

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 65

2020 | 10

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

POLSKIE TOWARZYSTWO STATYSTYCZNE
POLISH STATISTICAL ASSOCIATION



WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

PAŹDZIERNIK / OCTOBER
ROK / VOLUME 65

2020 | 10 (713)

RADA NAUKOWA / SCIENTIFIC COUNCIL

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

sekretarz/secretary: Paulina Kucharska-Singh

KOLEGIUM REDAKCYJNE / EDITORIAL BOARD

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

ZESPÓŁ REDAKCYJNY / EDITORIAL STAFF

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

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

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

redaktor merytoryczny / substantive editor: Wioletta Wrzaszcz

sekretarz/secretary: Małgorzata Zygmunt

ADRES REDAKCJI / EDITORIAL OFFICE ADDRESS

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

Redakcja językowa: Wydział Czasopism Naukowych, Główny Urząd Statystyczny

Language editing: Scientific Journals Division, Statistics Poland

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

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



Zakład Wydawnictw
Statystycznych

Druk i oprawa / Printed and bound:

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

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

© Copyright by Główny Urząd Statystyczny

Indeks 381306

Informacje w sprawie nabywania czasopism / Information on purchasing of the journal:

Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment

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

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

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

SPIS TREŚCI

CONTENTS

Od redakcji	4
From the editorial team	
Statystyka w praktyce	
Statistics in practice	
Piotr Matuszak	
Temporary part-time employment and future wages of male employees – a panel data analysis of the German labour market	7
Praca tymczasowa w niepełnym wymiarze a przyszłe wynagrodzenie mężczyzn – analiza panelowa niemieckiego rynku pracy	
Jan Zawadzki	
Prognozowanie brakujących danych w szeregach czasowych przy zastosowaniu modeli hybrydowych – podejście teoretyczne i empiryczne	24
Forecasting missing data in time series with the application of hybrid models – theoretical and empirical approach	
Studia interdyscyplinarne. Wyzwania badawcze	
Interdisciplinary studies. Research challenges	
Kornelia Przedworska	
Niezdolność do odpowiedniego ogrzania mieszkania w gospodarstwach domowych w Unii Europejskiej i krajach stowarzyszonych	48
Inability to adequately heat a dwelling in the European Union and associated countries' households	
Dyskusje. Recenzje. Informacje	
Discussions. Reviews. Information	
Justyna Gustyn	
Wydawnictwa GUS. Wrzesień 2020	66
Publications of Statistics Poland. September 2020	
Dla autorów	68
For the authors	
Zakres tematyczny działów	77
Thematic scope of sections	

OD REDAKCJI

W październikowym wydaniu „Wiadomości Statystycznych. The Polish Statistician” znajdują Państwo artykuły z zakresu praktycznych zastosowań statystyki oraz studiów interdyscyplinarnych.

Numer otwiera praca Piotra Matuszaka *Temporary part-time employment and future wages of male employees – a panel data analysis of the German labour market*. Autor bada związek między wykonywaniem przez mężczyzn pracy tymczasowej w niepełnym wymiarze a płacą, jaką otrzymują oni, będąc później zatrudnieni w pełnym wymiarze czasu pracy. Zależność ta jest analizowana w okresie od dwóch do pięciu oraz od sześciu do dziesięciu lat przed momentem obserwacji. Autor posługuje się danymi za lata 1984–2014 dotyczącymi niemieckiego rynku pracy, które zostały pozyskane z Panelu Społeczno-Ekonomicznego. Stosuje estymator z efektami stałymi, rozwiązujący problem nieobserwowalnej heterogeniczności poprzez transformację wewnątrzgrupową. Z badania wynika, że doświadczenie zatrudnienia w niepełnym wymiarze czasu pracy niekorzystnie wpływa na wysokość wynagrodzenia w przyszłości. Modele za lata 1994–2014 pozwalają sformułować wniosek, że w analizowanym okresie praca w niepełnym wymiarze w ciągu ostatnich dwóch do pięciu lat prowadziła do obniżenia płacy średnio o 4,4% w porównaniu z zatrudnieniem w pełnym wymiarze czasu pracy, ale po upływie pięciu lat związek ten był już nieistotny statystycznie.

Prof. dr hab. Jan Zawadzki w artykule *Prognozowanie brakujących danych w szeregach czasowych przy zastosowaniu modeli hybrydowych – podejście teoretyczne i empiryczne* rozpatruje użyteczność modeli hybrydowych w prognozowaniu brakujących danych w szeregach czasowych na przykładzie zapotrzebowania na energię elektryczną w okresach godzinnych w wybranej aglomeracji. Autor analizuje modele szeregu czasowego oraz modele regresyjne ze złożonymi wahaniami sezonowymi, wykorzystując dane za trzy kolejne lata pierwszej dekady XXI w. Wykazuje, że regresyjne modele hybrydowe charakteryzują się lepszymi własnościami prognostycznymi i niższymi ocenami błędów prognoz inter- i ekstrapolacyjnych niż modele szeregu czasowego, przy czym różnice dokładności są większe w przypadku prognoz ekstrapolacyjnych. We wszystkich przypadkach najlepszymi modelami hybrydowymi okazują się pojedyncze modele hierarchiczne opisujące wahania o cyklu 24-godzinnym.

Przedmiotem artykułu Kornelii Przedworskiej *Niezdolność do odpowiedniego ogrzania mieszkania w gospodarstwach domowych w Unii Europejskiej i krajach stowarzyszonych* jest zjawisko ubóstwa energetycznego gospodarstw domowych oceniane na podstawie niezdolności do odpowiedniego ogrzania mieszkania. Do określenia jego poziomu autorka wykorzystuje dane Eurostatu i Głównego Urzędu Statystycznego, które pochodzą z badań EU-SILC obejmujących lata 2009–2018. Z analizy wynika, że średni odsetek gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania w Europie maleje, ale w niektórych krajach, takich jak Macedonia, Litwa czy Bułgaria, liczba gospodarstw domowych doświadczających ubóstwa energetycznego jest nadal duża. W Polsce poziom tego zjawiska jest niższy niż średni poziom w UE. Autorka bada ponadto korelację odsetka gospodarstw domowych w Polsce, które nie są w stanie odpowiednio ogrzać mieszkania, z wysokością przeciętnego dochodu *per capita* w gospodarstwie domowym oraz z poziomem ubóstwa ekonomicznego.

Ostatnią pozycją w numerze jest omówienie nowości wydawniczych GUS przygotowane przez Justynę Gustyn.

Zapraszamy do lektury.

FROM THE EDITORIAL TEAM

The October issue of *Wiadomości Statystyczne. The Polish Statistician* presents articles devoted to the practical applications of statistics and interdisciplinary studies.

Piotr Matuszak's article *Temporary part-time employment and future wages of male employees – a panel data analysis of the German labour market* opens the issue. The author investigates the connection between temporary part-time employment of male employees and their wages in subsequent full-time employment. The above-mentioned relationship is analysed in the periods from two to five and from six to ten years prior to the start of the observation. The author uses data on the German labour market for the years 1984–2014 which were drawn from the Socio-Economic Panel. He applies the fixed effects estimator which solves the unobserved heterogeneity problem by a 'within' transformation. The study demonstrates that the experience of part-time employment negatively affects future wages. It can be concluded from the models for the years 1994–2014 that in the analysed period, part-time work experience from two to five years before starting a full-time job resulted in wages lower by 4.4% on average than wages of those who had had only full-time work experience. However, five or more years after starting full-time employment, this relationship becomes statistically insignificant.

Jan Zawadzki, PhD, DSc, ProfTit, in his article *Forecasting missing data in time series with the application of hybrid models – theoretical and empirical approach*, discusses the usefulness of hybrid models in forecasting missing data in time series based on the example of the demand for electricity in hourly periods in a selected agglomeration. The author analyses time series models and regressive models with complex seasonal fluctuations using data for three consecutive years of the first decade of the 2000s. Regressive hybrid models demonstrate better forecasting properties and generate fewer mistakes in interpolative and extrapolative forecasts than time series models, and precision differences are greater in the case of extrapolative forecasts. In all the cases, individual hierarchical models describing 24-hour-cycle fluctuations proved to be the most effective in this respect of all hybrid models.

The subject matter of the paper *Inability to adequately heat a dwelling in the European Union and associated countries' households* by Kornelia Przedworska is the phenomenon of energy poverty of households, which is diagnosed when a household is unable to sufficiently heat its dwelling. To determine the level of energy poverty, the author uses data from Eurostat and Statistics Poland, including data from EU-SILC research conducted in the years 2009–2018. The analysis demonstrates that the average percentage of households unable to sufficiently heat their dwellings is decreasing in Europe, but in some countries, such as Macedonia, Lithuania or Bulgaria, the number of households experiencing energy poverty is still high. In Poland, this phenomenon occurs on a smaller scale than in the EU on average. The author also investigates the correlation of the percentage of Polish households unable to sufficiently heat their homes with the households' average income per capita and the level of economic poverty.

The issue closes with a selection of Statistics Poland's new publications compiled by Justyna Gustyn.

We wish you pleasant reading.

Temporary part-time employment and future wages of male employees – a panel data analysis of the German labour market

Piotr Matuszak^a

Abstract. The aim of this study is to assess the relationship between temporary part-time employment and the wages male employees receive in subsequent full-time employment within the first five and the first ten years from the date of starting their full-time employment. The study uses data from the German labour market, obtained from the Socio-Economic Panel for the years 1984–2014. The fixed effects estimator, which solves the unobserved heterogeneity issue by removing time-invariant individual effects by a ‘within’ transformation, was applied in the empirical analysis. The results indicate that having experience as a part-time worker is associated with lower future wages – a one-year increase in the number of years in part-time work in the last two to five years leads to a reduction in future wages in a full-time job by 4.4% on average, compared to having solely a full-time job experience. However, this relationship becomes statistically insignificant after five years of being employed full-time. The results are robust to different specifications and it is indicated that an inverse relationship between working below regular hours and future wages in full-time employment is related to working part-time in low- and medium-skilled occupations. At the same time, working part-time is less detrimental to future wages than periods of unemployment.

Keywords: part-time employment, part-time wages effects, German labour market

JEL: C33, J24, J31

Praca tymczasowa w niepełnym wymiarze a przyszłe wynagrodzenie mężczyzn – analiza panelowa niemieckiego rynku pracy

Streszczenie. Celem badania przedstawionego w artykule jest ocena związku między pracą tymczasową w niepełnym wymiarze wykonywaną przez mężczyzn a płacą, jaką otrzymują oni w ciągu pierwszych pięciu i dziesięciu lat od momentu zatrudnienia w pełnym wymiarze czasu pracy. W analizie empirycznej posłużono się danymi dotyczącymi niemieckiego rynku pracy pochodzącymi z Panelu Społeczno-Ekonomicznego za lata 1984–2014. Zastosowano estymator z efektami stałymi, który rozwiązuje problem nieobserwowalnej heterogeniczności poprzez transformację wewnątrzgrupową. Wyniki wskazują, że doświadczenie zatrudnienia w niepełnym wymiarze czasu pracy wiąże się z niższym przyszłym wynagrodzeniem – dodatkowy rok pracy w niepełnym wymiarze w okresie ostatnich dwóch do pięciu lat prowadzi do obniżenia płacy średnio o 4,4% w porównaniu z doświadczeniem pracy w pełnym wymiarze. Jednak po pięciu latach zatrudnienia w pełnym wymiarze czasu pracy związek ten jest już nieistotny statystycznie. Wyniki są stabilne przy różnych specyfikacjach modeli. Wykazano, że odwrotna zależność między pracą w niepełnym wymiarze a przyszłą płacą ma związek z doświadczeniem

^a Uniwersytet Ekonomiczny w Poznaniu, Instytut Ekonomii. ORCID: <https://orcid.org/0000-0003-4572-3168>.

zatrudnienia w zawodach wymagających niskich i średnich kwalifikacji. Jednocześnie doświadczenie pracy w niepełnym wymiarze jest mniej niekorzystne dla przyszłego wynagrodzenia niż okresy braku zatrudnienia.

Słowa kluczowe: niepełny wymiar czasu pracy, płaca, niemiecki rynek pracy

1. Introduction

The phenomenon of part-time employment has been broadly discussed in literature as a result of an increasing number of employees working fewer than regular hours. Researchers in this field focus not only on the impact of part-time work on current wages and job positions, but also on the issue of future employment and salaries. Part-time jobs may be seen as non-optimal positions and they can affect employees in various ways, for example in the form of lower future wages once becoming full-time workers. The main reasons for such a deterioration in future earnings might lie in the negative signalling effects and reduced human capital accumulation.

While part-time employment may be perceived as a means to signal employees' skills when they are otherwise difficult to evaluate, and, consequently, become a stepping stone towards full-time employment (Bollé, 1997; Buddelmeyer et al., 2005; Farber, 1999; Kyyrä et al., 2017; Månsson & Ottosson, 2011), a part-time job experience is often interpreted by employers as a sign of low productivity and/or a lack of high labour force attachment and motivation (Biewen et al., 2018; Connolly & Gregory, 2009). This, in turn, leads to lower future wages and reduces the probability of being offered a full-time job (Connolly & Gregory, 2009; Tam, 1997).

