2019; Radziukiewicz, 2006; Sekhampu, 2013). Determinanty ubóstwa monetarnego i niemonetarnego przedstawiono w zestawieniu 2.

Zestawienie 2. Determinanty ubóstwa monetarnego i niemonetarnego

Cechy	Sposób zdefiniowania determinanty
Członkowie gospodarstwa domowego	
Wiek	Wiek w latach
Płeć	1 – kobieta 0 – mężczyzna
Wykształcenie	1 – podstawowe lub niższe 2 – gimnazjalne 3 – zasadnicze zawodowe lub średnie (liceum ogólnokształcące, liceum profilowane, technikum, policealne, pomaturalne) 4 – średnie po kolegium nauczycielskim lub języków obcych lub wyższe (z tytułem inżyniera, licencjata, magistra, lekarza lub równorzędnym, wyższe po studiach podyplomowych, wyższe ze stopniem naukowym co najmniej doktora)
Status aktywności na rynku pracy	1 – zatrudniony lub emeryt 2 – uczeń lub student 3 – bezrobotny 4 – rencista 5 – nieaktywny zawodowo z innych przyczyn
Problemy zdrowotne	1 – poważne ograniczenie wykonywania czynności 2 – niezbyt poważne ograniczenie 3 – brak ograniczeń
Gospodarstwo domowe	
Liczba osób w gospodarstwie domowym	Liczba osób
Wielkość miejscowości zamieszkania	1 – bardzo duże miasto 2 – duże miasto 3 – średnie miasto 4 – małe miasto 5 – bardzo małe miasto 6 – wieś

Uwaga. Bardzo duże miasto – powyżej 500 tys. mieszkańców, duże miasto – 200–500 tys. mieszkańców, średnie miasto – 100–200 tys. mieszkańców, małe miasto – 20–100 tys. mieszkańców, bardzo małe miasto – poniżej 20 tys. mieszkańców.

Źródło: opracowanie własne na podstawie: GUS i Urząd Statystyczny w Łodzi (2013); Jalan i Ravallion (1998); Kurowska (2008); Nawawi i in. (2019); Radziukiewicz (2006); Sekhampu (2013).

Na trzecim etapie dla każdej domeny ubóstwa oszacowano parametry modelu MIMIC obrazujące zależności pomiędzy zdeprywowanymi możliwościami badanych jednostek (czyli zmiennymi ukrytymi mierzącymi zagrożenie ubóstwem) a obserwowalnymi determinantami i symptomami ubóstwa. Pozwoliło to – zgodnie z koncepcją Sena (1979, 1985, 1992) – na ocenę stopnia zagrożenia ubóstwem poprzez oszacowanie wartości zdeprywowanych możliwości jednostek przy jednoczesnym określeniu wpływu czynników oddziałujących na zwiększenie (stymulanty) albo zmniejszenie (destymulanty) stopnia zagrożenia ubóstwem. Parametry modelu MIMIC oszacowano metodą największej wiarygodności (MNW; ang. *maximum likelihood),* z uwzględnieniem braków danych, przy użyciu pakietu STATA16. Meto-

da ta umożliwia analizę całego zbioru danych dla wszystkich badanych jednostek, nawet jeśli w przypadku niektórych jednostek nie wszystkie wartości zmiennych są znane.

Na kolejnym etapie dla każdego członka badanych gospodarstw domowych oszacowano wartości zmiennej ukrytej reprezentującej zdeprywowane możliwości oddzielnie w domenie ubóstwa monetarnego i niemonetarnego. Oszacowane wartości zdeprywowanych możliwości poszczególnych osób zostały następnie wykorzystane do oszacowania grupowego wskaźnika stopnia zagrożenia ubóstwem RPI, oddzielnie dla wymiaru monetarnego i niemonetarnego:

$$RPI_{hi} = \frac{x_{ehi} - x^{*}_{h\,0}}{x^{**}_{h\,100} - x^{*}_{h\,0}}, \tag{1}$$

gdzie:

x_{ehi} – empiryczna wartość zmiennej ukrytej (zdeprywowanych możliwości) dla i-tej osoby w h-tej domenie ubóstwa,

$x^{**}_{h\,100}$ – wartość krytyczna maksimum zmiennej ukrytej (zdeprywowanych możliwości) w h-tej domenie ubóstwa (najwyższe możliwe do osiągnięcia zagrożenie ubóstwem),

$x^{*}_{h\,0}$ – wartość krytyczna minimum zmiennej ukrytej (zdeprywowanych możliwości) w h-tej domenie ubóstwa (brak zagrożenia ubóstwem).

Jako maksymalną wartość zmiennej ukrytej odzwierciedlającej najwyższe możliwe zagrożenie ubóstwem przyjęto wartość tej zmiennej dla fikcyjnej osoby, która posiada najmniej pożądane możliwe do osiągnięcia wartości wszystkich symptomów i determinant zagrożenia ubóstwem. Jako minimalną wartość zmiennej ukrytej odzwierciedlającej brak zagrożenia ubóstwem przyjęto wartość tej zmiennej dla fikcyjnej osoby, która posiada najbardziej pożądane możliwe do osiągnięcia wartości wszystkich symptomów zagrożenia ubóstwem i jego determinant. Krytyczne wartości zdeprywowanych możliwości zostały określone w sposób pozwalający na porównanie stopnia zagrożenia ubóstwem w obu domenach ubóstwa dla różnych grup społeczno-ekonomicznych i dla różnych jednostek terytorialnych, a także porównanie w czasie. Grupowy RPI przyjmuje wartości z przedziału [0; 1]. Im wyższa wartość wskaźnika, tym wyższy stopień zagrożenia ubóstwem. Wartość 1 oznacza najwyższy możliwy do osiągnięcia stopień zagrożenia ubóstwem, natomiast wartość 0 wskazuje na całkowity brak zagrożenia ubóstwem.

Na ostatnim etapie oszacowano stopień zagrożenia ubóstwem dla wymiarów monetarnego i niemonetarnego łącznie. Dokonano agregacji wartości grupowych RPI w jeden wskaźnik syntetyczny. Parametry modelu MIMIC łącznie dla obu wymiarów

ubóstwa oszacowano z uwzględnieniem tych samych determinant co w submodelach zagrożenia ubóstwem monetarnym i niemonetarnym. Jako symptomy ubóstwa przyjęto grupowe wskaźniki ubóstwa: wskaźnik stopnia zagrożenia ubóstwem monetarnym (RPI^{um}) oraz wskaźnik stopnia zagrożenia deprywacją materialną (RPI^{dm}).

Wykorzystując oszacowane wartości syntetycznej zmiennej zdeprywowanych możliwości (czyli stopnia zagrożenia ubóstwem), obliczono wartości syntetycznego wskaźnika stopnia zagrożenia ubóstwem dla każdej badanej osoby:

$$RPI_i = \frac{x_{ei} - x_0^*}{x_{100}^{**} - x_0^*}, \tag{2}$$

gdzie:

x_{ei} – empiryczna wartość zmiennej ukrytej zdeprywowanych możliwości dla i-tej osoby w obu domenach ubóstwa łącznie,

x_{100}^{**} – wartość maksimum zmiennej ukrytej (zdeprywowanych możliwości) w obu domenach ubóstwa łącznie (najwyższe możliwe do osiągnięcia zagrożenie ubóstwem),

x_0^* – wartość krytyczna minimum zmiennej ukrytej (zdeprywowanych możliwości) w obu domenach ubóstwa łącznie (brak zagrożenia ubóstwem).

Jako maksymalną wartość zmiennej ukrytej odzwierciedlającej najwyższe zagrożenie ubóstwem przyjęto wartość tej zmiennej dla osoby fikcyjnej równą 1, a jako minimalną wartość zmiennej ukrytej odzwierciedlającej brak zagrożenia ubóstwem – wartość tej zmiennej dla osoby fikcyjnej równą 0.

3. Wyniki

3.1. Ocena stopnia zagrożenia ubóstwem monetarnym i niemonetarnym oraz łącznie (wielowymiarowo)

Oszacowania parametrów modeli MIMIC, obrazujących zależności pomiędzy zdeprywowanymi możliwościami badanych jednostek (zmiennymi ukrytymi mierzącymi zagrożenie ubóstwem) a obserwowalnymi determinantami i symptomami ubóstwa, przeprowadzone dla wymiaru monetarnego i niemonetarnego oraz obu wymiarów łącznie, przedstawiono w tabl. 1.

Tabl. 1. Oszacowanie parametrów modeli MIMIC dla zagrożenia ubóstwem monetarnym i niemonetarnym oraz wielowymiarowego zagrożenia ubóstwem

Zmienne	Oszacowania parametrów	Błąd standardowy	z	$P > z$	Przedział ufności 95%	
					od	do
Zagrożenie ubóstwem monetarnym						
Submodel strukturalny						
Wiek	0,024	0,007	3,42	0	0,010	0,037
Wiek^2	–0,186	0,007	–26,47	0	–0,199	–0,172
Płeć	0,008	0,005	1,70	0,09	–0,001	0,018
Wykształcenie	–0,201	0,007	–27,62	0	–0,216	–0,187
Status aktywności na rynku pracy	0,123	0,006	21,28	0	0,112	0,135
Problemy zdrowotne	–0,049	0,006	–8,28	0	–0,060	–0,037
Liczba osób w gospodarstwie domowym	–0,108	0,006	–18,56	0	–0,120	–0,097
Wielkość miejscowości zamieszkania	0,108	0,005	20,84	0	0,098	0,118
Submodel pomiarowy						
Osoba uboga	1[a]	.	.	.	.	.
Luka dochodowa^2	0,545	0,019	28,93	0	0,508	0,582
Zagrożenie ubóstwem niemonetarnym						
Submodel strukturalny						
Wiek	0,057	0,005	11,34	0	0,047	0,067
Wiek^2	–0,152	0,005	–29,46	0	–0,162	–0,142
Płeć	0,012	0,004	3,31	0	0,005	0,019
Wykształcenie	–0,213	0,006	–37,97	0	–0,224	–0,202
Status aktywności na rynku pracy	0,098	0,004	22,79	0	0,089	0,106
Problemy zdrowotne	–0,109	0,004	–24,49	0	–0,117	–0,100
Liczba osób w gospodarstwie domowym	–0,073	0,004	–17,20	0	–0,082	–0,065
Wielkość miejscowości zamieszkania	0,039	0,004	10,08	0	0,032	0,047
Submodel pomiarowy						
Wyjazd	1[a]	.	.	.	.	.
Jedzenie mięsa	0,768	0,012	63,11	0	0,745	0,792
Ogrzewanie	0,758	0,012	62,49	0	0,734	0,781
Niespodziewane wydatki	1,059	0,012	91,51	0	1,036	1,082
Opłaty	0,446	0,010	43,40	0	0,426	0,467
Telewizor	0,109	0,010	11,37	0	0,090	0,127
Samochód	0,582	0,011	54,71	0	0,561	0,602
Pralka	0,173	0,010	17,87	0	0,154	0,191
Telefon	0,134	0,010	14,05	0	0,116	0,153

a Skale zmiennej nieobserwowalnej nie są znane, dlatego konieczne było zastosowanie reguły skali, tzn. założenia, że skale zmiennych ukrytych muszą zostać w modelu ustalone, ponieważ w przeciwnym razie model nie będzie identyfikowalny. Skale można ustalić poprzez ograniczenie jednego ładunku czynnikowego każdej zmiennej ukrytej do jedności (czyli przyjmując, że jego wartość wynosi 1).

Tabl. 1. Oszacowanie parametrów modeli MIMIC dla zagrożenia ubóstwem monetarnym i niemonetarnym oraz wielowymiarowego zagrożenia ubóstwem (dok.)

Zmienne	Oszacowania parametrów	Błąd standardowy	z	$P > z$	Przedział ufności 95%	
					od	do
Wielowymiarowe zagrożenie ubóstwem						
Submodel strukturalny						
Wiek	0,017	0,001	16,30	0	0,015	0,019
Wiek^2	–0,049	0,001	–35,82	0	–0,052	–0,046
Płeć	0,004	0,001	5,05	0	0,002	0,005
Wykształcenie	–0,066	0,002	–42,35	0	–0,070	–0,063
Status aktywności na rynku pracy	0,032	0,001	30,40	0	0,030	0,034
Problemy zdrowotne	–0,032	0,001	–32,11	0	–0,034	–0,030
Liczba osób w gospodarstwie domowym	–0,025	0,001	–24,88	0	–0,027	–0,023
Wielkość miejscowości zamieszkania	0,015	0,001	17,67	0	0,014	0,017
Submodel pomiarowy						
RPI^{um}	1[a]	.	.	.	.	.
RPI^{dm}	0,778	0,016	49,400	0	0,747	0,809

a Skale zmiennej nieobserwowalnej nie są znane, dlatego konieczne było zastosowanie reguły skali, tzn. założenia, że skale zmiennych ukrytych muszą zostać w modelu ustalone, ponieważ w przeciwnym razie model nie będzie identyfikowalny. Skale można ustalić poprzez ograniczenie jednego ładunku czynnikowego każdej zmiennej ukrytej do jedności (czyli przyjmując, że jego wartość wynosi 1).

Uwaga. z – wartość statystyki o asymptotycznym rozkładzie normalnym. $P > z$ – empiryczny poziom istotności, przy którym odrzuca się hipotezę zerową o nieistotności parametru.

Źródło: opracowanie własne na podstawie danych z EU-SILC.

Wskaźniki ogólnego dopasowania modeli do analizowanej (empirycznej) macierzy kowariancji wskazują na ich dobre dopasowanie. Otrzymano następujące wartości miar:

- CFI: dla zagrożenia ubóstwem monetarnym – 0,975, dla zagrożenia ubóstwem niemonetarnym – 0,760, dla wielowymiarowego zagrożenia ubóstwem – 0,966;
- RMSEA: dla zagrożenia ubóstwem monetarnym – 0,037, dla zagrożenia ubóstwem niemonetarnym – 0,051, dla wielowymiarowego zagrożenia ubóstwem – 0,043.