Another explanation of the potentially lower wages of former part-time employees is connected to limited human capital accumulation. Employees working below the standard number of hours accumulate less human capital and experience than the full-time employed in the same period (Hirsch, 2005). Moreover, part-time employment is often related to occupational downgrading and working below one's qualifications, which leads to the further deterioration in human capital (Connolly & Gregory, 2008; Manning & Petrongolo, 2008). Additionally, as Nelen and de Grip (2009) demonstrate, part-time workers receive less support from companies in terms of investment in their human capital, compared to their full-time counterparts.

Empirical analyses suggest rather negative effects of part-time employment on future wages – former part-time workers are likely to receive lower salaries in both part-time and full-time employment compared to workers with only full-time work experience. Connolly and Gregory (2009) show that part-time work leads to low returns in future earnings for British women, substantially more so in low-skilled jobs. Similarly, Blundell et al. (2016) indicate that part-time jobs for women in the UK are of low value in terms of human capital accumulation; moreover, they are associat-

ed with low wage trajectories compared to full-time employment. The causal effect of part-time work on female workers in Germany was researched by Paul (2016) and the results suggest that part-time employment negatively affects future wages (when compared to full-time work), which refers in particular to women working less than 15 hours a week.

Despite the fact that there are several empirical studies discussing the effects of part-time employment on future wages, they focus solely on the careers of female employees. As indicated by Biewen et al. (2018, p. 5 and 6), no detailed analysis of this issue with regard to men has been conducted to date.

Therefore, the aim of this study is to assess the relationship between temporary part-time employment and the wages male employees receive in subsequent full-time employment within the first five and the first ten years from the date of starting their full-time employment. The aforementioned aim allows the formulation of two research hypotheses:

- H1: temporary part-time work is associated with lower future wages (when compared to full-time employment) after the first year of being employed full-time;¹
- H2: temporary part-time work is associated with lower future wages (when compared to full-time employment) for longer than five years.

2. Research method

Data from the German Socio-Economic Panel (SOEP), Wave 31, covering the years 1984–2014, were used in the analysis.² This dataset was chosen for two reasons. Firstly, it is a very comprehensive set covering a broad range of characteristics of respondents over an extensive period, which makes it possible to employ a within variation of a number of individuals in an econometric analysis. Secondly, Germany is a country where the rate of men working part-time has increased substantially over the last decades, i.e. from approximately 2% in the early 1990s to over 9% in recent years (OECD, 2019), which is one of the most dynamic changes noted in the EU.

The sample consisted of men aged 18–58 whose course of employment was observed for at least five successive years. Men who were self-employed were excluded from the study the year they entered this status. The definition of part-time employment used in this study is based on the classification of OECD (2016), which adopted the term ‘part-time employee’ to describe those who work less than 30 hours per

¹ The distinction of the first year after entering full-time employment is made to control, to some extent, for the effects of the possible change of employer and the loss of job-specific human capital (Fitzenberger et al., 2016), which tend to be larger directly after the change.

² For a more detailed description of the SOEP, see Wagner et al. (2007). This time span was selected as the full version of SOEP’s Wave 31 was available at the time when the research herein was initiated.

week. The minimum number of working hours required to be classified in the analysis was one. To calculate the hourly wage, the gross monthly real wage (in euros, as of 2005) was divided by the number of hours worked. Since the SOEP data contain information on wages in the current month and the question on the number of hours worked relates to the week prior to the SOEP interview, the application of the actual hours worked could have led to misleading results. Therefore, the information on contractual hours was used to measure the hours worked – this approach was employed by, e.g., Paul (2016). Although the question about contractual hours also relates to the week before the SOEP interview, it may be argued that this value tends to be less volatile than the actual hours worked.³

The data used in the analysis consist of 126,023 observations. 14,348 men were observed, on average, for over 13 years. The minimum number of observed periods for one person was five, and the maximum 31. Table 1 presents a summary of the statistics describing the sample applied in this study. The average wage per hour was EUR 17.27 (in euros, as of 2005) in the period of 1984–2014. The mean value of contractual hours per week was 38.61; 39.20 for full-time and 19.46 for part-time workers. The average age of the men whose careers were analysed was 38.97 and the age group of 50–58 was represented most extensively in the analysis. 15% of the observations referred to men who lived in East Germany.

Table 1. Summary statistics, men aged 18–58, years examined 1984–2014

Variable	Number of observations	Mean
<i>Wage per hour</i>	81,404	17.27
<i>Hours</i>	103,703	38.61
<i>if full-time</i>	100,691	39.20
<i>if part-time</i>	3,012	19.46
<i>Age</i>	126,023	38.97
18–23	126,023	0.11
24–29	126,023	0.14
30–34	126,023	0.13
35–39	126,023	0.13
40–44	126,023	0.13
45–49	126,023	0.13
50–58	126,023	0.23
<i>Place of residence: East Germany</i>	126,023	0.15

Note. *Wage* is presented in the 2005 constant euros. *Hours* refers to job hours per week. *Age* groups and *Place of residence: East Germany* are dummy variables – their average values represent the share of individuals classified in a particular group in the sample. Weighted statistics are presented. Individuals' cross-sectional weights for all SOEP samples (W11105 in the SOEP database) were employed in order to compensate for unequal probabilities of selection and sample attrition and to obtain populations-based statistics.

Source: author's work based on SOEP data for the years 1984–2014.

³ To verify the robustness of the results, an analysis employing actual worked hours was conducted. See Table 5, Model 7.

As previously mentioned, the rate of part-time employment among men has increased substantially in Germany over recent years. Based on the SOEP data, the percentage of men in part-time employment increased from nearly 2% in the years 1984–1989 to almost 8% in 2014. Since the growing role of part-time employment suggests that the profile of an employee working less than 30 hours per week might also have changed, periods of lower and higher rates of part-time employment need to be distinguished. Therefore, the following analysis presents also data for the periods 1984–2001 and 2002–2014 separately.⁴

Table 2 presents the employment status of the men from the sample used in this study. It should be pointed out that in the years 1984–2014, part-time employment (PT) constituted on average 2.47% of the total employment of men and it increased after the year 2001. More than 80% of the men from the sample were working full-time (FT) and 17.21% were non-employed (NE) in the period 1984–2014.

Table 2. Structure of employment status, men aged 18–58

Employment status	1984–2014	1984–2001	2002–2014
	in %		
FT	80.32	81.47	78.51
PT	2.47	2.22	2.87
NE	17.21	16.31	18.62

Source: author's work based on the SOEP data for the years 1984–2014.

Table 3. Structure of occupations by employment status, men aged 18–58

Skill level of occupation	FT	PT		
		1984–2001	2002–2014	
	in %			
High-skilled	36.71	53.14	62.98	41.64
Medium-skilled	57.50	28.74	24.83	33.77
Low-skilled	5.79	18.12	12.19	24.59

Source: author's work based on SOEP data for the years 1984–2014.

Skill levels of occupations by employment status are presented in Table 3.⁵ It should be noted that most part-time workers were employed in high-skilled occupa-

⁴ The year 2002 is treated as a threshold as it was the first year with a higher-than-average rate of part-time employment within the period of 1984–2014.

⁵ Skill levels of occupations are based on the first digit of the International Standard Classification of Occupations (ISCO): high-skilled jobs – managers, professionals, technicians and associate professionals, armed forces (1–3; 10); medium-skilled jobs – clerical support workers, service and sales workers, skilled agricultural and fishery workers, craft and related trades workers, plant and machine operators, assemblers (4–8); low-skilled jobs – elementary occupations (9). This classification is used in, e.g., OECD (2012).

tions, whereas among those employed full-time there were proportionally fewer high-skilled employees. Likewise, the proportion of part-time employees in low-skilled occupations was higher than a similar proportion of full-time employees. Among all types of occupations, part-time employees were employed in medium-skill-level occupations least often. However, the percentage of part-time employees in high-skilled occupations decreased after 2001 (from 62.98% to 41.64%), and at the same time increased in low-skilled occupations (from 12.19% to 24.59%).

The fact that there is a large number of employees working less than 30 hours a week in high-skilled jobs is not surprising when analysing a more detailed classification of occupations (based on the 2-Digit Industry Code in the SOEP). Most part-time employees, 25.46%, were employed in jobs from the Education and Sport category, where 42.25% of teachers had less than 30 contractual hours per week in the analysed period. However, the percentage of part-time teachers fell from over 50% in the years 1989–1993 to 33% in 2014. When teachers were excluded from the analysis, high-skilled occupations accounted for 39.28% of the whole part-time employment sector, and low-skilled occupations for 23.16%. Other job categories with the highest part-time employment rates were Restaurants (12.97%) and Health Services (7.13%).

Table 4. Average wages per hour by employment status, men aged 18–58

Employment status	1990–2014	1990–2001	2002–2014
	in euros of 2005		
FT	17.11	16.72	17.58
PT	22.81	26.64	18.54
PT, excluding teachers	19.06	21.78	16.66

Source: author's work based on SOEP data for the years 1990–2014.

Employees working part-time earned higher hourly wages on average in the years 1990–2014, both when including and when excluding teachers (see Table 4). However, after the year 2001, there was a wage gap between part-time employees and full-time employees when teachers were not considered.

The descriptive statistics presented in this part of the study demonstrate that with the increasing role of part-time employment its characteristics changed. Before the year 2002, working less than 30 hours a week was primarily associated with high-skilled occupations with a wage premium (compared to full-time employment), mostly teachers. After 2002, more part-time employees were working in medium- and low-skilled jobs, receiving lower hourly wages.

This study uses data which comprise time series for each cross-sectional object. In order to choose between panel data models and a simple ordinary least squares (OLS) regression in the empirical analysis, the Breusch-Pagan Lagrange multiplier

test (Breusch & Pagan, 1980) was conducted. Its results show that the null hypothesis on zero variance across entities can be rejected (at the 0.01 significance level), which, in turn, implies that significant unobserved differences across individuals occur in the dataset, and therefore the OLS regression would not be a suitable method in this study. Fixed effect and random effect models are considered the basic solutions to the unobserved heterogeneity issue in panel data (Wooldridge, 2002). The choice between fixed effect and random effect models is based on the possible rejection of the orthogonality assumption, which means that one needs to verify whether time-invariant unobserved heterogeneity is correlated with explanatory variables. The Hausman test (Hausman, 1978), which examines differences between the coefficients obtained when employing both methods, was conducted, and its results suggest that the null hypothesis on non-systematic differences in coefficients can be rejected (at the 0.01 significance level). Therefore, the fixed effects model, which makes it possible for time-invariant omitted variables to be arbitrarily correlated with explanatory variables, should be the one applied in this analysis.

The fixed effects estimator is a pooled ordinary least squares estimator from the regression on time-demeaned data (Wooldridge, 2002, pp. 265–279). This transformation is most often called the ‘within’ or ‘fixed effects’ transformation and it solves the unobserved heterogeneity issue by removing time-invariant individual effects by means of time-demeaning. In this analysis, the relationship between part-time work experience and future wages is identified through employees who change their employment status, and the potential time-invariant determinants of wage levels – including the unobservable ones, such as intelligence, innate skills and talent – are controlled for by the fixed effects model. Since this study focuses on the relationship between temporary part-time work and the future wages in full-time employment, observations are limited to the then current full-time employees, which allows a comparison of the wages of full-time employees with part-time work experience to the wages of full-time workers with full-time work experience and the wages of full-time employees who were formerly non-employed.

The estimated model within the fixed-effects framework takes the following form:

$$lwage_{it} = \beta_0 + \beta_1 PT_{it-1} + \beta_2 \sum_{s=2}^5 PT_{it-s} + \beta_3 \sum_{s=6}^{10} PT_{it-s} + \beta_4 NE_{it-1} + \\ + \beta_5 \sum_{s=2}^5 NE_{it-s} + \beta_6 \sum_{s=6}^{10} NE_{it-s} + \mathbf{c}'_{it} \boldsymbol{\alpha} + \varepsilon_{it},$$

where:

lwage – log of hourly wages of the current full-time employees,

- PT – part-time employment status,
 NE – non-employment status,
 c – vector of control variables consisting of age groups, dummy variables indicating if a person lived in East Germany or rural areas and time dummies to control for common shocks,
 β_2, β_3 – parameters of interest by the variables indicating part-time experience,
 α – vector of parameters by control variables.

Clustered standard errors were calculated.

3. Results

Tables 5 and 6 present the results of the fixed effects estimation.⁶

Table 5. Past employment states and future full-time wages, dependent variable: log of real hourly wages

Specification	Model (years)						
	1 (1994–2014)	2 (1994–2001)	3 (2002–2014)	4 (1990–2014)	5 (1994–2014)	6 (1994–2014)	7 (1994–2014)
PT	0.315*** (0.05)	.	.
PT_{t-1}	-0.104*** (0.03)	-0.071* (0.04)	-0.107** (0.05)	-0.106*** (0.03)	-0.172*** (0.03)	-0.108*** (0.04)	-0.068** (0.03)
$\sum_{s=2}^5 PT_{t-s}$	-0.044** (0.02)	-0.023 (0.02)	-0.031* (0.02)	-0.043*** (0.01)	-0.029* (0.01)	-0.046** (0.02)	-0.040** (0.02)
$\sum_{s=6}^{10} PT_{t-s}$	-0.018 (0.01)	0.001 (0.02)	-0.012 (0.01)	.	-0.014* (0.01)	-0.019 (0.01)	-0.001 (0.01)
NE_{t-1}	-0.115*** (0.02)	-0.103*** (0.02)	-0.114*** (0.02)	-0.109*** (0.01)	-0.121*** (0.02)	-0.115*** (0.02)	-0.113*** (0.02)
$\sum_{s=2}^5 NE_{t-s}$	-0.053*** (0.01)	-0.078*** (0.01)	-0.034*** (0.01)	-0.082*** (0.01)	-0.054*** (0.01)	-0.053*** (0.01)	-0.051*** (0.01)
$\sum_{s=6}^{10} NE_{t-s}$	-0.025*** (0.00)	-0.028*** (0.01)	-0.016*** (0.01)	.	-0.022*** (0.00)	-0.025*** (0.00)	-0.024*** (0.00)
<i>Age groups, East Germany, rural area, year dummies, intercept</i>	yes	yes	yes	yes	yes	yes	yes
<i>FT</i>	ref.	.	.
<i>Teacher dummies</i>	no	no	no	no	no	yes	no
N	20,332	6,936	13,396	43,543	20,701	20,332	20,341

Note. Clustered standard errors are reported in parentheses. Asterisks denote significance levels: *** – 0.01, ** – 0.05, * – 0.1; ref. – reference group. Observations consist of current full-time employees. Model 5 contains also current part-time workers. In Model 7, actual worked hours were used. Detailed results are presented in the Appendix, Tables A1–A4.

Source: author's work supported by Stata 15 (xtreg package) and data from SOEP.

⁶ For detailed results, see the Appendix, Tables A1–A5.

Table 6. Past employment states and future wages – skill levels of occupations, dependent variable: log of real hourly wages

Specification	Model 8 (years 1994–2014)	Specification	Model 8 (years 1994–2014)	Specification	Model 8 (years 1994–2014)
<i>FThigh</i>	ref.	$\sum_{s=2}^5 PThigh_{t-s}$	-0.012 (0.02)	$\sum_{s=2}^5 NE_{t-s}$	-0.095*** (0.01)
<i>PThigh</i>	0.309*** (0.05)	$\sum_{s=2}^5 PTmedium_{t-s}$	-0.042** (0.02)	$\sum_{s=6}^{10} NE_{t-s}$	-0.023*** (0.01)
<i>PTmedium</i>	0.318*** (0.10)	$\sum_{s=2}^5 PTlow_{t-s}$	-0.090*** (0.03)	Age groups, East Germany, rural area, year dummies, intercept	yes
<i>PTlow</i>	0.319*** (0.09)	$\sum_{s=6}^{10} PThigh_{t-s}$	0.002 (0.01)	FT occupations dummies	yes
<i>PThigh_{t-1}</i>	-0.164*** (0.03)	$\sum_{s=6}^{10} PTmedium_{t-s}$	-0.041*** (0.01)	Teacher dummies	yes
<i>PTmedium_{t-1}</i>	-0.199*** (0.04)	$\sum_{s=6}^{10} PTlow_{t-s}$	-0.020 (0.03)	N	20,701
<i>PTlow_{t-1}</i>	-0.142* (0.08)	<i>NE_{t-1}</i>	-0.127*** (0.01)		

Note. Clustered standard errors are reported in parentheses. Asterisks denote significance levels: *** – 0.01, ** – 0.05, * – 0.1; ref. – reference group. Observations consist of the current full-time and part-time employees. Detailed results are presented in the Appendix, Table A5.

Source: author’s work supported by Stata 15 (xtreg package) and data from SOEP.

The first column includes the estimates from the analysis of the data from the years 1994–2014⁷ (Model 1). It shows that an increase of one year in part-time work experience in the last five years (excluding the last year) decreases the current hourly wage by 4.4% on average (compared to full-time work experience) and this relationship is statistically significant at the 0.05 level. At the same time, no statistically significant (of 0.1 level) relation was observed between working part-time six to ten years earlier and the wages earned in the future full-time job. In comparison, the coefficients of both variables indicating non-employment in the past years suggest a stronger negative correlation (5.3% for the last five years; 2.5% six to ten years earlier) and are significant at the 0.01 level.

In Models 2 and 3, observations were divided into periods 1994–2001 and 2002–2014. An inverse relationship between working part-time and future wages proved statistically insignificant after the first year from switching to full-time employment in the years 1994–2001. In the second analysed period, the coefficient of the variable indicating previous experience of part-time work within five years equalled –3.1% and was significantly different from zero at the 0.1 level. As previously, the experi-

⁷ This time span is necessitated by the fact that the model includes variables indicating the job status in the last ten years, and, while having access to SOEP data for the period from 1984 to 2014, variables on the past ten years are available starting only from 1994.

ence of part-time employment was not statistically significant (at the 0.1 level) after five years of working full-time. In both models, an increase in the number of years of non-employment in the past was associated with lower future wages and this relationship was of high statistical significance.

These results suggest that the H1 hypothesis cannot be rejected while there is sufficient evidence to reject the H2 hypothesis.

4. Robustness check

In the first step of the robustness check,⁸ the variable indicating work experience of six to ten years was excluded, which allowed to roughly double the number of observations in regression and increased the time span to the years 1990–2014⁹ (Table 5, Model 4). Then, both the current full-time and part-time workers were considered. Consequently, the relationship between part-time work experience and wages of employees working currently less than 30 hours a week (Model 5) was possible to examine. Since the characteristics of teachers in part-time employment differ from those of other workers, interaction terms capturing part-time experience of this professional group were included (Model 6). In the next step, the actual worked hours instead of contractual working hours were used (Model 7). Finally, skill levels of occupations were taken into consideration to see whether the obtained results held for the different job groups (Table 6, Model 8).

The results of the analysis including only variables indicating work experience in the last five years (Model 4) demonstrate that each additional year in part-time work in the previous two to five years was associated with, on average, a 4.3%-decrease in wages, compared to the experience in full-time employment, and this relationship is highly significant. The magnitude is very close to the value obtained in Model 1 (–4.4%).

Both those working less than 30 hours a week and those working at least 30 hours a week in the years 1994–2014 were included in Model 5. Coefficients of the variables indicating part-time experience equalled –2.9% (two to five years) and –1.4% (six to ten years), and are significant at the 0.1 level.

When interaction terms to control for teachers were introduced (Model 6), the coefficient of the variable indicating experience in part-time employment in the previous two to five years was of a slightly greater magnitude (–4.6%) than in the basic specification, and was significant at the 0.05 level. The coefficient of the variable representing part-time experience of over five years was insignificantly different from zero.

⁸ Detailed results of the robustness check are available in the Appendix, Tables A1–A5.

⁹ Technically, the time span could also cover the year 1989, however, the analysis starts with the year 1990 as it is the first period taking into account former East Germany.

When the actual worked hours were considered (Model 7), a link between part-time experience up to five years (excluding the previous year) and future wages was identified as negative and significant at the 0.05 level. Its magnitude was slightly lower than in the basic specification (−4.0%). The relationship between part-time work experience from before six to ten years and the future full-time wages was statistically insignificant.

In Table 6 (Model 8) skill levels of occupation were examined, both in part-time and full-time employment. The results demonstrate that there were no statistically significant differences between part-time and full-time work experience in high-skilled occupations. At the same time, the coefficients of the variable representing a past medium-skilled part-time worker were negative (−4.2% for the experience from two to five years before the start of full-time employment, and −4.1% for the experience from six to ten years before the start of full-time employment) and significant at the 0.05 and 0.01 levels, respectively. A coefficient for low-skilled part-time work experience has a negative sign and is of a greater magnitude (−9.0%) than the other coefficients, and is highly significant. However, this correlation after five years became statistically insignificant. When comparing the occupational skill groups in full-time and part-time employment, it should be pointed out that the inverse relationship between working in medium- or low-skilled occupations and the future wages was stronger for part-time employees. Therefore, reduced work experience both in terms of working fewer hours and working below one's qualifications might negatively affect future wages.

The analysis presented above demonstrates that the inverse relationship between part-time work experience and the future full-time wages, hinted at in the previous part of this paper, is robust to different specifications of the model, which supports the H1 hypothesis. The results of Model 8 suggest that that link applies to part-time work in medium- and low-skilled occupations. The results for part-time work experience after five years from entering full-time employment differ between specifications, which leads to the conclusion that there is insufficient evidence to support the H2 hypothesis.

5. Discussion

The relationship between part-time work experience and lower future wages in full-time employment (up to five years after switching to full-time employment) presented in this study is in line with previous empirical research on the effects of part-time employment on future wages (Blundell et al., 2016; Connolly & Gregory, 2009;

Paul, 2016; one has to remember, though, that these studies concern solely female employees).

This paper also shows, through analysing part-time work experience across different occupational groups, that both working fewer hours and working below one's qualifications can be a reason for the negative impact of part-time work on future wages. Similar conclusions were presented by Connolly and Gregory (2009) with regard to women in the UK.

While part-time work experience leads to lower future earnings as compared to salaries earned by employees having solely full-time work experience, the decrease in the hourly wage resulting from former part-time employment is smaller than the decrease resulting from former non-employment. Paul (2016) finds a similar relationship for German women. Blundell et al. (2016) observed a related pattern for women from the UK, namely that part-time employment reduces the depreciation of skills resulting from non-employment.

Concerning the possible limitations of this study, it must be mentioned that the fixed effects 'within' estimator assumes time-invariant unobserved heterogeneity, and therefore some unobserved characteristics of individuals which change over time (e.g. mid-career educational enhancements) may not be taken into consideration. Another issue that might limit the possibility of drawing unambiguous conclusions as to the causal effects is the problem of selection into part-time employment, which might be affected by a number of factors, including changes in legislation. Finally, reasons for switching to part-time work differ across employees, which might also have heterogeneous effects on future wages. Therefore, the results presented in this study should be approached with caution in terms of causality.

6. Conclusions

The aim of this study was to address a substantial research gap in the assessment of the relationship between temporary part-time employment and the subsequent wages in full-time employment of male workers (within five and ten years from the date of switching to full-time employment), and so to contribute to literature devoted to this matter. To account for unobserved time-invariant characteristics of employees, the fixed effects 'within' estimator was employed. Data from the German SOEP covering the years 1984–2014 were used in the analysis.

The results indicate that part-time work experience is associated with lower future wages (when compared to full-time work experience) within five years after switching to full-time employment. A one-year increase in the number of years in part-time work in the last two to five years decreases full-time job wages on average by

4.4%. Therefore, the first research hypothesis, that previous temporary part-time work experience is associated with lower wages after the first year of being employed full-time, when compared to full-time work experience, cannot be rejected. However, the link between the former experience of temporary part-time work and future wages was statistically insignificant after five years of having worked full-time, which suggests that the second research hypothesis, which says that this link would be statistically significant, should be rejected. Further analyses show that the results were driven mainly by the observations after the year 2001 and that an inverse relationship between previous part-time employment and future full-time job wages was observable among former workers of part-time low- and medium-skilled jobs. The results were robust when testing both full-time and part-time workers, controlling for teachers and examining the actual worked hours.

The increasing scale of part-time employment among men in the EU leads to the question of the long-term economic effects of working less than regular hours. An empirical analysis in this study shows that former part-time employees earn considerably less after switching to full-time employment. However, this inverse relationship is no longer significant after five years from having started working full-time. Non-employment seems to be substantially more detrimental than working fewer than regular hours in terms of the subsequent full-time wages. These results indicate the need for further research in the field of part-time employment among men, especially in other EU countries, with the application of methods that would make it possible to draw unambiguous causal inferences.

Acknowledgements

I would like to especially thank Bernd Fitzenberger, Jakob de Lazzar and the participants of 'Econometric Projects' seminar at the Humboldt University of Berlin in the Winter Semester of 2016/2017 for their insightful comments and suggestions. Should any errors be detected in this paper, I am the only person responsible for them.