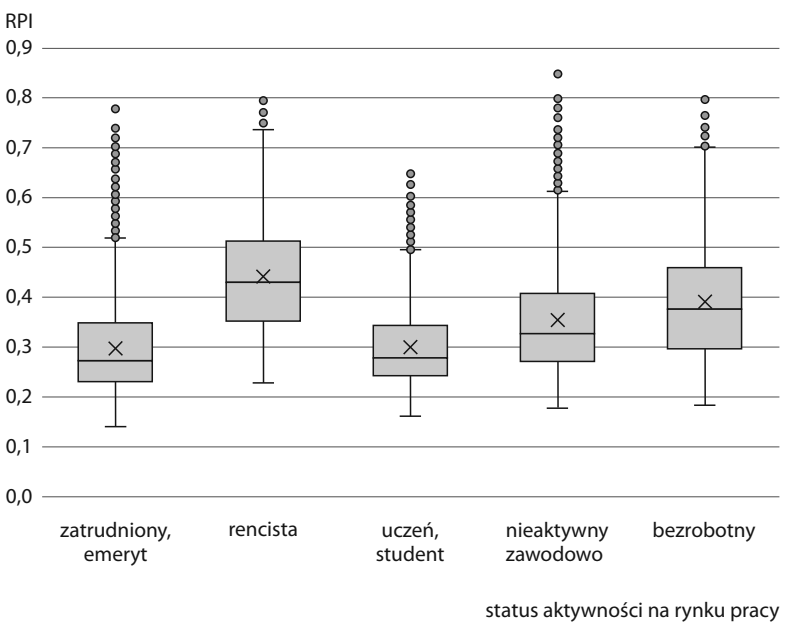
Symptomy zagrożenia ubóstwem we wszystkich modelach są istotne statystycznie, co oznacza, że zdeprywowane możliwości są dobrze odzwierciedlane przez proponowane wskaźniki. Podobnie wykorzystane w modelach determinanty zagrożenia ubóstwem, czyli cechy osób i ich gospodarstw domowych wpływające na wzmocnienie albo osłabienie stopnia zagrożenia ubóstwem, okazały się istotne statystycznie (z wyjątkiem płci w modelu zagrożenia ubóstwem monetarnym), co świadczy o ich prawidłowym doborze.

3.2. Siła oddziaływania determinant ubóstwa

Najsilniejszy wpływ zarówno na stopień zagrożenia ubóstwem w wyróżnionych domenach, jak i na wartość syntetycznego wskaźnika zagrożenia ubóstwem mają te same determinanty, chociaż oddziałują z różną siłą na wskaźniki grupowe i wskaźnik syntetyczny.

Na wzrost stopnia zagrożenia ubóstwem monetarnym i niemonetarnym oraz zagrożenia wielowymiarowego najsilniej wpływa niski status aktywności na rynku pracy. Największe wielowymiarowe zagrożenie ubóstwem w zależności od statusu na rynku pracy występuje wśród rencistów (wykr. 1). Spadek aktywności ekonomicznej na rynku pracy powoduje zmniejszanie się dochodów, a w konsekwencji – wzrost zagrożenia ubóstwem.

Wykr. 1. Stopień zagrożenia ubóstwem monetarnym i niemonetarnym łącznie w zależności od statusu aktywności na rynku pracy w 2018 r.

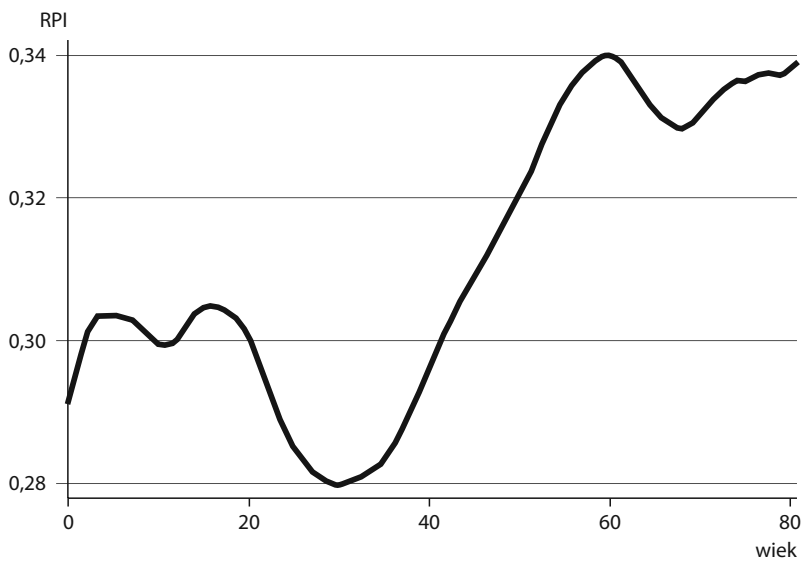


Źródło: opracowanie własne na podstawie danych z EU-SILC z wykorzystaniem pakietu STATA 16.

Stopień wielowymiarowego zagrożenia ubóstwem silnie wzrasta także wraz z wiekiem (wykr. 2). W przypadku osób w wieku od 0 do 20 lat zależy on przede wszystkim od sytuacji materialnej rodziców lub opiekunów utrzymujących te osoby i ulega tylko nieznacznym wahaniom. Osoby w wieku od 20 do 30 lat podejmują aktywność

zawodową na rynku pracy, co pozwala im uzyskać dochody umożliwiające zaspokojenie ich podstawowych potrzeb na minimalnym akceptowalnym poziomie i tym samym znacząco redukuje ich zagrożenie ubóstwem. W grupie osób w wieku od 30 do 60 lat występuje największy wzrost stopnia zagrożenia ubóstwem wielowymiarowym, co wynika z pogarszania się wraz z wiekiem stanu zdrowia i związanym z tym ograniczeniem możliwości aktywności ekonomicznej. Osoby w wieku od 60 do 70 lat mają stałe źródło dochodu w postaci emerytury, dlatego ich stopień zagrożenia ubóstwem spada tylko nieznacznie. Natomiast osoby w wieku 70 lat i więcej potrzebują intensywniejszej opieki medycznej, co skutkuje wzrostem wydatków na cele zdrowotne i wzrostem zagrożenia ubóstwem.

Wykr. 2. Stopień zagrożenia ubóstwem monetarnym i niemonetarnym łącznie w zależności od wieku w 2018 r.



Źródło: opracowanie własne na podstawie danych z EU-SILC z wykorzystaniem pakietu STATA 16.

Determinantą, która istotnie oddziałuje na zmniejszenie stopnia zagrożenia ubóstwem (zarówno monetarnym, jak i niemonetarnym oraz łącznym) jest przede wszystkim poziom wykształcenia. Wraz z jego wzrostem rosną bowiem możliwości podjęcia dobrze płatnej pracy, czyli uzyskania dochodów umożliwiających zaspokojenie podstawowych potrzeb na minimalnym, akceptowalnym poziomie.

3.3. Analiza zagrożenia ubóstwem w różnych przekrojach

Grupowy RPI umożliwia porównanie stopnia zagrożenia ubóstwem pomiędzy jednostkami terytorialnymi i grupami typologicznymi gospodarstw domowych oraz porównanie w czasie. Wartość RPI w poszczególnych przekrojach należy interpretować jako średni stopień zagrożenia ubóstwem. Im większa wartość RPI, tym wyższy jest stopień zagrożenia ubóstwem. W tabl. 2 przedstawiono porównanie stopnia zagrożenia ubóstwem według makroregionów, wielkości miejscowości zamieszkania, źródła utrzymania gospodarstw domowych i typu rodziny. Do szacunków wykorzystano dane z badania EU-SILC przeprowadzonego przez GUS w 2018 r.

Tabl. 2. Stopień zagrożenia ubóstwem monetarnym i niemonetarnym oraz wielowymiarowego zagrożenia ubóstwem w różnych przekrojach w 2018 r.

Wyszczególnienie	Wielowymiarowe zagrożenie ubóstwem	Zagrożenie ubóstwem	
		monetarnym	niemonetarnym
Makroregiony			
Południowy	0,303	0,148	0,205
Północno-zachodni	0,306	0,140	0,205
Południowo-zachodni	0,296	0,151	0,188
Północny	0,310	0,167	0,208
Centralny	0,310	0,171	0,211
Wschodni	0,324	0,232	0,226
Woj. mazowieckie	0,286	0,138	0,181
P o l s k a	0,305	0,162	0,204
Wielkość miejscowości zamieszkania			
Bardzo duże miasto	0,254	0,096	0,159
Duże miasto	0,279	0,111	0,187
Średnie miasto	0,289	0,136	0,188
Małe miasto	0,300	0,135	0,200
Bardzo małe miasto	0,312	0,153	0,209
Wieś	0,328	0,214	0,224
P o l s k a	0,305	0,162	0,204
Gospodarstwa domowe według źródła utrzymania			
Pracowników	0,282	0,094	0,181
Rolników	0,309	0,326	0,185
Utrzymujących się z pracy na własny rachunek	0,267	0,153	0,150
Emerytów	0,329	0,200	0,230
Rencistów	0,454	0,462	0,353
Utrzymujących się ze źródeł niezarobkowych	0,356	0,328	0,267
P o l s k a	0,305	0,162	0,204

Tabl. 2. Stopień zagrożenia ubóstwem monetarnym i niemonetarnym oraz wielowymiarowego zagrożenia ubóstwem w różnych przekrojach w 2018 r. (dok.)

Wyszczególnienie	Wielowymiarowe zagrożenie ubóstwem	Zagrożenie ubóstwem	
		monetarnym	niemonetarnym
Gospodarstwa domowe według typu rodziny			
Jednorodzinne: małżeństwo / partnerzy bez dzieci	0,297	0,121	0,187
małżeństwo / partnerzy z jednym dzieckiem	0,292	0,130	0,186
małżeństwo / partnerzy z dwójką dzieci	0,283	0,142	0,179
małżeństwo / partnerzy z trójką i więcej dzieci	0,293	0,185	0,196
Rodzina niepełna	0,360	0,222	0,276
Wielorodzinne	0,295	0,112	0,200
Nierodzinne: jednoosobowe	0,361	0,295	0,266
wieloosobowe	0,374	0,273	0,288
P o l s k a	0,305	0,162	0,204

Uwaga. Jak przy zestawieniu 2. Województwa wchodzące w skład makroregionów: południowy: małopolskie, śląskie; północno-zachodni: lubuskie, wielkopolskie, zachodniopomorskie; południowo-zachodni: dolnośląskie, opolskie; północny: kujawsko-pomorskie, pomorskie, warmińsko-mazurskie; centralny: łódzkie, świętokrzyskie; wschodni: lubelskie, podkarpackie, podlaskie.

Źródło: opracowanie własne na podstawie badania przeprowadzonego z wykorzystaniem danych z EU-SILC.

3.4. Porównanie wielowymiarowej oceny zagrożenia ubóstwem z oceną w ujęciu jednowymiarowym

W myśl przyjętej w pracy ekonomicznej definicji ubóstwa analizy ubóstwa w ujęciu jednowymiarowym prowadzą do nie w pełni poprawnej identyfikacji osób ubogich. W podejściu jednowymiarowym uwzględnia się bowiem tylko jeden czynnik, np. poziom zaspokojenia potrzeb gospodarstw domowych wyłącznie przez pryzmat ich bieżących dochodów (bądź wydatków), a nie bierze się pod uwagę dochodów z poprzednich okresów (oszczędności). W takiej sytuacji jako ubogie gospodarstwa domowe mogą być zakwalifikowane również te, które dzięki oszczędnościom z poprzednich lat mają wystarczające środki finansowe na zaspokojenie swoich potrzeb. Dlatego przeprowadzono porównanie, pozwalające określić, w jakim stopniu wyniki oceny zagrożenia ubóstwem w ujęciu jednowymiarowym (tu: wyrażonego wskaźnikiem monetarnym – stopą ubóstwa) różnią się od wyników uzyskanych w ujęciu wielowymiarowym (z wykorzystaniem wskaźników monetarnych i niemonetarnych). Tego rodzaju analizy nie dokonano dotychczas w żadnej pracy badawczej.

Ponieważ nie jest możliwe porównanie stopnia zagrożenia ubóstwem wyrażonego w procentach całkowitego zagrożenia ubóstwem ze stopą ubóstwa monetarnego mierzoną jako odsetek osób ubogich, uzyskane wyniki zaprezentowano w postaci rankingów, w których pierwsze miejsce oznacza największe wielowymiarowe zagrożenie ubóstwem lub najwyższy odsetek osób ubogich w ujęciu jednowymiarowym (tabl. 3).

Tabl. 3. Ranking pod względem stopnia wielowymiarowego zagrożenia ubóstwem (ujęcie wielowymiarowe) i stopy ubóstwa monetarnego (ujęcie jednowymiarowe) w 2018 r.

Wyszczególnienie	Wielowymiarowe zagrożenie ubóstwem	Stopa ubóstwa
Makroregiony		
Wschodni	1	1
Północny	2	3
Centralny	2	2
Północno-zachodni	3	6
Południowy	4	5
Południowo-zachodni	5	4
Woj. mazowieckie	6	6
Wielkość miejscowości zamieszkania		
Wieś	1	1
Bardzo małe miasto	2	2
Małe miasto	3	4
Średnie miasto	4	3
Duże miasto	5	5
Bardzo duże miasto	6	6
Gospodarstwa domowe według źródła utrzymania		
Rencistów	1	1
Utrzymujących się ze źródeł niezarobkowych	2	2
Emerytów	3	4
Rolników	4	3
Pracowników	5	6
Utrzymujących się z pracy na własny rachunek	6	5
Gospodarstwa domowe według typu rodziny		
Nierodzinne: wieloosobowe	1	2
jednoosobowe	2	1
Rodzina niepełna	3	3
Jednorodzinne – małżeństwo / partnerzy bez dzieci	4	7
Wielorodzinne	5	8
Jednorodzinne: małżeństwo / partnerzy z trójką i więcej dzieci	6	4
małżeństwo / partnerzy z jednym dzieckiem	7	6
małżeństwo / partnerzy z dwójką dzieci	8	5

Uwaga. Jak przy tabl. 2.

Źródło: opracowanie własne na podstawie badania przeprowadzonego z wykorzystaniem danych z EU-SILC.

Na podstawie wyników przeprowadzonego badania empirycznego i powyższego rankingu sformułowano wnioski dotyczące stopnia zagrożenia ubóstwem w poszczególnych przekrojach.