References

- Biewen, M., Fitzenberger, B., de Lazzar, J. (2018). The role of employment interruptions and part-time work for the rise in wage inequality. *IZA Journal of Labor Economics*, 7(1), 1–34. <https://doi.org/10.1186/s40172-018-0070-y>
- Blundell, R., Dias, M. C., Meghir, C., Shaw, J. (2016). Female Labor Supply, Human Capital, and Welfare Reform. *Econometrica*, 84(5), 1705–1753. <https://doi.org/10.3982/ECTA11576>
- Bollé, P. (1997). Part-time work: Solution or trap?. *International Labour Review*, 136(4), 557–579. <https://www.ilo.org/public/english/revue/articles/97-4.htm>

- Breusch, T. S., Pagan, A. R. (1980). The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. *The Review of Economic Studies*, 47(1), 239–253. <https://www.jstor.org/stable/2297111>
- Buddelmeyer, H., Mourre, G., Ward-Warmedinger, M. (2005). *Part-Time Work in EU Countries – Labour Market Mobility, Entry and Exit* (ECB Working Paper No. 460). <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp460.pdf>
- Connolly, S., Gregory, M. (2008). Moving down: women's part-time work and occupational change in Britain 1991–2001. *The Economic Journal*, 118(526), 52–76. <https://doi.org/10.1111/j.1468-0297.2007.02116.x>
- Connolly, S., Gregory, M. (2009). The Part-time Pay penalty: Earnings Trajectories of British Women. *Oxford Economic Papers*, 61, 76–97. <https://doi.org/10.1093/oep/gpn043>
- Farber, H. S. (1999). Alternative and Part-Time Employment Arrangements as a Response to Job Loss. *Journal of Labour Economics*, 17(S4), 142–169. <https://doi.org/10.1086/209946>
- Fitzenberger, B., Steffes, S., Strittmatter, A. (2016). Return-to-job during and after parental leave. *The International Journal of Human Resource Management*, 27(8), 803–831. <https://doi.org/10.1080/09585192.2015.1037328>
- Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6), 1251–1271. <https://www.jstor.org/stable/1913827>
- Hirsch, B. T. (2005). Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills. *ILR Review*, 58(4), 525–551. <https://doi.org/10.1177/001979390505800401>
- Kyyrä, T., Arranz, J. M., García-Serrano, C. (2017). *Does Part-Time Work Help Unemployed Workers to Find Full-Time Work?. Evidence from Spain* (IZA Working Papers No. 10770). <http://ftp.iza.org/dp10770.pdf>
- Manning, A., Petrongolo, B. (2008). The part-time penalty for women in Britain. *The Economic Journal*, 118(526), 28–51. <https://doi.org/10.1111/j.1468-0297.2007.02115.x>
- Månsson, J., Ottosson, J. (2011). Transitions from part-time unemployment: Is part-time work a dead end or a stepping stone to the labour market?. *Economic and Industrial Democracy*, 32(4), 569–589. <https://doi.org/10.1177/0143831X10387836>
- Nelen, A., de Grip, A. (2009). Why Do Part-time Workers Invest Less in Human Capital than Full-timers?. *Labour*, 23(s1), 61–83. <https://doi.org/10.1111/j.1467-9914.2008.00439.x>
- OECD. (2012). *Settling In: OECD Indicators of Immigrant Integration 2012*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264171534-en>
- OECD. (2016). *OECD Labour Force Statistics 2015*. Paris: OECD Publishing. https://doi.org/10.1787/oecd_lfs-2015-en
- OECD. (2019). *OECD Employment Outlook 2019: The Future of Work*. Paris: OECD Publishing. <https://doi.org/10.1787/9ee00155-en>
- Paul, M. (2016). Is There a Causal Effect of Working Part-Time on Current and Future Wages?. *The Scandinavian Journal of Economics*, 118(3), 494–523. <https://doi.org/10.1111/sjoe.12157>
- Tam, M. (1997). *Part-Time Employment: A Bridge or a Trap?*. Aldershot: Avebury.
- Wagner, G. G., Frick, J. R., Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. *Schmollers Jahrbuch: Journal of Applied Social Science Stu-*

dies/Zeitschrift für Wirtschafts- und Sozialwissenschaften, 127(1), 139–169. <http://dx.doi.org/10.2139/ssrn.1028709>

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge,

Appendix

Table A1. Past employment states and future wages, dependent variable: log of real hourly wages, detailed results

Specification	Model (years)			
	1 (1994–2014)	2 (1994–2001)	3 (2002–2014)	4 (1990–2014)
PT_{t-1}	-0.104*** (0.03)	-0.071* (0.04)	-0.107** (0.05)	-0.106*** (0.03)
$\sum_{s=2}^5 PT_{t-s}$	-0.044** (0.02)	-0.023 (0.02)	-0.031* (0.02)	-0.043*** (0.01)
$\sum_{s=6}^{10} PT_{t-s}$	-0.018 (0.01)	0.001 (0.02)	-0.012 (0.01)	.
NE_{t-1}	-0.115*** (0.02)	-0.103*** (0.02)	-0.114*** (0.02)	-0.109*** (0.01)
$\sum_{s=2}^5 NE_{t-s}$	-0.053*** (0.01)	-0.078*** (0.01)	-0.034*** (0.01)	-0.082*** (0.01)
$\sum_{s=6}^{10} NE_{t-s}$	-0.025*** (0.00)	-0.028*** (0.01)	-0.016*** (0.01)	.
Age: 18–23	-0.292*** (0.02)
24–29	-0.074*** (0.01)	-0.048*** (0.01)	-0.065*** (0.02)	-0.080*** (0.01)
35–39	0.042*** (0.01)	0.020** (0.01)	0.046*** (0.01)	0.042*** (0.01)
40–44	0.067*** (0.01)	0.035** (0.01)	0.071*** (0.02)	0.054*** (0.01)
45–49	0.070*** (0.01)	0.040** (0.02)	0.077*** (0.02)	0.042*** (0.01)
50–58	0.054*** (0.01)	0.035* (0.02)	0.062*** (0.02)	0.011 (0.01)
30–34	ref.	ref.	ref.	ref.
East Germany	0.075 (0.07)	0.146** (0.06)	-0.026 (0.04)	0.006 (0.05)
Rural area	0.014 (0.03)	0.002 (0.02)	0.047* (0.03)	0.013 (0.02)
Year dummies	yes	yes	yes	yes
Intercept	2.879*** (0.02)	2.935*** (0.02)	2.842*** (0.02)	2.921*** (0.02)
N	20,332	6,936	13,396	43,543
R-squared within	0.115	0.112	0.050	0.217
R-squared between	0.023	0.060	0.115	0.211
Rho	0.905	0.880	0.918	0.872
σ_u	0.396	0.298	0.417	0.405
σ_e	0.128	0.110	0.125	0.155

Note. Clustered standard errors are reported in parentheses. Asterisks denote significance levels: *** – 0.01, ** – 0.05, * – 0.1. ref. – reference group. *PT* – part-time employment, *NE* – non-employment. Observations consist of current full-time employees.

Source: author's work supported by Stata 15 (xtreg package) and data from SOEP.

Table A2. Past employment states and future wages, dependent variable:
log of real hourly wages, part-time employees included, detailed results

Specification	Model 5 (years 1994–2014)	Specification	Model 5 (years 1994–2014)	Specification	Model 5 (years 1994–2014)
<i>PT</i>	0.315*** (0.05)	<i>Age: 18–23</i>	<i>Rural area</i>	0.025 (0.03)
<i>PT</i> _{<i>t</i>-1}	-0.172*** (0.03)	24–29	-0.074*** (0.01)	<i>Year dummies</i>	yes
$\sum_{s=2}^5 PT_{t-s}$	-0.029* (0.01)	35–39	0.042*** (0.01)	<i>Intercept</i>	2.877*** (0.02)
$\sum_{s=6}^{10} PT_{t-s}$	-0.014* (0.01)	40–44	0.069*** (0.01)	<i>N</i>	20,701
<i>NE</i> _{<i>t</i>-1}	-0.121*** (0.02)	45–49	0.071*** (0.01)	<i>R-squared within</i>	0.132
$\sum_{s=2}^5 NE_{t-s}$	-0.054*** (0.01)	50–58	0.056*** (0.01)	<i>R-squared between</i>	0.021
$\sum_{s=6}^{10} NE_{t-s}$	-0.022*** (0.00)	30–34	ref.	<i>Rho</i>	0.902
		<i>East Germany</i>	0.071 (0.07)	σ_u	0.404
				σ_e	0.133

Note. Clustered standard errors are reported in parentheses. Asterisks denote significance levels: *** – 0.01, ** – 0.05, * – 0.1. *PT* – part-time employment, *NE* – non-employment. Observations consist of current full-time and part-time employees.

Source: author's work supported by Stata 15 (xtreg package) and data from SOEP.

Table A3. Past employment states and future wages, dependent variable:
log of real hourly wages, teacher dummies, detailed results

Specification	Model 6 (years 1994–2014)	Specification	Model 6 (years 1994–2014)	Specification	Model 6 (years 1994–2014)
<i>PT</i> _{<i>t</i>-1}	-0.108*** (0.04)	<i>Age: 35–39</i>	0.042*** (0.01)	<i>Teacher PT</i> _{<i>t</i>-1}	0.031 (0.05)
$\sum_{s=2}^5 PT_{t-s}$	-0.046** (0.02)	40–44	0.067*** (0.01)	<i>Teacher</i> $\sum_{s=6}^{10} PT_{t-s}$	0.016 (0.02)
$\sum_{s=6}^{10} PT_{t-s}$	-0.019 (0.01)	45–49	0.069*** (0.01)	<i>Year dummies</i>	0.005
<i>NE</i> _{<i>t</i>-1}	-0.115*** (0.02)	50–58	0.054*** (0.01)	<i>Intercept</i>	yes 2.879*** (0.02)
$\sum_{s=2}^5 NE_{t-s}$	-0.053*** (0.01)	30–34	ref.	<i>N</i>	20,332
$\sum_{s=6}^{10} NE_{t-s}$	-0.025*** (0.00)	<i>East Germany</i>	0.076 (0.07)	<i>R-squared within</i>	0.115
<i>Age: 18–23</i>	–	<i>Rural area</i>	0.015 (0.03)	<i>R-squared between</i>	0.022
24–29	-0.074*** (0.01)	<i>Teacher</i>	-0.031 (0.04)	<i>Rho</i>	0.905
				σ_u	0.396
				σ_e	0.128

Note. As in Table A1.

Source: author's work supported by Stata 15 (xtreg package) and data from SOEP.

Table A4. Past employment states and future wages, dependent variable: log of real hourly wages, actual worked hours, detailed results

Specification	Model 7 (years 1994–2014)	Specification	Model 7 (years 1994–2014)	Specification	Model 7 (years 1994–2014)
PT_{t-1}	-0.068** (0.03)	<i>Age: 18–23</i>	x	<i>East Germany</i>	0.055 (0.06)
$\sum_{s=2}^5 PT_{t-s}$	-0.040** (0.02)	24–29	-0.066*** (0.01)	<i>Rural area</i>	0.031 (0.03)
$\sum_{s=6}^{10} PT_{t-s}$	-0.001 (0.01)	35–39	0.038*** (0.01)	<i>Year dummies</i>	yes
NE_{t-1}	-0.113*** (0.02)	40–44	0.054*** (0.01)	<i>Intercept</i>	2.815*** (0.02)
$\sum_{s=2}^5 NE_{t-s}$	-0.051*** (0.01)	45–49	0.054*** (0.01)	N	20,341
$\sum_{s=6}^{10} NE_{t-s}$	-0.024*** (0.00)	50–58	0.040*** (0.01)	<i>R-squared within</i> ...	0.092
		30–34	ref.	<i>R-squared between</i>	0.023
				Rho	0.886
				σ_u	0.381
				σ_e	0.137

Note. As in Table A1.

Source: author’s work supported by Stata 15 (xtreg package) and data from SOEP.

Table A5. Past employment states and future wages, dependent variable: log of real hourly wages, occupational levels, detailed results

Specification	Model 8 (years 1994–2014)	Specification	Model 8 (years 1994–2014)	Specification	Model 8 (years 1994–2014)
<i>FThigh</i>	ref.	$\sum_{s=6}^{10} PThigh_{t-s}$	0.002 (0.01)	<i>Age: 45–49</i>	0.069*** (0.01)
<i>FTmedium</i>	-0.016** (0.01)	$PTmedium_{t-1}$	-0.199*** (0.04)	50–58	0.054*** (0.01)
<i>FTlow</i>	-0.031*** (0.01)	$\sum_{s=2}^5 PTmedium_{t-s}$	-0.042** (0.02)	30–34	ref.
$FTmedium_{t-1}$	-0.017** (0.01)	$\sum_{s=6}^{10} PTmedium_{t-s}$	-0.041*** (0.01)	<i>East Germany</i>	0.076 (0.07)
$\sum_{s=2}^5 FTmedium_{t-s}$	-0.014*** (0.00)	$PTlow_{t-1}$	-0.142* (0.08)	<i>Rural area</i>	0.015 (0.03)
$\sum_{s=6}^{10} FTmedium_{t-s}$	-0.010*** (0.00)	$\sum_{s=2}^5 PTlow_{t-s}$	-0.090*** (0.03)	<i>Teacher</i>	-0.031 (0.04)
$FTlow_{t-1}$	-0.019* (0.01)	$\sum_{s=6}^{10} PTlow_{t-s}$	-0.020 (0.03)	<i>Teacher</i> PT_{t-1}	0.031 (0.05)
$\sum_{s=2}^5 FTlow_{t-s}$	-0.026*** (0.01)	NE_{t-1}	-0.135*** (0.02)	<i>Teacher</i> $\sum_{s=2}^5 PT_{t-s}$	0.016 (0.02)
$\sum_{s=6}^{10} FTlow_{t-s}$	-0.016*** (0.00)	$\sum_{s=2}^5 NE_{t-s}$	-0.064*** (0.01)	<i>Teacher</i> $\sum_{s=6}^{10} PT_{t-s}$	0.005 (0.02)
<i>PThigh</i>	0.309*** (0.05)	$\sum_{s=6}^{10} NE_{t-s}$	-0.028*** (0.00)	<i>Year dummies</i>	yes
<i>PTmedium</i>	0.318*** (0.10)	<i>Age: 18–23</i>	x	<i>Intercept</i>	2.966***
<i>PTlow</i>	0.319*** (0.09)	24–29	-0.074*** (0.01)	N	20,701
$PThigh_{t-1}$	-0.164*** (0.03)	35–39	0.042*** (0.01)	<i>R-squared within</i> ...	0.145
$\sum_{s=2}^5 PThigh_{t-s}$	-0.012 (0.02)	40–44	0.067*** (0.01)	<i>R-squared between</i>	0.168
				Rho	0.890
				σ_u	0.376
				σ_e	0.132

Note. As in Table A2. *FT* – full-time employment.