Po pierwsze, najwyższą wartość RPI odnotowano w makroregionie wschodnim, obejmującym województwa: lubelskie, podkarpackie i podlaskie. Ten makroregion jest także najbardziej zagrożony ubóstwem pod względem stopy ubóstwa monetarnego. Jednak w makroregionach znajdujących się na dalszych miejscach w rankingach występują różnice w stopniu zagrożenia ubóstwem w ujęciu wielowymiarowym i jednowymiarowym. Największa rozbieżność dotyczy makroregionu północno-zachodniego, w skład którego wchodzą województwa: lubuskie, wielkopolskie i zachodniopomorskie. W ujęciu wielowymiarowym jest on bowiem trzecim z kolei obszarem najbardziej zagrożonym ubóstwem, natomiast w jednowymiarowym pomiarze ubóstwa wystąpił tam najmniejszy odsetek osób ubogich.

Po drugie, najbardziej zagrożony ubóstwem w ujęciu zarówno wielowymiarowym, jak i jednowymiarowym jest obszar wiejski. Jeśli chodzi o miasta, to wartość RPI maleje wraz ze wzrostem liczby ludności, natomiast w ujęciu jednowymiarowym ta prawidłowość nie występuje – większy odsetek osób ubogich jest notowany w miastach średnich (100–200 tys. mieszkańców) niż w małych (20–100 tys. mieszkańców).

Po trzecie, najbardziej zagrożoną grupą osób w zależności od źródła utrzymania gospodarstw domowych w obu rozpatrywanych ujęciach są osoby z gospodarstw domowych rencistów, a w drugiej kolejności – osoby z gospodarstw utrzymujących się ze źródeł niezarobkowych. Na trzecim miejscu w rankingu pod względem wielowymiarowego zagrożenia ubóstwem znajdują się osoby z gospodarstw emerytów, natomiast w rankingu pod względem odsetka ubogich – osoby z gospodarstw rolników, które w ujęciu wielowymiarowym zajmują pozycję czwartą. Także w przypadku pozostałych grup występują rozbieżności w ocenie stopnia zagrożenia ubóstwem wielowymiarowym i jednowymiarowym.

Po czwarte, pod względem typu rodziny najbardziej zagrożone ubóstwem w ujęciu wielowymiarowym są osoby z nierodzinnych wieloosobowych gospodarstw domowych. Natomiast w ujęciu jednowymiarowym największy odsetek ubogich występuje wśród osób z gospodarstw nierodzinnych jednoosobowych. Członkowie małżeństw lub partnerzy bez dzieci zajmują czwartą pozycję w rankingu pod względem stopnia wielowymiarowego zagrożenia ubóstwem, a dopiero siódmą (przedostatnią) – pod względem odsetka osób ubogich. Osoby z gospodarstw wielorodzinnych lokują się na piątym miejscu w ujęciu wielowymiarowym, a pod względem bieżących dochodów – na ostatnim miejscu, co znaczy, że w ujęciu jednowymiarowym w tej grupie osób występuje najniższy odsetek osób ubogich. Jak już wspomniano, powodem znaczących różnic w porównywanych rankingach jest nieuwzględnianie w ocenie ubóstwa w ujęciu jednowymiarowym oszczędności z poprzednich okresów.

4. Podsumowanie

Zaproponowana operacjonalizacja pomiaru ubóstwa w ramach podejścia możliwości Sena przy wykorzystaniu modelu strukturalnego MIMIC pozwala na ocenę stopnia zagrożenia ubóstwem z uwzględnieniem zbioru rzeczywistych, zdeprywowanych sposobów funkcjonowania badanych osób (symptomów ubóstwa), odzwierciedlających deprywację ich możliwości, oraz czynników wzmacniających lub osłabiających zagrożenie jednostek ubóstwem (charakterystyk jednostek). Umożliwia też dokonanie wielowymiarowego pomiaru ubóstwa – wszechstronniejszego niż za pomocą innych stosowanych w praktyce metod analizy ubóstwa. Posługując się wskaźnikiem stopnia zagrożenia ubóstwem (Risk of Poverty Index – RPI), można oceniać stopień zagrożenia ubóstwem w różnych przekrojach: według jednostek terytorialnych i grup typologicznych gospodarstw domowych, a także w czasie.

Ujęcie wielowymiarowe zapewnia również precyzyjniejszą identyfikację osób ubogich zgodnie z przyjętą definicją ubóstwa niż ujęcie jednowymiarowe. Podstawę do sformułowania takiego wniosku daje porównanie wyników oceny stopnia zagrożenia ubóstwem w obu ujęciach. Pokazuje ono, że w analizie ubóstwa ujęcie jednowymiarowe jest niewystarczające do identyfikacji osób ubogich i oceny stopnia zagrożenia ubóstwem. Uwidacznia różnice pomiędzy rankingami według makroregionów, wielkości miejscowości zamieszkania oraz grup typologicznych gospodarstw domowych w ujęciu wielowymiarowym i jednowymiarowym, potwierdzając tym samym mankamenty ujęcia jednowymiarowego. Powodem znaczących różnic w porównywanych rankingach jest to, że identyfikacja osób ubogich w ujęciu jednowymiarowym bazuje tylko na bieżących dochodach lub wydatkach (nie uwzględnia oszczędności posiadanych przez jednostkę), w wyniku czego osoba, która faktycznie dysponuje wystarczającymi środkami finansowymi na zaspokojenie podstawowych potrzeb, może zostać uznana za osobę ubogą.

Przeprowadzona w badaniu identyfikacja czynników zmniejszających lub zwiększających stopień zagrożenia ubóstwem pokazała, że na zmniejszenie stopnia zagrożenia ubóstwem najsilniej wpływa wzrost poziomu wykształcenia, ponieważ wraz z nim zwiększają się możliwości podjęcia dobrze płatnej pracy, czyli uzyskania dochodów, które pozwalają jednostce na zaspokojenie jej potrzeb na minimalnym akceptowalnym poziomie. Natomiast za zwiększenie stopnia zagrożenia ubóstwem odpowiada przede wszystkim niska aktywność zawodowa osób na rynku pracy, powodująca spadek dochodów.

Przedstawiona wielowymiarowa metoda badania zagrożenia ubóstwem powinna wzbogacić warsztat metodologiczny urzędów statystycznych i badaczy zjawiska ubóstwa oraz stanowić narzędzie wspomagające dla instytucji prowadzących politykę

społeczną nakierowaną na walkę z ubóstwem. Charakterystykę osób i ich gospodarstw domowych zaklasyfikowanych jako ubogie można wykorzystać przy dystrybucji pomocy społecznej w celu zwalczania ubóstwa. Charakterystyka ta umożliwi nie tylko identyfikację osób ubogich, w przypadku których udzielenie bieżącej pomocy jest niezbędne do zaspokojenia ich podstawowych potrzeb, lecz także efektywne określenie czynników powodujących wzrost stopnia zagrożenia ubóstwem.

Bibliografia

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons. https://doi.org/10.1002/9781118619179.

European Commission. (2010). *Europe 2020. A strategy for smart, sustainable and inclusive growth*. Communication from the Commission, COM(2010) 2020.

European Commission. (2011, 27 stycznia). *The measurement of extreme poverty in the European Union*. https://ec.europa.eu/social/main.jsp?catId=89&langId=en&newsId=982&furtherNews=yes.

European Commission. (2019). *Towards a Sustainable Europe by 2030*. Reflection Paper, COM(2019) 22.

Eurostat. (b.r.). *EU statistics on income and living conditions (EU-SILC) methodology – introduction*. Pobrane 18 stycznia 2022 r. z https://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_introduction#Scope_and_audience.

Główny Urząd Statystyczny. (2015). *Praca badawcza pt. „Pomiar ubóstwa na poziomie powiatów (LAU 1) – etap II"*. https://stat.gov.pl/statystyki-eksperymentalne/jakosc-zycia/pomiar-ubostwa-na-poziomie-powiatow-lau-1-popt-2007-2013,4,1.html.

Główny Urząd Statystyczny. (2019). *Dochody i warunki życia ludności Polski – raport z badania EU-SILC 2018*. https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/dochody-i-warunki-zycia-ludnosci-polski-raport-z-badania-eu-silc-2018,6,12.html.

Główny Urząd Statystyczny, Urząd Statystyczny w Łodzi. (2013). *Jakość życia, kapitał społeczny, ubóstwo i wykluczenie społeczne w Polsce*. https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/jakosc-zycia-kapital-spoleczny-ubostwo-i-wykluczenie-spoleczne-w-polsce,1,1.html.

Hair jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E. (2014). *Multivariate Data Analysis* (7th edition). Pearson Education Limited. https://files.pearsoned.de/inf/ext/9781292035116.

Hauser, R. M., Goldberger, A. S. (1971). The Treatment of Unobservable Variables in Path Analysis. *Sociological Methodology*, *3*, 81–117. https://doi.org/10.2307/270819.

Jalan, J., Ravallion, M. (1998). *Determinants of Transient and Chronic Poverty. Evidence from Rural China* (Policy Research Working Paper No. 1936). https://doi.org/10.1596/1813-9450-1936.

Jöreskog, K. G., Goldberger, A. S. (1975). Estimation of Model with Multiple Indicators and Multiple Causes of a Single Latent Variable. *Journal of the American Statistical Association*, *70*(351), 631–639. https://doi.org/10.2307/2285946.

Konarski, R. (2014). *Modele równań strukturalnych. Teoria i praktyka*. Wydawnictwo Naukowe PWN.

Krishnakumar, J. (2008). Multidimensional Measures of Poverty and Well-Being Based on Latent Variable Models. W: N. Kakwani, J. Silber (red.), *Quantitative Approaches to Multidimensional Poverty Measures* (s. 118–134). Palgrave Macmillan. https://doi.org/10.1057/9780230582354_7.

Krishnakumar, J., Ballon, P. (2008). Estimating Basic Capabilities: A Structural Equation Model Applied to Bolivia. *World Development*, *36*(6), 992–1009. https://doi.org/10.1016/j.worlddev.2007.10.006.

Kurowska, A. (2008). *Skąd się bierze bieda?*. Forum Obywatelskiego Rozwoju. https://for.org.pl/upload/File/zeszyty/Zeszyty_Skad_sie_bierze_bieda_Kurowska.pdf.

McDonald, R. P., Ho, M.-H. R. (2002). Principles and Practice in Reporting Structural Equation Analyses. *Psychological Methods*, *7*(1), 64–82. https://doi.org/10.1037/1082-989X.7.1.64.

Muthen, B. O., Siek-Toon, K., Goff, G. N. (1994). *Multidimensional Description of Subgroup Differences in Mathematics Achievement Data from the 1992 National Assessment of Educational Progress*. National Center for Research on Evaluation, Standards, and Student Testing.

Nawawi, S. A., Busu, I., Fauzi, N., Amin, M. F. M. (2019). Determinants of Poverty: A Spatial Analysis. *Journal of Tropical Resources and Sustainable Science*, *7*(2), 83–87. https://doi.org/10.47253/jtrss.v7i2.514.

Organizacja Narodów Zjednoczonych. (2015). *Rezolucja przyjęta przez Zgromadzenie Ogólne w dniu 25 września 2015 r. Przekształcamy nasz świat: Agenda na rzecz zrównoważonego rozwoju 2030*. http://unic.un.org.pl/files/164/Agenda%202030_pl_2016_ostateczna.pdf.

Panek, T., Zwierzchowski, J. (2021). *Economic Growth, Poverty, Inequality and Social Transfers in the European Union*. SGH Publishing House. https://agathos.sgh.waw.pl/spisy/pelny_tekst/d108028.pdf.

Panek, T., Zwierzchowski, J. (2022). Examining the Degree of Social Exclusion Risk of the Population Aged 50 + in the EU Countries Under the Capability Approach. *Social Indicators Research*, *163*(3), 973–1002. https://doi.org/10.1007/s11205-022-02930-9.

Radziukiewicz, M. (2006). Propozycja budowy miernika biedy skumulowanej w Polsce. *Polityka Społeczna*, (11–12), 3–8. https://www.ipiss.com.pl/wp-content/uploads/downloads/2012/10/ps_11-12_2006_m_radziukiewicz.pdf.

Sekhampu, T. J. (2013). Determinants of Poverty in a South African Township. *Journal of Social Sciences*, *34*(2), 145–153. https://doi.org/10.1080/09718923.2013.11893126.

Sen, A. (1979). *Equality of what?*. https://www.ophi.org.uk/wp-content/uploads/Sen-1979_Equality-of-What.pdf.

Sen, A. (1985). *Commodities and capabilities*. North-Holland.

Sen, A. (1992). *Inequality Reexamined*. Clarendon Press.

Wiadomości Statystyczne. The Polish Statistician, 2022, vol. 67, 10, 48–61 DOI: 10.5604/01.3001.0016.0659
Statystyka w praktyce / Statistics in practice
Zgłoszony/submitted: 30.06.2022, poprawiony/revised: 19.09.2022, zaakceptowany/accepted: 28.09.2022

Efektywność zajęć zdalnych w czasie pandemii COVID-19[1]

Elżbieta Zalewska[a], Kamila Trzcińska[b]

Streszczenie. Szkolnictwo, zarówno pod względem treści, jak i formy nauczania, powinno nadążać za zmianami dokonującymi się w otaczającej nas rzeczywistości. W ostatnim czasie zmiany te były stymulowane – oprócz postępu technologicznego – przez pandemię COVID-19 i związane z nią obostrzenia, które wymusiły przejście w tryb nauki zdalnej. Nasuwa się więc pytanie, jak efektywne było kształcenie online.

Celem badania omawianego w artykule jest porównanie efektywności kształcenia akademickiego w trybie stacjonarnym i zdalnym. Dane do analiz zaczerpnięto z bazy wyników uzyskanych przez studentów I roku studiów dziennych na kierunku finanse i rachunkowość, prowadzonych na Wydziale Ekonomiczno-Socjologicznym Uniwersytetu Łódzkiego (UŁ), z egzaminów końcowych z matematyki w latach akademickich 2018/2019 (nauka stacjonarna) i 2020/2021 (nauka zdalna) oraz – jako punktu odniesienia – z egzaminów maturalnych z matematyki na poziomie podstawowym zdanych przez badane osoby. W podanych latach przez cały cykl zajęć obowiązywał jeden tryb nauczania. Efektywność kształcenia została oceniona na podstawie krzywej Lorentza i współczynnika Giniego, powszechnie używanego do określania nierówności dochodów, ale możliwego do zastosowania w przypadku każdego rodzaju danych o nierównym rozkładzie.

Na podstawie przeprowadzonych analiz stwierdzono, że na UŁ nie występują istotne różnice pomiędzy efektami nauki zdalnej i stacjonarnej, co świadczy o tym, że uczelnia zachowała wysoką jakość kształcenia pomimo zmiany trybu pracy. Statystyki opisowe wyników weryfikacji wiedzy z matematyki dla badanych grup studentów – w szczególności lewostronna asymetria – świadczą o nieznacznej przewadze pozytywnych wyników egzaminów końcowych w okresie nauki zdalnej w porównaniu z wynikami uzyskanymi podczas nauki stacjonarnej. Dotyczy to również egzaminów maturalnych. Jednak różnice te nie są znaczące, wynoszą bowiem ok. 0,1 p.proc.

Słowa kluczowe: efektywność kształcenia, nauka zdalna, pandemia COVID-19, współczynnik Giniego

JEL: I21, I23, C02

[1] Artykuł został opracowany na podstawie referatu wygłoszonego na XV Międzynarodowej Konferencji Naukowej im. Profesora Aleksandra Zeliasia na temat „Modelowanie i prognozowanie zjawisk społeczno-gospodarczych", która odbyła się w dniach 9–12 maja 2022 r. w Zakopanem. / The article is based on a paper delivered at the 15th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena, held on 9–12 May 2022 in Zakopane.

[a] Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Polska / University of Lodz, Faculty of Economics and Sociology, Poland. ORCID: https://orcid.org/0000-0003-1544-300X.
Autor korespondencyjny / Corresponding author, e-mail: elzbieta.zalewska@eksoc.uni.lodz.pl.

[b] Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Polska / University of Lodz, Faculty of Economics and Sociology, Poland. ORICD: https://orcid.org/0000-0002-4714-4074.
E-mail: kamila.trzcinska@uni.lodz.pl.

© Elżbieta Zalewska, Kamila Trzcińska
Artykuł udostępniony na licencji CC BY-SA 4.0 / Article available under the CC BY-SA 4.0 licence

Effectiveness of distance learning during the COVID-19 pandemic

Abstract. Education should be able to keep up with the changes taking place in the surrounding reality, both in terms of its content and mode of work. In addition to technological progress, these changes have been recently stimulated by the COVID-19 pandemic and the related restrictions which enforced distance learning. These circumstances pose a question about the effectiveness of online learning.

The aim of this paper is to compare the effectiveness of distance and classroom learning. The analyses are based on the results of the final exams in mathematics obtained by first-year full-time students majoring in finance and accounting at the University of Lodz in the years 2018/2019 (classroom learning) and 2020/2021 (distance learning). These students' scores achieved in the basic-level matura exam in mathematics served as a point of reference. The teaching mode remained the same throughout the entire cycle of classes in the aforementioned years. The assessment of the effectiveness of learning was based on the Lorenz curve and the Gini coefficient, which is commonly used by economists to determine income inequality. However, it can also be applied to any kind of data of an unequal distribution.

The conducted research showed no significant differences in terms of the effects of distance learning and classroom learning, which proves that the university maintained a high quality of education despite the introduced change in the mode of teaching. The descriptive statistics relating to the verification of the mathematical knowledge in the selected student groups – in particular the left-sided asymmetry – indicate a slight predominance of positive results achieved in the final exams during distance learning compared to the results obtained during classroom education, which also applies to the matura exams. However, these differences are not significant as they differ by approximately 0.1 p.p.

Keywords: education efficiency, distance learning, COVID-19 pandemic, Gini coefficient

1. Wprowadzenie

Z powodu pandemii COVID-19, która wybuchła pod koniec 2019 r., i związanych z nią obostrzeń szkoły, w tym jednostki prowadzące studia wyższe, zostały zmuszone do szybkiego zorganizowania zajęć dydaktycznych w trybie zdalnym. Ta nagła sytuacja była testem na zdolność adaptacji zarówno nauczycieli/wykładowców, jak i uczniów/studentów do nowych realiów. Powstaje pytanie, czy w tych warunkach udało się zachować odpowiednią jakość kształcenia.

Jednym z głównych elementów oceny jakości kształcenia jest jego efektywność. Wzrost efektywności kształcenia definiuje się jako osiągnięcie lepszych wyników przy niezmienionych kosztach/nakładach (Ryl-Zaleska, 2022). Badania kwestionariuszowe dotyczące skuteczności nauki online podczas pandemii COVID-19 wśród studentów i wykładowców prowadzili m.in. Alaghawat i Alshamailah (2022), Asgari i in. (2021), Bahasoan i in. (2020), Butnaru i in. (2021) oraz Hebebci i in.

(2020). Z kolei Bonnini i Cavallo (2021) oraz Mohammed Ali (2021) analizowali wpływ pandemii na efektywność nauczania wśród osób niepełnosprawnych.

Od wybuchu pandemii COVID-19 metody nauczania akademickiego zmieniły się zasadniczo na całym świecie. Celem wprowadzenia zajęć zdalnych było zapewnienie dystansu społecznego i powstrzymanie rozprzestrzeniania się infekcji. Dystans społeczny, w szczególności edukacyjny, wymagał wdrożenia nowego sposobu kształcenia w skali całego kraju, co okazało się trudnym procesem. Uczelnie musiały wypróbować innowacyjne rozwiązania w tym zakresie i utrzymać ciągłość edukacji, zachowując dotychczasowe standardy jakości. Jakość kształcenia zdalnego jest w tym kontekście kwestią kluczową, która wymaga odpowiedniej uwagi (Sahu, 2019). Polska była jednym z pierwszych europejskich krajów, które wprowadziły obowiązek nauki zdalnej. Początkowo wydawało się, że jest to rozwiązanie chwilowe, ale ostatecznie ta forma kształcenia utrzymała się przez dłuższy czas – na polskich uczelniach obowiązywała prawie dwa lata. W ten sposób pandemia COVID-19 zapoczątkowała cyfrową transformację szkolnictwa wyższego (Strielkowski, 2020). W związku z tym ważna jest ocena efektywności nowych sposobów nauczania należących do grupy e-kształcenia.

Pojęcie *e-kształcenia* jest trudne do precyzyjnego zdefiniowania. Pod tą nazwą kryje się wiele innych pojęć, m.in. takich jak: *nauka komputerowa*, *nauka przez internet*, *e-learning*, *nauka z wykorzystaniem technologii mobilnych* lub *kształcenie na odległość*. Różnica pomiędzy e-learningiem a kształceniem na odległość polega na sposobie wykorzystania internetu w trakcie prowadzenia zajęć. Podczas kształcenia na odległość wykładowca znajduje się w innym miejscu niż studenci i prowadzi zajęcia w sposób korespondencyjny, np. listownie, a pojęcie *e-learningu* wiąże się z nauką na odległość i jednoczesnym wykorzystaniem internetu jako platformy kształcenia. Ponadto e-kształcenie może być zarówno uzupełnieniem zajęć tradycyjnych, jak i jedynym sposobem prowadzenia zajęć. Często pojęcia te są utożsamiane. E-kształcenie to dziedzina o charakterze interdyscyplinarnym, której „celem jest znalezienie odpowiedzi na pytanie, jak kształcić w sposób najbardziej efektywny" (Kopciał, 2013, s. 81). Kształcenie na odległość charakteryzuje się inną formą organizacji procesów dydaktycznych niż nauczanie tradycyjne. Przez *e-learning* najczęściej rozumiemy naukę z wykorzystaniem internetu, czyli system uczenia się przy użyciu technologii informatycznej (Główny Urząd Statystyczny, b.r.). Pojawiają się również szersze definicje e-learningu, np. jako uczenie się wspomagane narzędziami i nośnikami elektronicznymi (Topol, 2020).

W literaturze przedmiotu istnieje podział metod e-learningowych ze względu na: dostępność w czasie (typ synchroniczny i typ asynchroniczny), charakter relacji między studentem a wykładowcą (kurs, w którym uczestniczą słuchacze wraz z wykładowcą, oraz kurs, w którym student posługuje się materiałami multimedialnymi i uczy się samodzielnie) oraz odniesienie do nauki tradycyjnej (nauka zdalna, inaczej online, i hybrydowa, zwana także *mieszaną* – ang. *blended learning*; Szymańska i Zalewska, 2015).

Formą nauki online najczęściej stosowaną na uczelniach przed pandemią był *blended learning*, czyli uzupełnianie metod tradycyjnych e-zajęciami. Główną zaletą nauki mieszanej jest możliwość wykorzystania zdalnych i bezpośrednich form aktywizacji studentów. E-learning był traktowany jako nauka wspomagająca metody tradycyjne, która może wnieść nową jakość do kształcenia akademickiego (Dąbrowski, 2013; Maleńczyk i Gładysz, 2019; Romaniuk, 2015). Chociaż zwracano uwagę, że jest to forma nauki w pewnym sensie nieunikniona (Zalewska, 2015), to nikt się nie spodziewał, że zostanie wdrożona z dnia na dzień. W pierwszych miesiącach pandemii przejście na naukę zdalną było pod wieloma względami chaotyczne (Alirezabeigi i in., 2020). W ramach polskiego systemu akademickiego nauczanie prowadzono głównie w formie zajęć synchronicznych (transmitowanych na żywo) i częściowo asynchronicznych (m.in. poprzez przesyłanie gotowych materiałów). Obawy dotyczące zdrowia i bezpieczeństwa społecznego doprowadziły do zawieszenia wszelkich kontaktów i działań bezpośrednich, w tym również praktyk i staży.

Głównym problemem, przed którym stanęły uczelnie, a w szczególności wykładowcy, był brak czasu i szkoleń informujących, w jaki sposób efektywnie przenieść tradycyjne kształcenie na zajęcia online. W pierwszych dniach pandemii większość uczelni nie miała przygotowanych wytycznych dla wykładowców, jak przeprowadzić zajęcia zdalne. Konieczność zapewnienia wysokiej jakości kształcenia – będącej głównym czynnikiem determinującym konkurencyjność uczelni – sprawiła, że w czasie pandemii nauka na odległość stała się najszybciej rozwijającym się segmentem szkolnictwa wyższego. Tylko elastyczne, nowoczesne i wyspecjalizowane instytucje świadczące wysokiej jakości usługi są bowiem w stanie zdobyć i utrzymać silną pozycję rynkową (Zalewska, 2021).

Celem badania omawianego w artykule jest porównanie efektywności kształcenia akademickiego w trybie stacjonarnym i zdalnym.

2. Metoda badania

W badaniu wykorzystano bazę wyników uzyskanych przez studentów I roku studiów dziennych na kierunku finanse i rachunkowość (FiR), prowadzonych na Wydziale Ekonomiczno-Socjologicznym Uniwersytetu Łódzkiego, z egzaminów końcowych z matematyki w latach akademickich 2018/2019 (nauka stacjonarna) i 2020/2021 (nauka zdalna) oraz egzaminów maturalnych z matematyki na poziomie podstawowym zdanych przez badane osoby. Porównywane grupy miały podobną liczebność: w roku akademickim 2018/2019 w badaniu uczestniczyło 132 studentów, a w roku 2020/2021 – 175. Dany tryb kształcenia obowiązywał przez cały cykl zajęć. Dokonując wyboru grup badawczych, sugerowano się jednakowym trybem prowadzenia zajęć i przeprowadzenia egzaminu z matematyki (online/stacjonarnie). Badane grupy były również podobne pod względem wyników egzaminu maturalnego. Umożliwiło to porównanie jakości kształcenia akademickiego prowadzonego w trybie stacjonarnym oraz zdalnym.

W roku akademickim 2018/2019 badana grupa studentów uczestniczyła w zajęciach z matematyki prowadzonych stacjonarnie. Zajęcia zakończyły się egzaminem w postaci testu, który również odbył się stacjonarnie, w sali komputerowej. Test składał się z 30 pytań wielokrotnego wyboru, na które studenci udzielali odpowiedzi na formularzu przygotowanym na platformie Moodle. Egzamin został przeprowadzony w tym samym czasie dla całego kierunku studiów. Studenci odpowiadali na te same pytania, ale ich kolejność była dobierana losowo. Dodatkowo nie można było cofać się do pytań wcześniejszych. Czas na rozwiązanie zadań był jednakowy we wszystkich grupach. Bezpośrednio po zakończeniu egzaminu studenci otrzymywali wynik końcowy. Aby zaliczyć test, należało uzyskać minimum 50% punktów możliwych do zdobycia.

W roku akademickim 2020/2021 wszystkie zajęcia były prowadzone w trybie zdalnym – synchronicznie, z wykorzystaniem platformy Microsoft Teams. Wykłady miały formę prezentacji, a ćwiczenia – konwersacji z użyciem tablicy Microsoft Whiteboard. Dzięki temu studenci mogli aktywnie uczestniczyć w zajęciach, zarówno wykładowych, jak i ćwiczeniowych. Dodatkowo zadawano im prace domowe oraz przesyłano materiały do pracy samodzielnej. Forma egzaminu była identyczna jak w roku 2018/2019, ale studenci rozwiązywali test w domu.

Porównując średnie wyniki egzaminu maturalnego z matematyki uzyskane przez studentów obu badanych grup, można zauważyć, że różnice są niewielkie. Przeciętny rezultat uzyskany przez studentów odbywających I rok studiów na kierunku FiR w roku akademickim 2018/2019 (czyli na maturze zdawanej w 2018 r.) wyniósł

84,24%±10,95% pkt możliwych do zdobycia, a przez studentów odbywających I rok studiów w roku 2020/2021 (matura w 2020 r.) – 84,31%±13,45% pkt. Również różnice w wynikach egzaminu końcowego z matematyki obu badanych grup są nieznaczne. W roku akademickim 2018/2019 studenci uzyskali średnio na osobę 63,61%±20,36% pkt możliwych do zdobycia, a w roku 2020/2021, kiedy prowadzono naukę zdalną – 73,13%±19,59% pkt. Średnia różnica pomiędzy wynikami egzaminów końcowych z matematyki w latach nauki zdalnej i stacjonarnej jest niewielka i wynosi 9 p.proc., co przekłada się na 3 pkt możliwe do uzyskania w teście (zob. tablica).