Source: author’s work supported by Stata 15 (xtreg package) and data from SOEP.

Prognozowanie brakujących danych w szeregach czasowych przy zastosowaniu modeli hybrydowych – podejście teoretyczne i empiryczne

Jan Zawadzki^a

Streszczenie. Celem artykułu jest przedstawienie możliwości prognozowania brakujących danych w szeregach czasowych dla danych godzinnych przy zastosowaniu modeli hybrydowych. W badaniu wykorzystano hybrydowe modele szeregu czasowego oraz modele regresyjne ze złożonymi wahaniami sezonowymi. Wahania złożone dla danych godzinnych mogą być sumą lub iloczynem wahań o cyklach rocznym, tygodniowym i dobowym, a wahania o długości cyklu wyrażonej liczbą parzystą (12-miesięczne i 24-godzinne) mogą być opisywane za pomocą regularnych modeli hierarchicznych.

Rozważania o charakterze teoretycznym podparto analizą empiryczną zapotrzebowania na energię elektryczną w okresach godzinnych w wybranej aglomeracji. Wykorzystano dane statystyczne obejmujące trzy kolejne lata pierwszej dekady XXI w., które zostały udostępnione przez przedsiębiorstwo zajmujące się dystrybucją energii elektrycznej i włączone do Banku Danych Katedry Zastosowań Matematyki w Ekonomii Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie. Przyjęto założenie, że luki niesystematyczne będą występowały w przypadku każdego rodzaju wahań sezonowych. Otrzymane wyniki wskazują na użyteczność modeli hybrydowych w prognozowaniu zjawisk ekonomicznych o bardzo dużej częstotliwości prowadzenia obserwacji.

Słowa kluczowe: dane godzinne, sezonowość złożona, modele hybrydowe, luki niesystematyczne

JEL: C32, C53

Forecasting missing data in time series with the application of hybrid models – theoretical and empirical approach

Abstract. The main goal of the article is to present the possibilities of forecasting missing observations in time series for hourly data with the application of hybrid models. Hybrid time series models and regression models with complex seasonal fluctuations were used in the study. Complex fluctuations for hourly data can be either a sum or a product of fluctuations of annual, weekly and daily cycles, while fluctuations the length of a cycle expressed by an even number (12-month and 24-hour ones) can be described using regular hierarchical models.

The theoretical considerations were illustrated by an empirical analysis of the demand for electricity in hourly periods in a selected agglomeration. The statistical data covered three consecutive years of the first decade of the 2000s. The data were provided by an electricity distribution company and included in the Data Bank of the Department of Applied Mathematics in Economics of the West Pomeranian University of Technology in Szczecin. It was assumed

^a Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, Wydział Ekonomiczny.
ORCID: <https://orcid.org/0000-0002-2996-3180>.

that non-systematic gaps occur with regard to all types of seasonal fluctuations. The obtained results indicate the usefulness of hybrid models in forecasting economic phenomena subject to very frequent observations.

Keywords: hourly data, complex seasonality, hybrid models, non-systematic gaps

1. Wprowadzenie teoretyczne

W modelowaniu i prognozowaniu zmiennych dla danych godzinnych znajdują zastosowanie klasyczne modele szeregu czasowego z potrójnie złożonymi wahaniami sezonowymi w postaci addytywnej lub multiplikatywnej, w których każdy rodzaj wahań jest opisywany za pomocą osobnych podzbiorów zmiennych zero-jedynkowych.

Zapis analityczny modelu szeregu czasowego wykazującego złożoną sezonowość w postaci addytywnej z trendem liniowym jest następujący (Kufel, 2010; Szmuksta-Zawadzka i Zawadzki, 2011):

$$Y_t = \alpha_1 t + \alpha_0 + \sum_{i=1}^{12} b_{0i} M_{it} + \sum_{j=1}^7 c_{0j} D_{jt} + \sum_{k=1}^{24} d_{0k} G_{kt} + U_t \quad (1)$$

$$t = 1, 2, \dots, n$$

przy warunkach

$$\sum_{i=1}^{12} b_{0i} = \sum_{j=1}^7 c_{0j} = \sum_{k=1}^{24} d_{0k} = 0,$$

gdzie:

M_{it} – zmienne zero-jedynkowe przyjmujące wartość 1 w i -tym miesiącu ($i = 1, 2, \dots, 12$),

D_{jt} – zmienne zero-jedynkowe przyjmujące wartość 1 w j -tym dniu tygodnia ($j = 1, 2, \dots, 7$),

G_{kt} – zmienne zero-jedynkowe przyjmujące wartość 1 w k -tej godzinie dnia tygodnia ($k = 1, 2, \dots, 24$),

U_t – składnik losowy.

Model w postaci wykładniczej wyraża się wzorem

$$Y_t = \exp \left(\alpha_1 t + \alpha_0 + \sum_{i=1}^{12} b_{0i} M_{it} + \sum_{j=1}^7 c_{0j} D_{jt} + \sum_{k=1}^{24} d_{0k} G_{kt} + U_t \right) \quad (2)$$

$$t = 1, 2, \dots, n$$

przy warunkach

$$\sum_{i=1}^{12} b_{0i} = \sum_{j=1}^7 c_{0j} = \sum_{k=1}^{24} d_{0k} = 0.$$

Objaśnienia symboli są takie same jak w modelu (1).

Po obustronnym zlogarytmowaniu za pomocą logarytmu naturalnego otrzymujemy jego postać zlinearyzowaną względem $\ln Y_t$.

W literaturze coraz częściej można spotkać przykłady wykorzystania modeli hierarchicznych z wahaniami sezonowymi w modelowaniu i prognozowaniu na podstawie kompletnych szeregów czasowych. Zazwyczaj dotyczy to szeregów dla danych miesięcznych i ewentualnie dziennych (m.in. Athanasopoulos i in., 2017; Hyndman i in., 2011), rzadziej dla danych godzinnych czy półgodzinnych (np. Kwac i in., 2014; Taieb i in., 2017).

Modele, w których przynajmniej jeden rodzaj wahań o długości cyklu wyrażonej liczbą parzystą jest opisywany za pomocą regularnych modeli hierarchicznych, noszą nazwę modeli hybrydowych (mieszanych). W modelach tych wahania o długości cyklu wyrażonej liczbą nieparzystą (7-dniowe) są opisywane tak jak w modelu (1), za pomocą zmiennych zero-jedynkowych. Metody i przykłady ich zastosowań można znaleźć m.in. w pracach Szmuksta-Zawadzkiej i Zawadzkiego (2014a) oraz Zawadzkiego (2018). Regularne modele hierarchiczne mogą być wykorzystywane do opisu wahań o parzystej długości cykli dla danych godzinnych (w analizowanym przypadku – o cyklach rocznym i dobowym).

Modele, w których przynajmniej jeden rodzaj wahań jest opisywany za pomocą modelu hierarchicznego, to pojedyncze modele hybrydowe. Jeżeli wykorzystuje się je do opisu obu rodzajów wahań, nazywa się je podwójnymi modelami hybrydowymi. Wahania o cyklu tygodniowym (nieparzystym) każdorazowo są opisywane za pomocą zmiennych zero-jedynkowych D_{jt} .

Definicja regularnego modelu hierarchicznego zamieszczona została w pracy Szmuksta-Zawadzkiej i Zawadzkiego (2002). Wynika z niej, że jest to model o podzielnikach p_i cyklu wahań okresowych (sezonowych) m , będącego liczbą parzystą, spełniający jednocześnie dwa warunki:

$$2 \leq p_i \leq \frac{m}{2} \quad \text{oraz} \quad \prod_i p_i = m. \quad (3)$$

Modele hierarchiczne dla wahań o cyklu 12-miesięcznym i 24-godzinym są oznaczane odpowiednio jako HM i HG. Liczby występujące po tych symbolach oznaczają kolejne podzielniki długości cyklu wahań.

Liczba regularnych modeli hierarchicznych dla danego cyklu wahań m jest równa liczbie permutacji oraz permutacji z powtórzeniami dzielników p_i . Dla danych miesięcznych o cyklu rocznym ($m = 12$) jest ich siedem:

- cztery dwustopniowe (HM26, HM34, HM43, HM62);
- trzy trzystopniowe (HM223, HM232, HM322).

Dla danych o cyklu dobowym ($m = 24$) jest ich 19:

- sześć dwustopniowych (HG2_12, HG38, HG46, H64, HG83, HG12_2);
- dziewięć trzystopniowych (HG226, HG234, HG243, HG262, HG342, HG423, HG432, HG226, HG622);
- cztery czterostopniowe (HG2223, HG2232, HG2322, HG3222).

Liczba szacowanych podwójnych modeli hybrydowych jest równa iloczynowi liczb modeli pojedynczych i wynosi $7 \cdot 19 = 133$. Dla jednej postaci analitycznej szacowanych jest zatem łącznie $7 + 19 + 133 = 159$ modeli hybrydowych. Do tego należy doliczyć jeden model klasyczny (z trzema zespołami zmiennych zero-jedynkowych) dany równaniem (1), służący jako punkt odniesienia dla modeli hybrydowych. Modele te pozwalają na „wyrównywanie” obserwacji w okresach będących skutkiem zakłóceń losowych lub zakłóceń występujących w okresach, w których istnieją luki w danych.

Poniżej podano zapisy analityczne dwu-, trzy- i czterostopniowych liniowych modeli z trendem liniowym (Szmuksta-Zawadzka i Zawadzki, 2014a).

Zapis dwustopniowego modelu hierarchicznego wygląda następująco:

$$Y_t = \alpha_1 t + \alpha_0 + \sum_{s=1}^{p_1} b_{0s} Q_{st} + \sum_{r=1}^{p_2} b_{0sr} Q_{srt} + U_{srt} \quad (4)$$

$$t = 1, 2, \dots, n$$

przy warunkach

$$\sum_{s=1}^{p_1} b_{0s} = \sum_{r=1}^{p_2} b_{0sr} = 0.$$

Model trzystopniowy wyraża się wzorem

$$Y_t = \alpha_1 t + \alpha_0 + \sum_{s=1}^{p_1} b_{0s} Q_{st} + \sum_{r=1}^{p_2} b_{0sr} Q_{srt} + \sum_{l=1}^{p_3} b_{0srll} Q_{srllt} + U_{srllt} \quad (5)$$

$$t = 1, 2, \dots, n$$

przy warunkach

$$\sum_{s=1}^{p_1} b_{0s} = \sum_{r=1}^{p_2} b_{0sr} = \sum_{l=1}^{p_3} b_{0srlt} = 0.$$

Model czterostopniowy można zapisać jako

$$Y_t = \alpha_1 t + \alpha_0 + \sum_{s=1}^{p_1} b_{0s} Q_{st} + \sum_{r=1}^{p_2} b_{0sr} Q_{srt} + \sum_{l=1}^{p_3} b_{0srlt} Q_{srlt} + \sum_{k=1}^{p_4} b_{0srlkt} Q_{srlkt} + U_{srlkt} \quad t = 1, 2, \dots, n \quad (6)$$

przy warunkach

$$\sum_{s=1}^{p_1} b_{0s} = \sum_{r=1}^{p_2} b_{0sr} = \sum_{l=1}^{p_3} b_{0srlt} = \sum_{k=1}^{p_4} b_{0srlkt} = 0.$$

Postać zmiennych Q zależy od liczby stopni hierarchii, liczby i kolejności podzielników modelu oraz długości cyklu wahań. Występujące we wzorach (4)–(6) zmienne zero-jedynkowe Q dla cyklu 24-godzinnego odnoszą się do 19 modeli dwu-, trzy- i czterostopniowych. Liczba zmiennych na każdym stopniu hierarchii, z uwagi na warunek sumowalności parametrów do 0, jest o 1 mniejsza od wartości podzielników p_1, p_2, p_3 lub p_4 . Przykładowo w modelu dwustopniowym HG12_2 ich liczba wynosi odpowiednio 11 i 1. Zmienne zero-jedynkowe Q_{st} ($s = 1, 2, \dots, 11$) przyjmują wartości równe 1 dla każdego okresu dwugodzinnego: $Q_{1t} = 1$ dla godziny 1 i 2 oraz 0 dla pozostałych godzin cyklu dobowego, Q_{2t} dla 3 i 4, ..., Q_{11t} dla 21 i 22. Natomiast zmienna Q_{srt} ($r = 1, 2$) jest tylko jedna, ponieważ $p_2 - 1 = 1$. Zmienna Q_{s1t} przyjmuje wartość 1 dla pierwszej (nieparzystej) godziny każdego okresu dwugodzinnego.

W siedmiu dwu- i trzystopniowych modelach hierarchicznych opisujących wachania o cyklu rocznym zmienne Q przyjmują inne wartości. Przykładowo dla modelu HM43 podzielniki wynoszą: $p_1 = 4$ i $p_2 = 3$. Liczby zmiennych Q są o 1 mniejsze. Zmienne zero-jedynkowe $Q_{st(M)}$ ($s = 1, 2, 3$) przyjmują wartości 1 w s -tym kwartale i 0 w pozostałych, a zmienne $Q_{srt(M)}$ ($r = 1, 2$) – wartości 1 w r -tych miesiącach każdego kwartału. Wszystkie omawiane w pracy modele wykazujące złożoną sezonowość: klasyczne szeregu czasowego, regresyjne oraz hierarchiczne (hybrydowe) są modelami jednorównaniowymi.

Idea budowy i konwencja zapisu regularnych modeli hierarchicznych zaproponowana przez Szumksta-Zawadzka i Zawadzkiego (2002) dla danych miesięcznych (roz-

szerzona potem na dane dekadowe, dzienne i godzinne) jest koncepcją uniwersalną i odnosi się do modeli opisujących wahania sezonowe o parzystej długości cykli, zarówno niezłożonych, jak i złożonych. Jest ona podobna do prezentowanej m.in. w pracy Athanasopoulou i in. (2017) dla danych miesięcznych. Podobieństwo dotyczy agregacji na okresy dwu-, trzy-, cztero- i sześciomiesięczne. Różnica, oprócz sposobu zapisu, polega natomiast na tym, że na najniższym stopniu hierarchii liczy się średnie dla okresów jednoimiennych. Przykładowo w przypadku wahań o okresie 24-godzinny w modelu HG12_2 liczy się średnie dla godzin parzystych i nieparzystych.

W modelach hierarchicznych parametry strukturalne są uśrednionymi parametrami klasycznych modeli szeregu czasowego ze zmiennymi zero-jedynkowymi. Sposób uśredniania zależy od liczby stopni hierarchii oraz liczby i kolejności podzielników.

Dla cyklu 24-godzinnego w modelach (1) i (2), przy uwzględnieniu warunku sumowalności do 0, szacowane są 23 parametry d_{0k} ($k = 1, 2, \dots, 23$). Ocenę parametru d_{024} oblicza się jako sumę 23 parametrów i nadaje się jej przeciwny znak. Natomiast w dwustopniowym modelu hierarchicznym HG12_2 szacowanych jest 11 parametrów b_{0s} ($s = 1, 2, \dots, 11$) na pierwszym stopniu hierarchii i 1 parametr na drugim stopniu – b_{0s1} . Parametry b_{0s} są średnimi z parametrów d_{0k} dla kolejnych okresów dwugodzinnych: $b_{01} = 1/2 \cdot (d_{01} + d_{02})$; $b_{02} = 1/2 \cdot (d_{03} + d_{04})$; ...; $b_{011} = 1/2 \cdot (d_{021} + d_{022})$. Parametr b_{012} wyznacza się z warunku sumowalności do 0. Jedynie parametr b_{0s1} jest szacowany bezpośrednio jako średnia z parametrów d_{0k} dla godzin nieparzystych: $b_{0s1} = 1/12 \cdot (d_{01} + d_{03} + d_{05} + \dots + d_{023})$. Z warunku sumowalności do 0 wynika, że $b_{0s2} = -b_{0s1}$.

Literatura poświęcona metodom prognozowania brakujących danych w szeregach czasowych jest dość obszerna, jednak dominują prace dotyczące prognozowania dla okresów jednostkowych o cyklu co najwyżej tygodniowym. Spośród nich należy przywołać monografie o charakterze metodologicznym autorstwa Allisona (2001), Little'a i Rubina (2002) oraz Raudenbusha i Bryka (2002). Prace podejmujące problematykę prognozowania brakujących danych w godzinnych szeregach czasowych lub dla okresów jednostkowych krótszych niż godzina należą raczej do rzadkości i na ogół odnoszą się do prognozowania zjawisk meteorologicznych. Przykładowo można tu wymienić pracę Henna i in. (2013), poświęconą prognozowaniu temperatury. Weron (2006) rozpatruje prognozowanie w szeregach z lukami, w których zapotrzebowanie na energię jest mierzone w odstępach 5-minutowych. Metoda polega na zastępowaniu brakujących danych średnimi z obserwacji z dłuższych odcinków czasu, z sąsiadujących obserwacji lub np. z poprzedniego tygodnia, po wcześniejszym wyeliminowaniu obserwacji odstających. Badania nad zastosowaniami metod prognozowania brakujących danych w szeregach czasowych z potrójnie złożonymi wahaniami sezonowymi (dla danych godzinnych) były prowadzone do roku 2016 przez Szumuksta-Zawadzką i Zawadzkiego (2011, 2014a, 2015), a następnie przez Zawadzkiego (2018).

Celem niniejszego artykułu jest przedstawienie możliwości prognozowania brakujących danych w szeregach czasowych dla danych godzinnych przy zastosowaniu

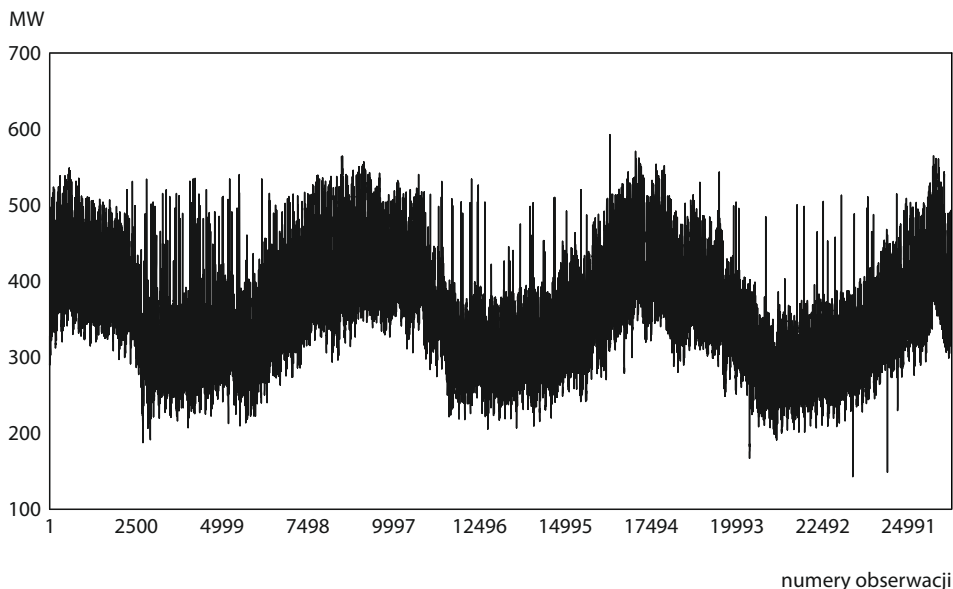
modeli hybrydowych. W omawianym badaniu wykorzystano hybrydowe modele szeregu czasowego oraz modele regresyjne¹ ze złożonymi wahaniami sezonowymi.

2. Metoda badania

Modelowaniu, a następnie prognozowaniu poddano zapotrzebowanie na energię elektryczną w okresach godzinnych w aglomeracji A. Dane statystyczne o kształtowaniu się zapotrzebowania na moc zostały udostępnione przez przedsiębiorstwo zajmujące się dystrybucją energii elektrycznej i włączone do Banku Danych Katedry Zastosowań Matematyki w Ekonomii Zachodniopomorskiego Uniwersytetu Technologicznego (ZUT) w Szczecinie. Odnoszą się one do trzech kolejnych lat pierwszej dekady XXI w.² Szereg wyjściowy (bez luk) obejmował dwa lata, tj. $2 \cdot 365 \cdot 24 = 17520$ obserwacji ($t = 1, 2, \dots, 17520$). Trzeci rok był okresem empirycznej weryfikacji prognoz i zawierał $365 \cdot 24 = 8760$ obserwacji ($t = 17521, \dots, 26280$).

Kształtowanie się zmiennej w okresie estymacyjnym i prognozowanym zostało przedstawione na wykr. 1.

Wykr. 1. Zapotrzebowanie na moc elektryczną



Źródło: Bank Danych Katedry Zastosowań Matematyki w Ekonomii ZUT w Szczecinie.

¹ Określenie *modele regresyjne* przyjęto m.in. za pracą Witkowskiej i in. (2000).

² Warunkiem udostępnienia danych było zobowiązanie do zachowania anonimowości firmy i nieujawniania konkretnych lat, których dotyczą.

W pracy Zawadzkiego (2018) rozpatrywano wariant, w którym luki niesystematyczne występowały w ośmiu godzinach: 2, 6, 8, 10, 14, 16, 18, 22, w okresie od 1 lipca pierwszego roku do 30 czerwca drugiego roku (wariant G). Łączna ich liczba wynosiła 2920 ($365 \cdot 8$). W badaniu omawianym w artykule założono, że luki niesystematyczne, oprócz wymienionych godzin, występują także w dwóch dniach każdego z 52 tygodni (wtorki i czwartki) i każdego z 62 dni dwóch miesięcy (sierpień i październik) – wariant GDM. W wyznaczeniu liczby luk dla tygodni należy wziąć pod uwagę fakt ich wcześniejszego uwzględnienia dla godzin. W każdym z dwóch dni tygodnia liczba luk zwiększy się o 16 godzin. Otrzymamy zatem $52 \cdot 2 \cdot 16$ godzin = 1664 godziny. Natomiast w przypadku wymienionych miesięcy należy uwzględnić występowanie luk dla godzin i tygodni. Przyrost liczby luk wyniesie więc $(62 - 19) \cdot 16 = 688$ godzin. Oznacza to zwiększenie łącznej liczby luk w tym wariantcie z 2920 do 5272, tj. o ponad 80%. Luki otrzymano przez wymazanie odpowiednich danych z pełnego szeregu.

Rozmieszczenie luk osobno w przypadku ośmiu godzin dla każdego z 365 dni, w przypadku dwóch dni tygodnia dla każdego z 52 tygodni i w przypadku 62 dni dla dwóch miesięcy przedstawiono w zestawieniach 1–3. Puste komórki oznaczają występowanie luk, a przez x oznaczono dostępne dane. Występowanie pustych komórek przynajmniej w jednej kolumnie oznacza luki systematyczne.

Zestawienie 1. Układ luk dla danych godzinnych

Lata i półrocza	Godziny																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
I: 1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2	x		x	x	x		x		x		x	x	x		x		x		x	x	x		x	x
II: 1	x		x	x	x		x		x		x	x	x		x		x		x	x	x		x	x
2	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Źródło: opracowanie własne.

Zestawienie 2. Układ luk dla dni tygodnia

Lata i półrocza	Dni tygodnia						
	poniedziałek	wtorek	środa	czwartek	piątek	sobota	niedziela
I: 1		x		x		x	x
2		x		x		x	x
II: 1	x			x		x	x
2		x		x		x	x

Źródło: opracowanie własne.

Zestawienie 3. Układ luk dla miesięcy

Lata i półrocza	Miesiące											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
I: 1	x	x	x	x	x	x	x	x	x	x	x	x
2	x	x	x	x	x	x	x	x	x	x	x	x
II: 1	x	x	x	x	x	x	x	x	x	x	x	x
2	x	x	x	x	x	x	x	x	x	x	x	x

Źródło: opracowanie własne.

Do modelowania i prognozowania wykorzystano pojedyncze i podwójne hybrydowe modele szeregu czasowego i regresyjne. Hybrydowe modele regresyjne, będące rozbudowanymi modelami szeregu czasowego, zawierają opóźnioną o 24 godziny zmienną endogeniczną oraz zmienne zero-jedynkowe opisujące występowanie świąt i jednego dnia przedświątecznego (Wielkiej Soboty). Wprowadzenie dni świątecznych uzasadnia się znacznie niższym poziomem zapotrzebowania na energię elektryczną w tych dniach (zob. Witkowska i in., 2000).