Tablica. Statystyki opisowe wyników weryfikacji wiedzy z matematyki dla badanych grup studentów

Egzaminy	Liczba studentów	Średnia	Mediana	I kwartyl	IV kwartyl	Odchylenie standardowe	Skośność
		w %					
Studenci na I roku studiów w roku 2018/2019							
Maturalny	132	84,242	88	76	92	10,949	–0,752
Końcowy na I roku studiów	132	63,614	67	55	77	20,363	–0,818
Studenci na I roku studiów w roku 2020/2021							
Maturalny	175	84,309	88	78	96	13,451	–1,241
Końcowy na I roku studiów	175	73,137	76	65	89	19,589	–1,519

Źródło: opracowanie własne na podstawie danych z UŁ.

Oceny efektywności kształcenia dokonano przy wykorzystaniu krzywej Lorenza (Lorenz, 1905) – powszechnie stosowanej krzywej opisującej nierówności.

Definicja 1. Niech X będzie nieujemną, ciągłą zmienną losową, z dodatnią i skończoną wartością oczekiwaną μ oraz dystrybuantą F. Krzywa Lorenza zmiennej losowej X jest zdefiniowana następująco:

$$L(p) = \frac{1}{\mu}\int_0^p F^{-1}(t)dt, \quad p \in [0,1]. \tag{1}$$

Za pomocą krzywej Lorenza można oszacować indeks Giniego (Gini, 1914), który jest miarą koncentracji rozkładu zmiennej losowej. Miary tej używa się do statystycznego opisu szerokiej gamy czynników społeczno-gospodarczych; najczęściej jest ona stosowana do określenia stopnia nierówności w podziale dochodów (Trzcińska, 2021).

Definicja 2. Niech X będzie ciągłą zmienną losową z krzywą Lorenza $L(p)$. Indeks Giniego G jest zdefiniowany wzorem

$$G = 1 - 2\int_0^1 L(p)\,dp \tag{2}$$

lub w sposób równoważny:

$$G = \int_0^1 2\big(p - L(p)\big)dp, \quad p \in [0,1]. \tag{3}$$

Współczynnik Giniego przyjmuje wartości z przedziału [0, 1]. Pełna równomierność rozkładu występuje wtedy, gdy jego wartość jest zerowa, natomiast wzrost wartości współczynnika oznacza zwiększanie się nierówności rozkładu.

Graficznie współczynnik Giniego można zobrazować jako stosunek pola obszaru koncentracji do jego teoretycznego maksimum. Obszar koncentracji to obszar między przekątną jednostkowego kwadratu a krzywą Lorenza; jego teoretyczne maksimum odpowiada obszarowi poniżej przekątnej (linii egalitarnej). Maksymalne pole obszaru koncentracji to pole trójkąta prostokątnego wyznaczonego przez prostą $y = x$ tudzież odcinki o końcach (0, 0) i (1, 0) oraz (1, 0) i (1, 1).

Współczynnik Giniego jest postrzegany jako spójna i solidna miara jakości edukacji (Rosthal, 1978; Sheret, 1988). Wniosek ten potwierdzają badania przeprowadzone na podstawie danych dotyczących osiągnięć dla różnych poziomów wykształcenia w wielu krajach (Crespo-Cuaresma i in., 2012; Thomas i in., 2001, 2002; Ziesemer, 2011).

Krzywa Lorenza nie zmieni się, jeżeli wektor wiedzy (utożsamiany z wektorem dochodów) pomnożymy przez dowolną dodatnią liczbę rzeczywistą. Wobec tego możemy powiedzieć, że współczynnik Giniego jest indeksem relatywnej nierównomierności. Zauważmy, że współczynnik ten można interpretować jako relatywny indeks nierówności mierzący skalę proporcji wiedzy (Biernacki i Ejsmont, 2011).

Załóżmy, że mamy dwie krzywe Lorenza dla dwóch różnych populacji A i B.

Definicja 3. Rozkład populacji A dominuje nad rozkładem populacji B w sensie Lorenza $L_A(x) \geq L_B(x)$ wtedy i tylko wtedy, gdy krzywa Lorenza dla populacji A jest położona nad krzywą Lorenza dla populacji B. Jeżeli krzywe się przecinają, to są nieporównywalne.

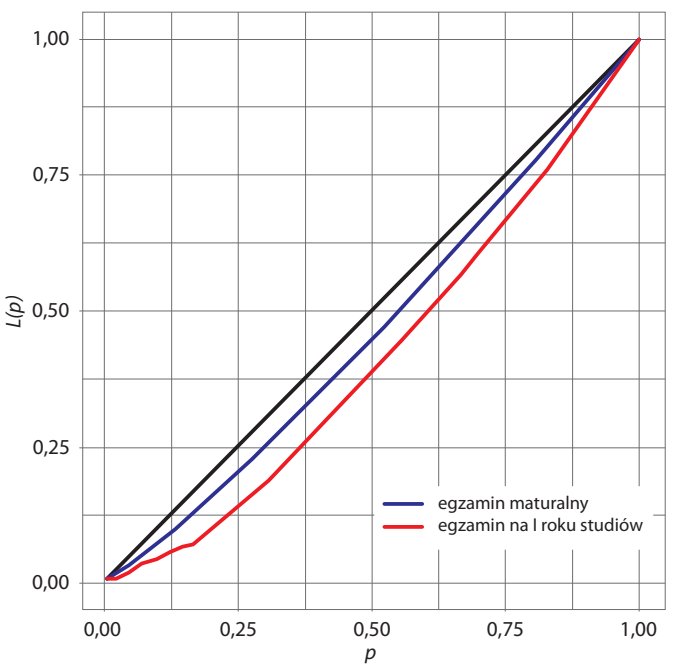
Jeżeli rozkład populacji A dominuje nad rozkładem populacji B w sensie Lorenza, to pole koncentracji dla A jest mniejsze niż pole koncentracji dla B – i tym samym w populacji A występuje mniejsza nierównomierność rozkładu niż w B. Na zbiorze

rozkładów wiedzy (utożsamianym z rozkładami dochodów), czyli w omawianym badaniu – wyników egzaminu końcowego z matematyki na I roku studiów i wyników egzaminu maturalnego z matematyki, porządek dominacji w sensie Lorenza jest więc w pewien sposób równoważny z porządkiem nierównomierności rozkładu wiedzy (Biernacki i Ejsmont, 2011).

3. Wyniki

W celu oceny jakości kształcenia podczas pandemii COVID-19 najpierw dokonano porównania wyników egzaminu końcowego z matematyki na I roku studiów oraz egzaminu maturalnego dla każdej z badanych grup studentów. W przypadku studentów I roku studiów w roku akademickim 2018/2019 – (nauka stacjonarna) rozkład wyników egzaminu końcowego z matematyki jest bardziej nierównomierny niż rozkład wyników egzaminu maturalnego, co przedstawiono na wykr. 1. Potwierdza to również wartość współczynnika Giniego, która w przypadku egzaminu maturalnego wyniosła 0,0719, a w przypadku egzaminu końcowego – 0,1734.

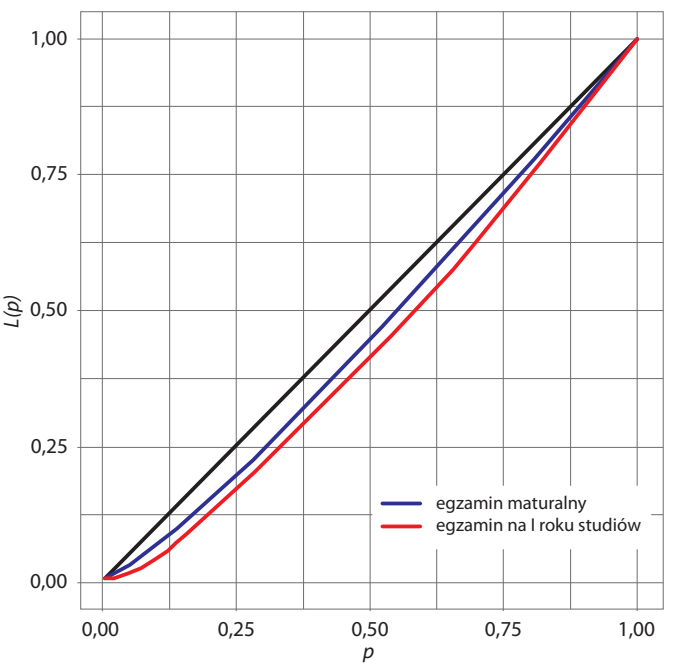
Wykr. 1. Krzywa Lorenza dla wyników egzaminu maturalnego z matematyki w 2018 r. i egzaminu końcowego z matematyki na I roku studiów w roku 2018/2019



Źródło: opracowanie własne na podstawie danych z UŁ.

Analogicznie przeanalizowano wyniki uzyskane przez studentów I roku studiów w roku 2020/2021, odbywających zajęcia w trybie zdalnym. Na ich podstawie można sformułować podobne wnioski jak dla grupy studiującej w trybie stacjonarnym, co świadczy o małym wpływie zmiany trybu kształcenia na jego efektywność (wykr. 2). Wyniki egzaminu końcowego na I roku studiów są tu bardziej nierównomierne od wyników egzaminu maturalnego. Wartość współczynnika Giniego wynosi w tym przypadku 0,0847 dla matur, a 0,1338 – dla egzaminów końcowych.

Wykr. 2. Krzywa Lorenza dla wyników egzaminu maturalnego z matematyki w 2020 r. i egzaminu końcowego z matematyki na I roku studiów w roku 2020/2021

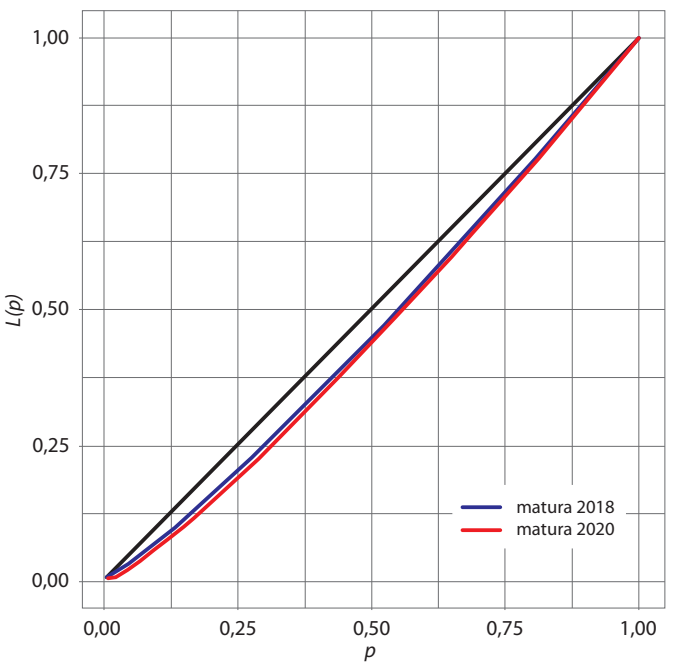


Źródło: opracowanie własne na podstawie danych z UŁ.

Wyniki zarówno egzaminu maturalnego, jak i egzaminu końcowego na I roku studiów nie uległy zmianie w porównywanych okresach, co potwierdzają wykr. 3 i 4. Rozkład wyników egzaminu maturalnego zdanego w 2020 r. okazał się bardziej nierównomierny od rozkładu wyników matury sprzed pandemii (wykr. 3). Wartość współczynnika Giniego wyniosła: 0,0847 dla matury w 2020 r. (w trybie zdalnym) i 0,0719 dla matury w 2018 r. (w trybie stacjonarnym). Można stąd również wysnuć

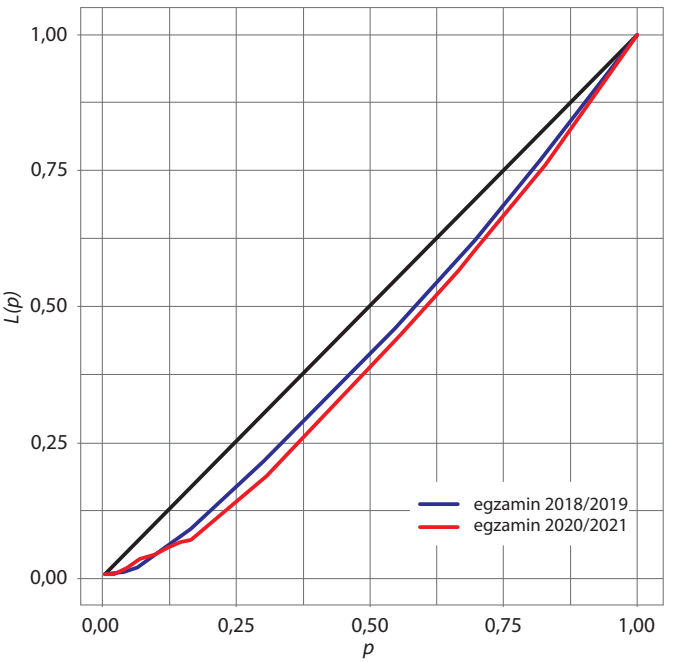
wniosek, że nierównomierności są bardzo zbliżone. Wyniki egzaminów przedstawione na wykr. 4 nie są porównywalne pod względem nierównomierności, ponieważ krzywe Lorenza się przecinają. Jednak zestawienie współczynników Giniego pokazuje, że wyniki egzaminów końcowych są zbliżone, tak samo jak wyniki matur.

Wykr. 3. Krzywa Lorenza dla wyników egzaminów maturalnych z matematyki w latach 2018 i 2020



Źródło: opracowanie własne na podstawie danych z UŁ.

Wykr. 4. Krzywa Lorenza dla wyników egzaminu końcowego z matematyki na I roku studiów w latach 2018/2019 i 2020/2021



Źródło: opracowanie własne na podstawie danych z UŁ.

4. Podsumowanie

Przeprowadzone badanie pozwoliło na porównanie efektywności zajęć z matematyki prowadzonych w trybie stacjonarnym i zdalnym, na które uczęszczali studenci I roku studiów na kierunku finanse i rachunkowość, prowadzonych na Wydziale Ekonomiczno-Socjologicznym UŁ, poprzez ocenę nierównomierności między wynikami egzaminów uzyskanymi przez studentów. Do analizy danych dotyczących wyników egzaminu końcowego z matematyki w latach 2018/2019 (tryb stacjonarny) i 2020/2021 (tryb zdalny) oraz porównawczo egzaminu maturalnego z matematyki zdanego przez badanych studentów zastosowano współczynnik Giniego oraz krzywą Lorentza.