Poniżej podano zapisy analityczne dwu-, trzy- i czterostopniowych hierarchicznych liniowych modeli regresyjnych. Zapis dwustopniowego regresyjnego modelu hierarchicznego jest następujący:

$$Y_t = \alpha_1 t + \alpha_0 + \sum_{s=1}^{p_1} b_{0s} Q_{st} + \sum_{r=1}^{p_2} b_{0sr} Q_{srt} + \beta Y_{t-24} + \sum_{z=1}^{10} \gamma_z S_{zt} + U_{srlt} \quad (7)$$

$$t = 1, 2, \dots, n$$

przy warunkach

$$\sum_{s=1}^{p_1} b_{0s} = \sum_{r=1}^{p_2} b_{0sr} = 0.$$

Model trzystopniowy wyraża się wzorem

$$Y_t = \alpha_1 t + \alpha_0 + \sum_{s=1}^{p_1} b_{0s} Q_{st} + \sum_{r=1}^{p_2} b_{0sr} Q_{srt} + \sum_{l=1}^{p_3} b_{0srll} Q_{srll} + \beta Y_{t-24} +$$

$$+ \sum_{z=1}^{10} \gamma_z S_{zt} + U_{srll} \quad t = 1, 2, \dots, n \quad (8)$$

przy warunkach

$$\sum_{s=1}^{p_1} b_{0s} = \sum_{r=1}^{p_2} b_{0sr} = \sum_{l=1}^{p_3} b_{0sr}l = 0.$$

Model czterostopniowy można zapisać następująco:

$$\begin{aligned} Y_t = & \alpha_1 t + \alpha_0 + \sum_{s=1}^{p_1} b_{0s} Q_{st} + \sum_{r=1}^{p_2} b_{0sr} Q_{srt} + \sum_{l=1}^{p_3} b_{0sr}l Q_{sr}l t + \\ & + \sum_{k=1}^{p_4} b_{0sr}l k t Q_{sr}l k t + \beta Y_{t-24} + \sum_{z=1}^{10} \gamma_z S_{zt} + U_{sr}l t \quad t = 1, 2, \dots, n \end{aligned} \quad (9)$$

przy warunkach

$$\sum_{s=1}^{p_1} b_{0s} = \sum_{r=1}^{p_2} b_{0sr} = \sum_{l=1}^{p_3} b_{0sr}l = \sum_{k=1}^{p_4} b_{0sr}l k t = 0.$$

Zmienne zero-jedynkowe S_{zt} przyjmują wartość 1 w następujące dni świąteczne: Nowy Rok, Wielkanoc, 1 i 3 Maja, Boże Ciało, Wniebowzięcie Najświętszej Marii Panny (15 sierpnia), Wszystkich Świętych (1 listopada), Dzień Niepodległości (11 listopada), Boże Narodzenie (25 grudnia) i Wielką Sobotę.

W modelach wykładniczych zamiast Y_{t-24} występuje $\ln Y_{t-24}$.

W modelach hybrydowych przynajmniej jeden rodzaj wahań o parzystej długości cyklu (24-godzinne lub/i 12-miesięczne) jest opisywany za pomocą regularnych modeli hierarchicznych.

Dla obu rodzajów modeli (szeregu czasowego i regresyjnych) szacowano także dwie postaci analityczne: liniową i wykładniczą.

Parametry strukturalne zarówno klasycznych modeli szeregu czasowego ze złożonymi wahaniami sezonowymi, dane równaniami (1) i (2), jak i modeli regresyjnych szacowano klasyczną metodą najmniejszych kwadratów, uwzględniając warunki sumowalności parametrów do 0. Estymację wykonano na podstawie dostępnych danych, tzn. z wyłączeniem okresów, w których wystąpiły luki. Tę samą metodę zastosowano do szacowania parametrów modeli hybrydowych, zarówno pojedynczych, jak i podwójnych. Do estymacji parametrów wykorzystano pakiet Statistica 13.3, a do obliczenia prognoz oraz ich błędów względnych – arkusz kalkulacyjny Excel³.

³ W literaturze (zob. np. Athanasopoulos i in., 2017) można znaleźć informacje o pakietach w języku R przeznaczonych do szacowania modeli hierarchicznych i prognozowania na ich podstawie; są to pakiety *thief* i *forecast*.

Na podstawie predyktorów opartych na oszacowanych równaniach wyznaczano dwa rodzaje prognoz: interpolacyjne i ekstrapolacyjne. Dla okresów, w których wystąpiły luki, wyznaczono prognozy interpolacyjne. Otrzymano je przez podstawienie wartości zmiennych występujących w tym okresie po prawej stronie każdego z oszacowanych modeli. W omawianym badaniu w wariancie GDM (w którym luki występują w ośmiu okresach godzinnych, dwóch dniach tygodnia oraz dwóch miesiącach) liczba ta wynosi 5272. Z kolei dla okresu empirycznej weryfikacji prognoz, wybiegającego poza przedział czasowy „próby”, dla $t = 17521, \dots, 26280$, wyznaczono 8760 prognoz ekstrapolacyjnych *ex post*. Otrzymano je przez podstawienie wartości zmiennych występujących po prawej stronie równań w okresie prognozowanym ($t > 17520$).

Ponieważ dla obu rodzajów prognoz znane były realizacje zmiennej prognozowanej, obliczono dla nich średnie błędy względne dane wzorami (10)–(13).

3. Wyniki modelowania i prognozowania dla modeli hybrydowych szeregu czasowego

Z badań dotyczących prognozowania brakujących danych, m.in. w szeregach czasowych z podwójnie złożonymi wahaniami sezonowymi (dla danych dziennych), wynika, że najmniejsze wartości oceny błędów prognoz interpolacyjnych i ekstrapolacyjnych otrzymuje się zazwyczaj dla modeli hybrydowych o innej strukturze lub innej postaci analitycznej (zob. np. Szmuksta-Zawadzka i Zawadzki, 2014b). Inna struktura oznacza, że modele różnią się liczbą lub kolejnością podzielników względnie rodzajem modelu (pojedynczy lub podwójny). Natomiast inna postać analityczna oznacza, że np. model o najmniejszych błędach prognoz interpolacyjnych może być liniowy, a prognoz ekstrapolacyjnych – wykładniczy. Ilustruje to przykład liniowego hybrydowego predyktora pojedynczego L_GDM_HG122_0/1 i wykładniczego predyktora podwójnego W_GDM_HG122_HM43.

Do prognozowania wybrano predyktory charakteryzujące się najmniejszymi wartościami oceny względnych błędów prognoz interpolacyjnych ($MAPE_I$) bądź prognoz ekstrapolacyjnych ($MAPE_E$). Przy zastosowaniu każdego z tych kryteriów obliczono, jako wielkości pomocnicze, błędy drugiego rodzaju prognoz. Oznaczono je odpowiednio $MAPE_{I,E}$ oraz $MAPE_{E,I}$. Pozwoliło to na porównanie wartości tego samego miernika przy zastosowaniu różnych kryteriów dotyczących wyboru modelu dla celów prognozowania, które polegają na minimalizacji błędów $MAPE_I$ lub $MAPE_E$. Jeżeli kryterium stanowi minimalizacja $MAPE_I$, to drugim rodzajem prognoz są prognozy ekstrapolacyjne, dla których oblicza się miernik $MAPE_{I,E}$. Można go porównać z miernikiem $MAPE_E$ wyznaczonym przy zastosowaniu kryterium

minimalizacji błędów prognoz ekstrapolacyjnych. To samo odnosi się do prognoz interpolacyjnych $MAPE_I$ oraz $MAPE_{E,I}$.

Wzory dotyczące wymienionych mierników mają postać:

$$MAPE_I = \frac{1}{n_1} \sum_{t \in I_{luk}} \frac{|Y_t - Y_{tp}|}{Y_t}, \quad (10)$$

$$MAPE_{I,E} = \frac{1}{T} \sum_{t=n+1}^T \frac{|Y_t - Y_{tp}|}{Y_t}, \quad (11)$$

$$MAPE_E = \frac{1}{T} \sum_{t=n+1}^T \frac{|Y_t - Y_{tp}|}{Y_t}, \quad (12)$$

$$MAPE_{E,I} = \frac{1}{n_1} \sum_{t \in I_{luk}} \frac{|Y_t - Y_{tp}|}{Y_t}, \quad (13)$$

gdzie:

Y – realizacje zmiennej prognozowanej,

Y_{tp} – prognozy,

n_1 – liczba brakujących danych (prognoz interpolacyjnych),

$t \in I_{luk}$ – t należy do okresów, w których występują luki,

n – liczba obserwacji dla pełnego szeregu (bez luk),

T – liczba prognoz ekstrapolacyjnych.

Wielkości wybranych parametrów struktury stochastycznej dla najlepszych pojedynczych i podwójnych modeli hybrydowych szeregu czasowego w postaci liniowej i wykładniczej charakteryzujących się najmniejszymi wartościami oceny błędów $MAPE_I$ zestawiono w tabl. 1. W ostatniej kolumnie podano oceny błędów prognoz ekstrapolacyjnych otrzymanych przy zastosowaniu kryterium $MAPE_{I,E}$.

Oszacowanym modelom hybrydowym nadano nazwy literowo-liczbowe. Pierwsza litera oznacza postać analityczną: liniową – L lub wykładniczą – W, a druga – kryterium wyboru najlepszego równania minimalizującego błędy prognoz: interpolacyjnych ($MAPE_I$) – I; ekstrapolacyjnych ($MAPE_E$) – E. Kolejne trzy litery – GDM, oznaczające występowanie luk w każdym z trzech rodzajów wahań sezonowych – wprowadzono ze względu na wykonane na dalszym etapie badania porównanie własności predyktywnych i prognoz otrzymanych dla najlepszych modeli z lukami występującymi w okresach godzinnych (G). Litery HG i HM odnoszą się do modeli hierarchicznych o cyklu odpowiednio dobowym i rocznym, a liczby po nich wystę-

pujące są kolejnymi podzielnikami długości cyklu wahań. W pojedynczych modelach hybrydowych występują oznaczenia 0/1, mówiące o tym, że dany rodzaj wahań jest opisywany za pomocą zmiennych zero-jedynkowych.

W dwóch ostatnich wierszach tabl. 1 i 2 podano charakterystyki liniowej i wykładniczej postaci modeli klasycznych L_GDM_0/1_0/1 i W_GDM_0/1_0/1. Dla każdej postaci analitycznej szacowano jeden model (z tego powodu po literze L lub W nie występują oznaczenia I lub E) i dlatego w przypadku kryterium $MAPE_I$ w nawiasach podano błędy prognoz ekstrapolacyjnych, a w przypadku $MAPE_E$ – błędy prognoz interpolacyjnych.

Tabl. 1. Oceny parametrów struktury stochastycznej i błędów prognoz modeli hybrydowych przy zastosowaniu kryterium $MAPE_I$

Modele	R^2	S_e w MW	V_{Se}	$MAPE_I$	$MAPE_{I,E}$
			w %		
LI_GDM_HG122_0/1	0,7980	32,76	8,969	6,665	9,817
WI_GDM_HG122_0/1	0,8106	32,49	8,823	6,584	9,307
LI_GDM_0/1_HM62	0,7827	33,99	9,306	7,123	10,480
WI_GDM_0/1_HM62	0,7982	33,55	9,108	6,985	9,947
LI_GDM_HG122_HM43	0,7710	34,87	9,547	7,330	10,620
WI_GDM_HG122_HM43	0,7860	34,53	9,375	7,248	10,110
L_GDM_0/1_0/1	0,8096	31,82	8,712	6,434	(9,076)
W_GDM_0/1_0/1	0,8227	31,45	8,540	6,308	(8,554)

Uwaga. R^2 – współczynnik determinacji, S_e – odchylenie standardowe składnika losowego, V_{Se} – współczynnik zmienności losowej.

Źródło: opracowanie własne.

Z informacji dotyczących kształtowania się wartości R^2 wynika, że wśród modeli klasycznych wyższą oceną – o 1,31 p.proc. – charakteryzuje się model wykładniczy (W_GDM_0/1_0/1), którego oceny parametrów S_e i V_{Se} mają wartości niższe odpowiednio o 0,37 MW i 0,172 p.proc. Najlepszym modelem hybrydowym okazał się jednostopniowy model wykładniczy WI_GDM_HG122_0/1 o współczynniku determinacji mniejszym o 1,21 p.proc. od lepszego z modeli klasycznych. W tym modelu wahań o cyklu dobowym zostały opisane za pomocą modelu hierarchicznego o podzielnikach 12 i 2. Oceny parametrów S_e i V_{Se} dla tego modelu przyjęły wartości nieznacznie większe od otrzymanych dla modelu klasycznego, odpowiednio o 1,04 MW i 0,28 p.proc.