Na podstawie wyznaczonych wartości współczynnika Giniego można stwierdzić, że efekty nauki stacjonarnej i zdalnej są podobne, co świadczy o utrzymaniu wysokiej jakości kształcenia przez uczelnię pomimo zmiany trybu pracy podczas pandemii COVID-19. Rozkład wyników matur z 2020 r. okazał się bardziej nierównomierny od rozkładu wyników matur z 2018 r. Statystyki opisowe, a w szczególności

asymetria lewostronna, świadczą o nieznacznej przewadze pozytywnych wyników egzaminu końcowego przeprowadzonego w roku akademickim, w którym zajęcia odbywały się w formie zdalnej; podobną przewagę zaobserwowano również dla egzaminu maturalnego, jednak różnice te nie są znaczące. W przypadku matury różnica wynosi ok. 0,1 p.proc., co stanowi 0,5 pkt z 50 pkt możliwych do uzyskania przez maturzystę. W przypadku egzaminu końcowego z matematyki na I roku studiów jest to ok. 9 p.proc., co przekłada się na 3 pkt możliwe do uzyskania przez studenta.

Nauka zdalna może wnieść nową jakość do kształcenia akademickiego. Zastosowanie jej na szeroką skalę w roku 2020/2021 wynikało z konieczności zmiany trybu kształcenia w związku z wybuchem pandemii COVID-19. Zajęcia e-learningowe stanowią jednak istotny element kształcenia studentów nie tylko w czasie pandemii, a rosnące oczekiwania ich uczestników i potencjał zaawansowanej technologii sprzyjają doskonaleniu tych technik nauczania. Dodatkowo osiągnięcie przez uczelnię wysokiej jakości e-learningu pozytywnie wpływa na jej postrzeganie i renomę w środowisku akademickim oraz w społeczeństwie.

Należy też zauważyć, że e-learning jest głównym sposobem zdobywania nowych umiejętności po ukończeniu studiów, ponieważ proces kształcenia przez całe życie bardzo trudno realizować w formie kursów tradycyjnych (stacjonarnych). Społeczeństwo informacyjne, a w szczególności rynek pracy, wymaga zaś ciągłego zdobywania nowych umiejętności i poznawania nowych dziedzin wiedzy.

Bibliografia

Alaghawat, M., Alshamailah, I. (2022). Distance Learning at the University of Jordan in the Setting of Covid-19 Pandemic from the Students' Perspectives. *International Journal of Emerging Technologies in Learning*, *17*(6). https://doi.org/10.3991/ijet.v17i06.27845.

Alirezabeigi, S., Masschelein, J., Decuypere, M. (2020). Investigating digital doings through breakdowns: A sociomaterial ethnography of a Bring Your Own Device school. *Learning, Media and Technology*, *45*(2), 193–207. https://doi.org/10.1080/17439884.2020.1727501.

Asgari, S., Trajkovic, J., Rahmani, M., Zhang, W., Lo, R. C., Sciortino, A. (2021). An observational study of engineering online education during the COVID-19 pandemic. *PLOS ONE*, *16*(4), 1–17. https://doi.org/10.1371/journal.pone.0250041.

Bahasoan, A., Ayuandiani, W., Mukhram, M., Rahmat, A. (2020). Effectiveness of Online Learning In Pandemic COVID-19. *International Journal of Science, Technology & Management*, *1*(2), 100–106. https://doi.org/10.46729/ijstm.v1i2.30.

Biernacki, M., Ejsmont, W. (2011). Efektywność kształcenia we wrocławskich liceach. *Acta Universitatis Lodziensis. Folia Oeconomica*, (253), 157–172.

Bonnini, S., Cavallo, G. (2021). A study on the satisfaction with distance learning of university students with disabilities: Bivariate regression analysis using a multiple permutation test. *Statistica Applicata – Italian Journal of Applied Statistics*, *33*(2), 143–162. https://doi.org/10.26398/IJAS.0033-008.

Butnaru, G. I., Niță, V., Anichiti, A., Brînză, G. (2021). The Effectiveness of Online Education during Covid 19 Pandemic – A Comparative Analysis between the Perceptions of Academic Students and High School Students from Romania. *Sustainability*, *13*(9), 1–20. https://doi.org/10.3390/su13095311.

Crespo-Cuaresma, J., Samir, K. C., Sauer, P. (2012). *Gini Coefficients of Educational Attainment: Age Group Specific Trends in Educational (In)equality*. https://paa2012.populationassociation.org/papers/121621.

Dąbrowski, M. (2013). E-learning w szkolnictwie wyższym. *Studia BAS*, (3), 203–212. https://orka.sejm.gov.pl/WydBAS.nsf/0/66B19F03A052497FC1257BDC0029FE66/$file/Strony%20odStudia_BAS_35i-10.pdf.

Gini, C. (1914). Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, *73*, 1203–1248.

Główny Urząd Statystyczny. (b.r.). *E-learning*. https://stat.gov.pl/metainformacje/slownik-pojec/pojecia-stosowane-w-statystyce-publicznej/1418,pojecie.html.

Hebebci, M. T., Bertiz, Y., Alan, S. (2020). Investigation of Views of Students and Teachers on Distance Education Practices during the Coronavirus (COVID-19) Pandemic. *International Journal of Technology in Education and Science*, *4*(4), 267–282. https://doi.org/10.46328/ijtes.v4i4.113.

Kopciał, P. (2013). Analiza metod e-learningowych stosowanych w kształceniu osób dorosłych. *Zeszyty Naukowe Warszawskiej Wyższej Szkoły Informatyki*, *7*(9), 79–99. https://doi.org/10.26348/znwwsi.9.79.

Lorenz, M. O. (1905). Methods of Measuring the Concentration of Wealth. *Publications of the American Statistical Association*, *9*(70), 209–219. https://doi.org/10.2307/2276207.

Maleńczyk, I., Gładysz, B. (2019). Academic E-learning in Poland: Results of a Diagnostic Survey. *International Journal of Research in E-learning*, *5*(1), 35–59. https://doi.org/10.31261/IJREL.2019.5.1.03.

Mohammed Ali, A. (2021). E-learning for Students With Disabilities During COVID-19: Faculty Attitude and Perception. *SAGE Open*, *11*(4). https://doi.org/10.1177/21582440211054494.

Romaniuk, M. W. (2015). E-learning in College on the Example of Academy of Special Education. *International Journal of Electronics and Telecommunications*, *61*(1), 25–29. https://doi.org/10.1515/eletel-2015-0003.

Rosthal, R. A. (1978). *Measures of Disparity. A Note. Research Report*. Killalea Associates.

Ryl-Zaleska, M. (2022). *Metody oceny efektywności kształcenia online*. http://e-edukacja.fundacja.edu.pl/_referaty/9_e-edukacja.pdf.

Sahu, P. (2019). Closure of Universities Due to Coronavirus Disease 2019 (COVID-19): Impact on Education and Mental Health of Students and Academic Staff. *Cureus*, *12*(4), 1–6. https://doi.org/10.7759/cureus.7541.

Sheret, M. (1988). Equality Trends and Comparisons for the Education System of Papua New Guinea. *Studies in Educational Evaluation*, *14*(1), 91–112.

Strielkowski, W. (2020). COVID-19 pandemic and the digital revolution in academia and higher education. *Preprints*, *1*, 1–6. https://doi.org/10.20944/preprints202004.0290.v1.

Szymańska, A., Zalewska, E. (2015). E-learning as a tool to improve the quality of education in quantitative subjects. *Didactics of Mathematics*, *12*(16), 125–134. https://doi.org/10.15611/dm.2015.12.13.

Thomas, V., Wang, Y., Fan, X. (2001). *Measuring Education Inequality: Gini Coefficients of Education* (World Bank Policy Research Working Paper No. 2525). https://openknowledge.worldbank.org/bitstream/handle/10986/19738/multi_page.pdf?sequence=1&isAllowed=y.

Thomas, V., Wang, Y., Fan, X. (2002). *A New Dataset on Inequality in Education. Gini an Theil Indices of Schooling for 140 Countries, 1960–2000*. The World Bank.

Topol, P. (2020). Metody i narzędzia kształcenia zdalnego w polskich uczelniach w czasie pandemii COVID-19 – Część 1, Dyskusja 2020. *Studia Edukacyjne*, (58), 69–83. https://pressto.amu.edu.pl/index.php/se/article/view/26007/23736.

Trzcińska, K. (2021). Kształtowanie się rozkładu dochodów ludności Polski dla regionów na podstawie wybranych modeli teoretycznych. *Acta Universitatis Lodziensis. Folia Oeconomica*, *1*(352), 111–126. https://doi.org/10.18778/0208-6018.352.06.

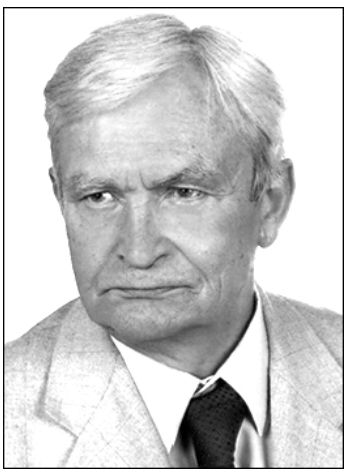
Zalewska, E. (2015). Jakość kursów e-learning. W: P. Wdowiński (red.), *Nauczyciel akademicki wobec nowych wyzwań edukacyjnych* (s. 105–113). Wydawnictwo Uniwersytetu Łódzkiego.

Zalewska, E. (2021). The Application of Continuous Quality Improvement Methods at Universities in the Opinion of Students and Lecturers of the University of Lodz. *Folia Oeconomica Stetinensia*, *21*(1), 175–190. https://doi.org/10.2478/foli-2021-0012.

Ziesemer, T. (2011). *What Changes Gini Coefficients of Education? On the dynamic interaction between education, its distribution and growth* (UNU-MERIT Working Papers No. 2011-053). http://collections.unu.edu/view/UNU:200#viewAttachments.

Wiadomości Statystyczne. The Polish Statistician, 2022, vol. 67, 10, 62–64
In memoriam

Profesor Krzysztof Zadora (1943–2021)



Fot. Archiwum UE w Katowicach

Prof. dr hab. Krzysztof Zadora urodził się 22 sierpnia 1943 r. w Chorzowie. Studiował w Wyższej Szkole Ekonomicznej w Katowicach (obecnie Uniwersytet Ekonomiczny w Katowicach), gdzie w 1967 r. uzyskał stopień zawodowy magistra. Po studiach podjął na tej uczelni pracę jako asystent, a potem starszy asystent w Katedrze Statystyki na Wydziale Przemysłu. W 1971 r. obronił rozprawę doktorską z dziedziny ekonometrii pt. *Metody konstrukcji prognoz krótkoterminowych na podstawie szeregów czasowych*, napisaną pod kierunkiem jednego z pionierów polskiej ekonometrii prof. Zbigniewa Pawłowskiego. W tym samym roku został mianowany adiunktem w Instytucie Ekonometrii i zajmował to stanowisko do 1978 r.

W tamtym czasie Instytut Ekonometrii, kierowany przez prof. Pawłowskiego, przeżywał rozkwit. Było tam zatrudnionych ok. 50 osób, głównie młodych i ambitnych naukowców. Podjęli się oni, zainicjowanego przez prof. Pawłowskiego, zadania stworzenia nowego kierunku studiów jednolitych o nazwie cybernetyka ekonomiczna i informatyka (pierwowzór takich kierunków, jak analityka gospodarcza oraz informatyka i ekonometria). Adiunkt Zadora prowadził wówczas zajęcia z cybernetyki ekonomicznej i procesów stochastycznych. Były to nowe przedmioty, więc na wykłady razem ze studentami uczęszczali młodzi pracownicy Instytutu. Począwszy od 1965 r., corocznie odbywała się (i odbywa się do dzisiaj) Konferencja Statystyków, Ekonometryków i Matematyków Uniwersytetów Ekonomicznych Polski Południowej, dająca możliwość poznania nowinek z zakresu zastosowań metod ilościowych w ekonomii i zarządzaniu. Krzysztof Zadora należał do najaktywniejszych uczestników konferencji.

Od czasu obrony doktoratu zainteresowania naukowe prof. Zadory skupiały się wokół ekonometrii. Ich owocem była również habilitacja, którą uzyskał w 1981 r. na podstawie książki *Ekonometryczna ekstrapolacja szeregów czasowych* (Wydawnictwo Akademii Ekonomicznej w Katowicach, 1980). Jako docent podjął się kierowania rozwojem młodszej kadry naukowo-dydaktycznej – został kierownikiem Zakładu Ekonometrii (1981–1984), a potem kierownikiem Zakładu Matematyki i Programowania (do 1993 r.). Ponadto w latach 1981–1987 był zastępcą dyrektora Instytutu

© Janusz L. Wywiał
Artykuł udostępniony na licencji CC BY-SA 4.0 / Article available under the CC BY-SA 4.0 licence

Ekonometrii ds. nauki, a w latach 1984–1990 – docentem etatowym. Tytuł naukowy profesora uzyskał 18 lutego 1992 r.

Profesor Zadora brał udział w pracach zespołu pod kierunkiem dr hab. Elżbiety Stolarskiej, prof. UE, realizującego projekt resortowy MR.I.30 dotyczący konstrukcji, estymacji i weryfikacji modelu polskiej gospodarki. Rezultaty tych prac zostały opublikowane w monografii pod redakcją Elżbiety Stolarskiej pt. *Ekonometryczny model gospodarki Polski – MES 1* (Wydawnictwo Akademii Ekonomicznej w Katowicach, 1987). W latach 1982–1990 sam kierował dużymi zespołami naukowo-badawczymi złożonymi z pracowników Instytutu, które realizowały projekty finansowane przez Ministerstwo Szkolnictwa. Pierwszym takim przedsięwzięciem były *Klasyczne i niekonwencjonalne metody predykcji* (1982–1985, w ramach projektu resortowego R.III.9), drugim – *Ekonometryczna analiza zmian struktury* (1986–1990, jako część projektu w ramach Centralnego Programu Badań Podstawowych CPBP 10.09). Ich wyniki istotnie przyczyniły się do rozwoju naukowego uczestniczących w nich badaczy; powstało też wiele wartościowych prac naukowych i publikacji. Warto w tym miejscu wymienić ważniejsze prace prof. Zadory z tego okresu: *O pewnym zagadnieniu aproksymacji i jego zastosowaniu* („Przegląd Statystyczny", 1982), *Metoda najmniejszych modułów w zastosowaniu do estymacji liniowych jednorównaniowych modeli ekonometrycznych* („Zeszyty Naukowe Akademii Ekonomicznej w Krakowie", 1992) i *Specyfikacja struktury ekonometrycznych modeli systemów sterowanych* (rozdział w monografii napisanej wspólnie ze Zbigniewem Czerwińskim, Wojciechem Maciejewskim i Antonim Smolukiem pt. *Ekonometria. Nadzieje, osiągnięcia, niedostatki*, 1987).

Podczas realizacji projektów naukowo-badawczych prof. Zadora nie narzucał swoich pomysłów wykonawcom, wręcz przeciwnie – oczekiwał od nich formułowania własnych pomysłów dotyczących poszczególnych aspektów projektów. Takie podejście pozwalało wykazać się wiedzą, inspirowało do jej poszerzania i poszukiwania nowych oryginalnych rozwiązań.

Warto zauważyć, że niektóre rezultaty prac prof. Zadory z dziedziny ekonometrii klasycznej, predykcji ekonometrycznej i statystyki wciąż inspirują do nowych badań. Chodzi zwłaszcza o poszukiwanie możliwości modelowania ekonometrycznego w warunkach, gdy nie są spełnione klasyczne założenia normalnej regresji liniowej. Profesor miał wiele wątpliwości co do wnioskowania na podstawie oszacowanych modeli ekonometrycznych w sytuacji, gdy pewne założenia nie są spełnione lub są nieweryfikowalne. Czas pokazał, że klasyczne modele ekonometryczne zastępuje się innymi podejściami do modelowania, takimi jak sieci neuronowe czy uczenie maszynowe. Kontestacyjny stosunek prof. Zadory do możliwości wnioskowania na podstawie modeli ekonometrycznych przerodził się z czasem w sceptyczne trakto-

wanie niektórych teorii ekonomicznych, m.in. koncepcji głoszących, że prawa ekonomiczne są w pełni weryfikowalne w badaniach empirycznych. Do zajęcia takiego stanowiska przyczyniło się przestudiowanie monografii Hansa-Hermanna Hoppego *Economic Science and the Austrian Method*. W późniejszym czasie wspomniane wątpliwości sprawiły, że Profesor częściowo zmienił swoje zainteresowania. Przewidywał zmierzch powszechnego używania klasycznych modeli ekonometrycznych i poszukiwał bardziej realnych możliwości wykorzystania m.in. metod ilościowych w praktyce funkcjonowania gospodarki, a w szczególności – przedsiębiorstwa. Możliwości mikroekonometrii w tym zakresie wydawały się niewystarczające, dlatego zaczął interesować się inną dziedziną ekonomii – rachunkowością, a zwłaszcza rachunkowością zarządczą.

W latach 80. i na początku lat 90. XX w., gdy w kraju dokonywały się przemiany polityczne, społeczne i gospodarcze, prof. Zadora zaangażował się w działania reformatorskie mające na celu wzmacnianie niezależności politycznej Polski, jak również stawianie czoła nowym wyzwaniom ekonomicznym i społecznym. Należał do ludzi o radykalnych poglądach, ale był zwolennikiem stopniowej i ostrożnej modernizacji, także w środowisku akademickim. Nie zaniedbywał przy tym swoich obowiązków naukowo-dydaktycznych. W latach 1990–1993 był rektorem Akademii Ekonomicznej w Katowicach. Razem z prorektor ds. nauki dr hab. Elżbietą Stolarską i prorektor ds. dydaktyki dr hab. Joanną Żabińską sukcesywnie wprowadzał potrzebne zmiany w funkcjonowaniu uczelni, zgodne z ówczesnym duchem reformatorskim. Co ważne, nie dopuścił do skłócenia społeczności akademickiej uczelni.

Od 1993 r. pracował jako profesor Akademii Ekonomicznej w Katedrze Rachunkowości Wydziału Ekonomii, a potem Wydziału Finansów i Ubezpieczeń, gdzie w latach 2010–2015 był zatrudniony na stanowisku profesora zwyczajnego. W latach 1993–2015 był kierownikiem Zakładu Rachunkowości Zarządczej. Działał również w Stowarzyszeniu Księgowych w Polsce.

Dorobek publikacyjny prof. Zadory liczy kilkadziesiąt pozycji z dziedziny ekonometrii oraz rachunkowości. Za swoją działalność naukowo-badawczą i dydaktyczną Profesor pięciokrotnie otrzymał Nagrodę Rektora Akademii Ekonomicznej w Katowicach (w latach 1997–2001). Wypromował 10 doktorów oraz ok. 400 magistrów. Zwieńczeniem jego zasług na polu naukowo-dydaktycznym oraz organizacyjnym było przyznanie mu w 1994 r. Krzyża Kawalerskiego Orderu Odrodzenia Polski.

Profesor Krzysztof Zadora zmarł 18 września 2021 r.

Janusz L. Wywiał

Wiadomości Statystyczne. The Polish Statistician, 2022, vol. 67, 10, 65–67
Dyskusje. Recenzje. Informacje / Discussions. Reviews. Information

WYDAWNICTWA GUS. WRZESIEŃ 2022
PUBLICATIONS OF STATISTICS POLAND. SEPTEMBER 2022

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: *Kultura i dziedzictwo narodowe w 2021 r.*
Title: *Culture and national heritage in 2021*
Język: polski, angielski
Language: Polish, English
Dodatkowe informacje: opracowanie dostępne w wersji drukowanej i elektronicznej
Additional information: publication available both in the printed and in the electronic form

Kontynuacja corocznych opracowań pt. *Kultura*, wspólna publikacja GUS i Urzędu Statystycznego w Krakowie, powstała we współpracy z Narodowym Instytutem Dziedzictwa, Narodowym Instytutem Muzealnictwa i Ochrony Zbiorów oraz Naczelną Dyrekcją Archiwów Państwowych; od 2022 r. włączona do serii Analizy statystyczne. Przedstawia wybrane dane charakteryzujące działalność podmiotów w obszarze kultury i dziedzictwa narodowego w Polsce oraz wskaźniki dotyczące uczestnictwa w przedsięwzięciach kulturalnych.

Publikacja składa się z dwóch części: analitycznej (z uwagami metodologicznymi) i tabelarycznej (dostępnej wyłącznie w formie elektronicznej). Część analityczna zawiera informacje na temat ochrony zabytków, funkcjonowania archiwów państwowych, bibliotek, działalności wydawniczej, centrów kultury, domów i ośrodków kultury, klubów i świetlic, szkolnictwa artystycznego, wystawiennictwa, galerii sztuki, działalności scenicznej, kinematografii, radiofonii i telewizji oraz imprez masowych, wydatków na kulturę i ochronę dziedzictwa narodowego oraz wydatków na kulturę w gospodarstwach domowych w 2021 r., a także funkcjonowania przemysłów kultury i kreatywnych w 2020 r. Ponadto zaprezentowano w niej rozkład terytorialny podmiotów działających w obszarze kultury w 2021 r., a niektóre zjawiska z zakresu kultury przedstawiono w ujęciu powiatowym.

© Główny Urząd Statystyczny
Artykuł udostępniony na licencji CC BY-SA 4.0 / Article available under the CC BY-SA 4.0 licence

Tytuł: *Nowoczesne technologie i nowe źródła danych w pomiarze inflacji*
Title: *Modern technologies and new sources of data in inflation measurement*
Język: polski
Language: Polish
Dodatkowe informacje: opracowanie dostępne w wersji drukowanej i elektronicznej
Additional information: publication available both in the printed and in the electronic form

Opracowanie monograficzne prezentujące i podsumowujące prace prowadzone w ramach projektu „Budowa zintegrowanego systemu statystyki cen detalicznych – INSTATCENY", zrealizowanego przez konsorcjum naukowe zawiązane przez GUS ze Szkołą Główną Handlową w Warszawie i z Instytutem Podstaw Informatyki Polskiej Akademii Nauk.

W monografii opisano metodykę statystycznych badań cen na rynku detalicznym i działania podjęte w celu jej modernizacji. Ujęto w niej wszystkie najważniejsze zagadnienia dotyczące produkcji danych statystycznych w tym obszarze. Przedstawiono wyniki prac związanych z uzyskiwaniem i przetwarzaniem danych pochodzących z nowych, alternatywnych źródeł, takich jak dane dostarczane przez sieci handlowe oraz pobierane automatycznie ze stron internetowych sklepów, i wdrożeniem ich do praktyki badań cen detalicznych.

We wrześniu br. ukazały się ponadto:
- *Bezrobocie rejestrowane 1–2 kwartał 2022 r.*;
- „Biuletyn statystyczny" nr 8/2022;
- *Budżety gospodarstw domowych w 2021 r.*;
- *Ceny robót budowlano-montażowych i obiektów budowlanych (lipiec 2022 r.)*;
- *Fizyczne rozmiary produkcji zwierzęcej w 2021 r.*;
- *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach 2000–2022 (wrzesień 2022 r.)*;
- *Nakłady i wyniki przemysłu – 1–2 kwartał 2022 r.*;
- *Pomoc społeczna i opieka nad dzieckiem i rodziną w 2021 r.*;
- *Produkcja ważniejszych wyrobów przemysłowych w sierpniu 2022 r.*;
- *Równoległy oraz wyprzedzający zagregowany wskaźnik koniunktury, zegar koniunktury – szereg do lipca 2022 r.*;
- „Statistics in Transitions new series" nr 3/2022;

- *Sytuacja makroekonomiczna w Polsce na tle procesów w gospodarce światowej w 2021 r.*;
- *Sytuacja społeczno-gospodarcza kraju w sierpniu 2022 r.*;
- *Sytuacja społeczno-gospodarcza województw nr 2/2022*;
- *Telekomunikacja w 2021 r.*;
- *Transport – wyniki działalności w 2021 r.*;
- „Wiadomości Statystyczne. The Polish Statistician" nr 9/2022;
- *Zatrudnienie i wynagrodzenia w gospodarce narodowej w pierwszym półroczu 2022 r.*;
- *Zeszyt metodologiczny. Działalność badawcza i rozwojowa.*

Wszystkie publikacje GUS w wersji elektronicznej są dostępne na stronie stat.gov.pl/publikacje/publikacje-a-z.
Electronic forms of all the publications by Statistics Poland are available at stat.gov.pl/en/publications.

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

DLA AUTORÓW
FOR THE AUTHORS

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

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

Od 2007 r. „WS" znajdują się na liście polskich punktowanych czasopism naukowych MEiN. Zgodnie z komunikatem Ministra Edukacji i Nauki z dnia 1 grudnia 2021 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych „WS" otrzymały 70 punktów.

„Wiadomości Statystyczne. The Polish Statistician" są udostępniane w następujących bazach, repozytoriach, katalogach i wyszukiwarkach: Agro, BazEkon, Central and Eastern European Academic Source (CEEAS), Central and Eastern European Online Library (CEEOL), Central European Journal of Social Sciences and Humanities (CEJSH), Directory of Open Access Journals (DOAJ), EBSCO Discovery Service, European Reference Index for the Humanities and Social Sciences (ERIH Plus), Exlibris Primo, Google Scholar, ICI Journals Master List, ICI World of Journals, Norwegian Register for Scientific Journals and Publishers (The Nordic List) oraz Summon.

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

1. Zgłaszanie artykułów

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

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

Jeżeli w pracy występują tablice, wykresy lub mapy, powinny być umieszczone w treści artykułu. Ponadto należy je przesłać osobno (najlepiej w formacie Excel); szczegółowe informacje są zawarte w pkt 4.2.14.

Autor jest zobowiązany do podania w artykule wszelkich źródeł finansowania badań będących podstawą pracy. Jeżeli doszło do zaprezentowania podobnych materiałów podczas konferencji lub sympozjum naukowego albo jeżeli artykuł został przez autora umieszczony w repozytorium internetowym lub zgłoszony w innym wydawnictwie do opublikowania w innej wersji językowej, to podczas składania tekstu do publikacji w „WS" należy poinformować o tym redakcję.

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

Razem z artykułem należy przesłać skan oświadczenia (do pobrania ze strony internetowej czasopisma) o oryginalności pracy, niezłożeniu jej w innym wydawnictwie i udzieleniu wydawcy „WS" licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0), zawierającego numer ORCID, afiliację lub afiliacje oraz dane kontaktowe autora, wraz ze wskazaniem proponowanego działu czasopisma.

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

2. Przebieg prac redakcyjnych

Zgłoszony artykuł jest oceniany i opracowywany w czteroetapowym procesie:

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

   Autorzy artykułów, które otrzymały pozytywne oceny, wprowadzają poprawki zalecane przez recenzentów i przesyłają do redakcji zmodyfikowaną wersję pracy. Jeśli pojawi się różnica zdań dotycząca zasadności proponowanych zmian, autorzy są zobligowani do uzasadnienia swojego stanowiska.
3. **Ocena Kolegium Redakcyjnego** (KR), decydująca o przyjęciu pracy do publikacji. Jest dokonywana na podstawie recenzji, z uwzględnieniem opinii redaktorów tematycznego i merytorycznego. Polega m.in. na weryfikacji dokonania przez autora zmian w artykule stosownie do uwag recenzentów. KR ocenia artykuł pod względem poprawności i spójności merytorycznej oraz zaleca autorowi wprowadzenie poprawek, jeśli są one konieczne, aby praca spełniała wymogi czasopisma. Autorowi przysługuje prawo do odwołania od decyzji o niepublikowaniu artykułu. W takim przypadku powinien on skontaktować się z redakcją „WS" i przedstawić uzasadnienie. Ostateczna decyzja w tej sprawie należy do redaktora naczelnego.

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

   Artykuły przyjęte przez KR do publikacji są zamieszczane na stronie internetowej czasopisma w zakładce Early View, gdzie znajdują się do czasu opublikowania w konkretnym wydaniu „WS".
4. **Opracowanie redakcyjne, autoryzacja i korekta.** Artykuł zakwalifikowany do druku jest poddawany opracowaniu merytorycznemu i językowemu. Redakcja zastrzega sobie prawo

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

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

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

3. Zasady etyki publikacyjnej COPE

Redakcja „WS" podejmuje wszelkie starania w celu utrzymania najwyższych standardów etycznych zgodnie z wytycznymi Komitetu ds. Etyki Publikacyjnej (COPE), dostępnymi na stronie internetowej www.publicationethics.org, oraz wykorzystuje wszystkie możliwe środki mające na celu zapobieżenie nadużyciom i nierzetelności autorskiej. Przyjęte zasady postępowania obowiązują autorów, Radę Naukową, Kolegium Redakcyjne, redakcję, pracowników Wydziału Czasopism Naukowych GUS, 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 opisania 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 opracowywaniu artykułu;
   - autorstwo grzecznościowe (*gift authorship*) – podawanie jako współautorów osób, których wkład jest oparty jedynie na słabym powiązaniu z badaniem.

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

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

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

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

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

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

3.3. Odpowiedzialność recenzentów

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

2. Recenzenci są zobligowani do zachowania obiektywności i poufności oraz powstrzymania się od osobistej krytyki. Zawsze powinni uzasadnić swoją ocenę, przedstawiając stosowną argumentację.
3. Recenzenci powinni wskazać ważne dla wyników badań opublikowane prace, które w ich ocenie powinny zostać przywołane w ocenianym artykule.
4. W razie stwierdzenia wysokiego poziomu zbieżności treści recenzowanej pracy z innymi opublikowanymi materiałami lub podejrzenia innych przejawów nierzetelności naukowej recenzenci są zobowiązani poinformować o tym redakcję.
5. Po ukończeniu recenzji przechowywanie przesłanych przez redakcję materiałów (w 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 w trybie otwartego dostępu, tj. bezpłatnie i bez technicznych ograniczeń, od 1 stycznia 2022 r. na licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0). W przypadku artykułów zgłoszonych do „WS" od 2022 r. dozwolone jest dzielenie się artykułem (kopiowanie i rozpowszechnianie go w dowolnym medium i formacie) oraz adaptowanie go (w dowolnym celu, także komercyjnym) na warunkach określonych w tej licencji. Z pozostałych artykułów zamieszczonych w czasopiśmie można korzystać w ramach otwartego dostępu, zgodnie z ustawą o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego.
3. Wydawca deklaruje gotowość do opublikowania poprawek, wyjaśnień oraz przeprosin.

4. Wymogi redakcyjne

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

4.1. Struktura i zawartość artykułu

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

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

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

   Wszystkie części powinny być opatrzone numerami.
8. Bibliografia, zawierająca pełny wykaz prac i materiałów przywołanych w artykule, przygotowana zgodnie z wymogami czasopisma.

4.2. Przygotowanie artykułu

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

9. Interlinia – 1,5 wiersza; tablice i przypisy – 1 wiersz; przed tytułami rozdziałów i podrozdziałów oraz po nich – pusty wiersz.
10. Wcięcie akapitowe – 0,4 cm; bibliografia – bez wcięcia, wysunięcie 0,4 cm.
11. Przy wyliczeniach należy posłużyć się listą punktowaną z punktorami w postaci kropek (wysunięcie 0,4 cm, wcięcie 0 cm); wiersze (oprócz ostatniego) zakończone średnikiem.
12. Strony ponumerowane automatycznie.
13. Tablice i elementy graficzne (wykresy, mapy, schematy) muszą być przywołane w tekście.
14. Wykresy, mapy i schematy należy zamieścić w tekście głównym. Wykresy powinny być edytowalne (optymalnie wykonane w programie Excel; w przypadku wykonania w programie graficznym powinny mieć postać wektorową). Wykresy i inne materiały graficzne należy przekazać osobno w pliku programu Excel lub innym edytowalnym w pakiecie Microsoft Office. W przypadku wykonania ich w innym programie odpowiedni format pliku będzie ustalany indywidualnie.
15. Tablice muszą być edytowalne. Nie należy stosować rastrów, cieniowania, pogrubiania czy też podwójnych linii itp.
16. Wskazówki dotyczące opracowywania map znajdują się w publikacji *Mapy statystyczne. Opracowanie i prezentacja danych*, dostępnej na stronie internetowej GUS.
17. Pod tablicami i każdym elementem graficznym należy podać źródło opracowania, a także objaśnić użyte w nich skróty i symbole.
18. Literowe symbole liczb i innych wielkości niezłożonych należy zapisywać małą lub dużą literą i pismem pochyłym (np. a, A, $y(x)$, a_i); wektorów – pismem pochyłym i pogrubionym (np. $\boldsymbol{a}$, $\boldsymbol{A}$, $\boldsymbol{w}$, $\boldsymbol{y}(x)$, $\boldsymbol{w}_i$); macierzy – pismem prostym i pogrubionym (np. $\mathbf{A}$, $\mathbf{a}$, $\mathbf{M}$, $\mathbf{Y}(x)$, $\mathbf{M}_i$).
19. Objaśnienia znaków umownych w tablicach: kreska (–) – zjawisko nie wystąpiło; zero (0) – zjawisko istniało w wielkości mniejszej od 0,5; (0,0) – zjawisko istniało w wielkości mniejszej od 0,05; kropka (.) – brak informacji, konieczność zachowania tajemnicy statystycznej, wypełnienie pozycji jest niemożliwe lub niecelowe; „w tym" – oznacza, że nie podaje się wszystkich składników sumy.
20. Stosowane są następujące skróty: tablica – tabl., wykres – wykr.
21. Wszystkie zawarte w artykule informacje, dane i stwierdzenia wykraczające poza wiedzę powszechną – np. wyniki badań innych autorów, zarówno o charakterze empirycznym, jak i koncepcyjnym – muszą być opatrzone przypisem bibliograficznym. Przez wiedzę powszechną należy rozumieć informacje ogólnie znane i niebudzące wątpliwości ani kontrowersji w danej grupie społecznej, np. utworzenie GUS w 1918 r. lub powstanie UE w 1993 r. na podstawie traktatu z Maastricht. Natomiast dane statystyczne udostępniane lub publikowane np. przez GUS lub Eurostat nie należą do takich informacji. Charakteru wiedzy powszechnej nie mają również stwierdzenia odnoszące się do idei, zjawisk i procesów społecznych, politycznych czy gospodarczych. Nawet pozornie zdroworozsądkowe idee zmieniają bowiem swój sens w zależności od kultury, języka lub dyscypliny naukowej, a także bywają w rozmaity sposób konceptualizowane, jak np. pojęcie poznania w naukach społecznych.

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

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

4.3. Zasady przywoływania publikacji w treści artykułu

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

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

4.4. Przykłady opisu bibliograficznego

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

Typ publikacji	Przykład opisu bibliograficznego
Artykuł w czasopiśmie	
W wersji drukowanej	Nazwisko, X. (rok). Tytuł artykułu. *Tytuł czasopisma*, *rocznik* (zeszyt), strona początku–strona końca.
Dostępny w internecie, z DOI	Nazwisko, X., Nazwisko 2, Y. (rok). Tytuł artykułu. *Tytuł czasopisma*, *rocznik*(zeszyt), strona początku–strona końca. https://doi.org/xxx.
Dostępny w internecie, bez DOI	Nazwisko, X., Nazwisko 2, Y., Nazwisko 3, Z. (rok). Tytuł artykułu. *Tytuł czasopisma*, *rocznik*(zeszyt), strona początku–strona końca. https://xxx.
Maszynopis	
Niepublikowany / przygotowywany przez autora / zgłoszony do publikacji, ale jeszcze niezaakceptowany	Nazwisko, X. (rok). *Tytuł* [maszynopis niepublikowany / w przygotowaniu / zgłoszony do publikacji].
Zaakceptowany do publikacji	Nazwisko, X. (w druku). Tytuł artykułu. *Tytuł czasopisma.*
Opublikowany nieformalnie (np. na stronie internetowej autora)	Nazwisko, X., Nazwisko 2, Y. (rok). *Tytuł artykułu*. https://xxx.
Opublikowany w trybie early view / ahead of print / online first	Nazwisko, X. (rok). Tytuł artykułu. *Tytuł czasopisma.* Early view / Ahead of print / Online first. https://xxx.
Książka	
W wersji drukowanej	Nazwisko, X. (rok). *Tytuł książki*. Wydawnictwo.
Dostępna w internecie, z DOI	Nazwisko, X. (rok). *Tytuł książki*. Wydawnictwo. https://doi.org/xxx.
Dostępna w internecie, bez DOI	Nazwisko, X. (rok). *Tytuł książki*. Wydawnictwo. https://xxx.
W przekładzie	Nazwisko, X. (rok). *Tytuł książki* (tłum. Y. Nazwisko). Wydawnictwo.
Wydanie wielotomowe: tom zatytułowany	Nazwisko, X. (rok). *Tytuł książki: nr tomu. Tytuł tomu*. Wydawnictwo.
tom niezatytułowany	Nazwisko, X. (rok). *Tytuł książki* (nr tomu). Wydawnictwo.
Kolejne wydanie	Nazwisko, X. (rok). *Tytuł książki* (nr wydania). Wydawnictwo.
Pod redakcją: w języku polskim	Nazwisko, X. (red.). (rok). *Tytuł książki*. Wydawnictwo.
w języku angielskim	Nazwisko, X. (Ed.). (rok). *Tytuł książki*. Wydawnictwo.
Rozdział w pracy zbiorowej	Nazwisko, X. (rok). Tytuł rozdziału. W: Y. Nazwisko, Z. Nazwisko 2 (red.), *Tytuł książki* (s. strona początku–strona końca). Wydawnictwo. https://doi.org/xxx lub https://xxx.
Inne prace	
Raport: autor indywidualny	Nazwisko, X. (rok). *Tytuł raportu*. Wydawnictwo.
autor instytucjonalny	Nazwa instytucji. (rok). *Tytuł raportu*. Wydawnictwo.
Working Papers	Nazwisko, X. (rok). *Tytuł pracy* (nazwa serii i numer). https://doi.org/xxx lub https://xxx.
Sesja konferencyjna / prezentacja / referat	Nazwisko, X. (rok, dzień i miesiąc). *Tytuł pracy* [typ wystąpienia, np. referat]. Nazwa konferencji, miejsce konferencji.
Rozprawa doktorska: nieopublikowana	Nazwisko, X. (rok). *Tytuł pracy* [niepublikowana rozprawa doktorska]. Nazwa instytucji nadającej tytuł doktorski.
opublikowana	Nazwisko, X. (rok). *Tytuł pracy* [rozprawa doktorska, nazwa instytucji nadającej tytuł doktorski]. https://xxx.
Akt prawny	Pełny tytuł aktu prawnego wraz z datą publikacji w dzienniku urzędowym.

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

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

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

DZIAŁY „WS" – TEMATYKA ARTYKUŁÓW
WS SECTIONS – TOPICS OF THE ARTICLES

Pełny opis zakresu tematycznego działów: ws.stat.gov.pl/AimScope
Description of the topics covered in each section: ws.stat.gov.pl/AimScope

Studia metodologiczne / Methodological studies

- Oryginalne teoretyczne rozwiązania metodologiczne ze wskazaniem ich praktycznej użyteczności
- Prace przeglądowe i porównawcze oraz dotyczące etyki w statystyce, które wnoszą pionierski wkład poznawczy do obecnego stanu wiedzy

Statystyka w praktyce / Statistics in practice

- Nowatorskie zastosowania narzędzi i modeli statystycznych oraz analiza i ocena statystyczna zjawisk społeczno-ekonomicznych i innych, prowadzona w szczególności na danych z zasobów statystyki publicznej
- Wykorzystanie narzędzi informatycznych do uzyskiwania i przetwarzania informacji statystycznych, naliczania danych wynikowych, ich prezentacji i rozpowszechniania
- Projektowanie badań statystycznych, uzyskiwanie, integracja i przetwarzanie danych oraz generowanie wynikowych informacji statystycznych i kontrola ich ujawniania

Studia interdyscyplinarne. Wyzwania badawcze / Interdisciplinary studies. Research challenges

- Wyzwania badawcze wynikające z rosnących potrzeb użytkowników danych statystycznych i wymagające zaangażowania znacznych środków oraz rozwiązań z różnych dziedzin nauki i techniki
- Wykorzystanie technologii informacyjnych i komunikacyjnych, innowacyjność, przetwarzanie i analiza zagadnień związanych z data science i big data
- Wyniki badań prowadzonych przez przedstawicieli dyscyplin innych niż statystyka z wykorzystaniem metod statystycznych

Spisy powszechne – problemy i wyzwania / Issues and challenges in census taking

- Propozycje rozwiązań – zarówno organizacyjnych, jak i metodologicznych – możliwych do zastosowania w spisach oraz rezultaty analiz danych spisowych
- Praktyczne aspekty związane z gromadzeniem i udostępnianiem danych ze spisów, w tym dotyczące obciążenia odpowiedzi i ochrony tajemnicy statystycznej

Edukacja statystyczna / Statistical education

- Metody i efekty nauczania statystyki oraz popularyzacja myślenia statystycznego i rzetelnego posługiwania się informacjami statystycznymi
- Problemy związane z kształceniem w zakresie umiejętności stosowania statystyki na wszystkich poziomach edukacji, a także dotyczące wykorzystywania nowoczesnych koncepcji i metod dydaktycznych oraz pomocy naukowych w nauczaniu statystyki

Z dziejów statystyki / From the history of statistics

- Historia prowadzenia obserwacji statystycznych oraz rozwoju ich metodologii i narzędzi
- Życie i osiągnięcia zawodowe wybitnych statystyków, jak również działalność najważniejszych instytucji i organizacji statystycznych w Polsce i za granicą

In memoriam

- Nekrologi i artykuły wspomnieniowe

Informacje. Recenzje. Dyskusje / Discussions. Reviews. Information

- Teksty nierecenzowane i niemające charakteru artykułów naukowych: sprawozdania z konferencji naukowych i innych wydarzeń dotyczących statystyki, recenzje książek, omówienia nowości wydawniczych GUS, polemiki i dyskusje