Wśród modeli hybrydowych opisujących dwa rodzaje wahań lepszy okazał się model w postaci wykładniczej (WI_GDM_HG122_HM43), dla którego wartość R^2 wyniosła 0,7860. Najniższą oceną R^2 charakteryzuje się model LI_GDM_HG122_HM43; jest ona o ok. 5,2 p.proc. niższa od otrzymanej dla modelu klasycznego w postaci wykładniczej.

Najmniejsza wartość oceny błędów $MAPE_I$ otrzymana dla modeli hybrydowych (model pojedynczy WI_GDM_HG122_0/1) jest o 0,276 p.proc. wyższa od oceny minimalnej otrzymanej dla modelu klasycznego W_GDM_0/1_0/1. Niższe oceny

błędów prognoz interpolacyjnych, podobnie jak charakterystyk własności prognozy stycznych, otrzymano dla predyktorów wykładniczych.

W tabl. 2 zestawiono modele hybrydowe charakteryzujące się najmniejszymi wartościami błędów $MAPE_E$ (przedostatnia kolumna). Ostatnia kolumna zawiera oceny błędów $MAPE_{E,I}$.

Tabl. 2. Oceny parametrów struktury stochastycznej i błędów prognoz modeli hybrydowych szeregu czasowego przy zastosowaniu kryterium $MAPE_E$

Modele	R^2	S_e w MW	V_{Se}	$MAPE_E$	$MAPE_{E,I}$
			w %		
LE_GDM_HG122_0/1	0,7980	32,76	8,969	9,820	6,665
WE_GDM_HG122_0/1	0,8106	32,49	8,823	9,310	6,584
LE_GDM_0/1_HM43	0,7371	37,39	10,236	9,670	7,880
WE_GDM_0/1_HM43	0,7530	37,11	10,075	9,260	7,728
LE_GDM_HG122_HM34	0,7258	38,16	10,447	9,830	8,086
WE_GDM_HG83_HM43	0,7262	39,05	10,602	9,650	8,145
L_GDM_0/1_0/1	0,8096	31,82	8,712	9,699	(6,434)
W_GDM_0/1_0/1	0,8227	31,45	8,540	9,129	(6,308)

Uwaga. Jak przy tabl. 1.

Źródło: opracowanie własne.

Najwyższą ocenę R^2 otrzymano dla pojedynczego modelu hybrydowego WE_GDM_HG122_0/1, tak samo jak w przypadku kryterium $MAPE_I$, a najniższą – dla modelu podwójnego WE_GDM_HG83_HM43. Maksymalną ocenę uzyskał ten sam model co w przypadku kryterium $MAPE_E$; jest ona niższa o 1,21 p.proc. od oceny wykładniczej postaci modelu klasycznego (W_GDM_0/1_0/1). Również dla tych samych modeli otrzymano minimalne i maksymalne oceny S_e i V_{Se} . Ocena minimalna była o ok. 0,28 p.proc. wyższa od otrzymanej dla modelu klasycznego.

Najmniejszą wartość oceny błędów $MAPE_E$ otrzymano dla predyktora opartego na modelu WE_GDM_0/1_HM43; była ona tylko o 0,053 p.proc. mniejsza od otrzymanej dla modelu o najlepszych własnościach predykcyjnych WE_GDM_HG122_0/1 i zaledwie nieznacznie większa od wartości oceny błędu otrzymanego dla modelu klasycznego. Z dwóch postaci analitycznych modeli hybrydowych w przypadku tego kryterium, podobnie jak w przypadku $MAPE_I$, niższe oceny otrzymano dla postaci wykładniczej.

4. Wyniki modelowania i prognozowania dla regresyjnych modeli hybrydowych

Regresyjne równania hybrydowe, jak już wspomniano, różnią się od równań hybrydowych szeregu czasowego tym, że zawierają dodatkowo opóźnioną o 24 godziny zmienną endogeniczną oraz zmienne zero-jedynkowe oznaczające dni świąteczne i Wielką Sobotę.

W tabl. 3 zestawiono regresyjne modele hybrydowe charakteryzujące się najmniejszymi wartościami oceny błędów prognoz interpolacyjnych ($MAPE_t$). Litera P występująca na początku ich nazwy oznacza model typu regresyjnego. Własności predyktywne oraz błędy prognoz interpolacyjnych porównano z odpowiednimi charakterystykami otrzymanymi dla lepszego z predyktorów klasycznych.

Tabl. 3. Oceny parametrów struktury stochastycznej i błędów prognoz regresyjnych modeli hybrydowych przy zastosowaniu kryterium $MAPE_t$

Modele	R^2	S_e w MW	V_{Se}	$MAPE_t$	$MAPE_{L,E}$
			w %		
PLI_GDM_HG122_0/1	0,8454	28,66	7,847	5,452	7,202
PWI_GDM_HG122_0/1	0,8508	28,60	7,823	5,485	6,961
PLI_GDM_0/1_HM62	0,8413	29,05	7,953	5,581	7,308
PWI_GDM_0/1_HM62	0,8477	28,90	7,912	5,596	7,081
PLI_GDM_HG122_HM62	0,8384	29,30	8,020	5,597	7,256
PWI_GDM_HG122_HM62	0,8441	29,23	8,002	5,647	7,046
PL_GDM_0/1_0/1	0,8490	28,34	7,758	5,408	(7,244)
PW_GDM_0/1_0/1	0,8553	28,18	7,716	5,406	(6,983)

Uwaga. Jak przy tabl. 1.

Źródło: opracowanie własne.

Najniższa ocena R^2 charakteryzowała podwójny model liniowy PLI_GDM_HG122_HM62, a najwyższa – model pojedynczy w postaci wykładniczej PWI_GDM_HG122_0/1; była niższa tylko o 0,45 p.proc. od oceny otrzymanej dla wykładniczej postaci modelu klasycznego PW_GDM_0/1_0/1. Zwraca uwagę fakt, że rozstępy R^2 dla regresyjnych modeli hybrydowych były znacznie mniejsze od uzyskanych dla modeli szeregu czasowego i wyniosły odpowiednio 1,24 p.proc. i 3,96 p.proc.

Najniższe oceny S_e i V_{Se} otrzymano także dla modelu wykładniczego PWI_GDM_HG122_0/1. Były one nieznacznie wyższe od ocen tych parametrów charakteryzujących lepszy z modeli klasycznych (PW_GDM_0/1_0/1), odpowiednio o 0,42 MW i 0,107 p.proc. Natomiast najmniejszą wartość oceny błędów prognoz interpolacyjnych otrzymano nie dla jednostopniowego modelu wykładniczego, o najlepszych własnościach predyktywnych (PWI_GDM_HG122_0/1), lecz dla postaci liniowej tego modelu, a różnica wyniosła 0,69 p.proc. Klasyczny model wykładniczy charakteryzował się błędem mniejszym tylko o 0,10 p.proc.

Z porównania ocen błędów prognoz interpolacyjnych otrzymanych na podstawie modeli w postaci liniowej i wykładniczej o strukturze hierarchicznej tego samego rodzaju wynika, że są one, odmiennie niż w przypadku modeli szeregu czasowego, nieco wyższe dla postaci wykładniczej.

W tabl. 4 zestawiono hybrydowe modele regresyjne charakteryzujące się najmniejszymi wartościami oceny błędów $MAPE_E$. Za punkt odniesienia w porówna-

niach własności predyktywnych oraz ocen prognoz ekstrapolacyjnych przyjęto model klasyczny typu regresyjnego ze zmiennymi zero-jedynkowymi o niższej ocenie błędu tego rodzaju (PW_GDM_0/1_0/1).

Tabl. 4. Oceny parametrów struktury stochastycznej i błędów prognoz regresyjnych modeli hybrydowych przy zastosowaniu kryterium $MAPE_E$

Modele	R^2	S_e w MW	V_{Se}	$MAPE_E$	$MAPE_{E,l}$
			w %		
PLE_GDM_HG2322_0/1	0,8222	30,73	8,413	6,877	5,700
PWE_GDM_HG2322_0/1	0,8233	31,12	8,519	6,809	5,850
PLE_GDM_0/1_HM43	0,8313	29,95	8,199	6,702	5,699
PWE_GDM_0/1_HM43	0,8364	29,96	8,202	6,535	5,693
PLE_GDM_HG2322_HM43	0,8148	31,35	8,584	6,591	5,792
PWE_GDM_HG38_HM43	0,8275	30,74	8,416	6,546	5,874
PL_GDM_0/1_0/1	0,8490	28,34	7,758	7,244	(5,408)
PW_GDM_0/1_0/1	0,8553	28,18	7,772	6,983	(5,406)

Uwaga. Jak przy tabl. 1.

Źródło: opracowanie własne.

Wśród modeli klasycznych wyższą ocenę R^2 uzyskała postać wykładnicza modelu klasycznego PW_GDM_0/1_0/1; jest ona o 1,89 p.proc. wyższa od najwyższej oceny otrzymanej dla najlepszego modelu hybrydowego PWE_GDM_0/1_HM43. Najniższą ocenę S_e otrzymano dla modelu PLE_GDM_0/1_HM43, który charakteryzował się też najniższą oceną V_{Se} , a najwyższą – dla podwójnego liniowego modelu hybrydowego PLE_GDM_HG2322_HM43.

Oceny błędów prognoz ekstrapolacyjnych ($MAPE_E$) były tylko nieznacznie zróżnicowane. Błąd prognozy dla lepszego z podwójnych modeli hybrydowych (PWE_GDM_HG38_HM43) był większy o 0,16 p.proc. od najmniejszego błędu, a dla lepszego z modeli klasycznych (PW_GDM_0/1_0/1) – o 0,48 p.proc. większy od najmniejszego błędu otrzymanego dla regresyjnych modeli hybrydowych. Model klasyczny w rankingu 160 modeli wykładniczych (159 hybrydowych i 1 modelu klasycznego) znalazł się na 158. miejscu.

Do zbadania identyczności dystrybuant obu rozkładów empirycznych modułów względnych błędów prognoz wykorzystano składowe miernika $MAPE_E$ w liczbie 8760, oznaczone jako APE_{Et} . Posłużono się testem λ Kołmogorowa-Smirnowa (np. Domański, 1990, s. 129–130); przyjęto poziom istotności $\alpha = 0,05$. Statystyka empiryczna λ tego testu dla 8760 obserwacji w każdym modelu wynosi: $\lambda = D_{nm}\sqrt{n} = 0,03744 \cdot \sqrt{\frac{8760 \cdot 8760}{8760 + 8760}} = 0,03744 \cdot 66,182 = 2,4780$.

Następnie z tablicy rozkładu granicznego λ dla poziomu istotności $\alpha = 0,05$ odczytano wartość krytyczną $\lambda_\alpha = 1,354$. W teście tym obszar krytyczny jest prawo-

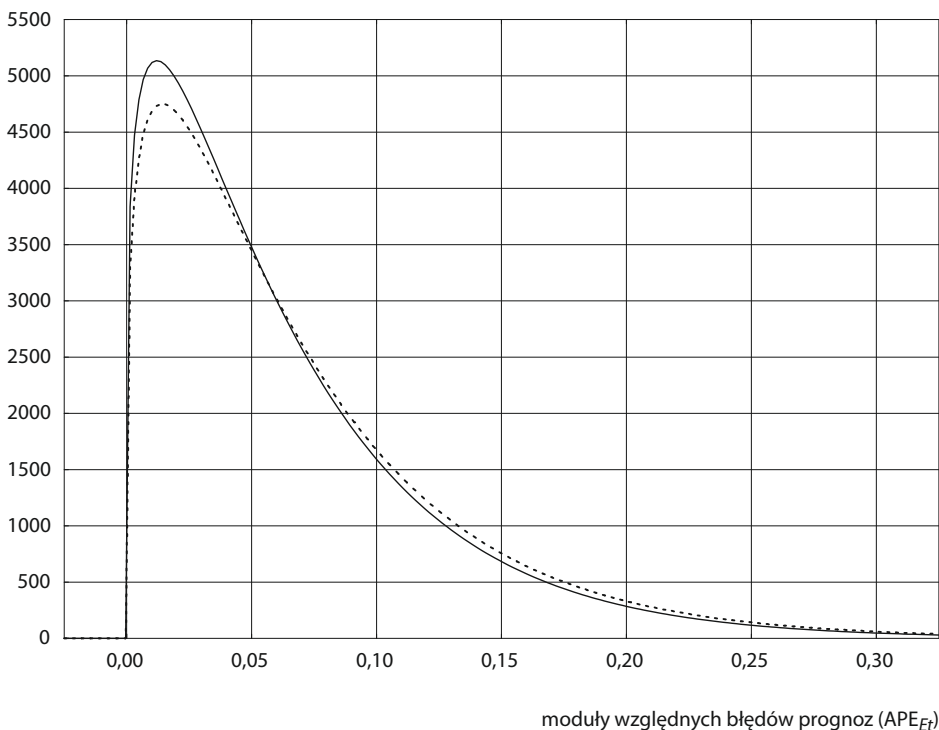
stronny. Ponieważ $\lambda = 2,4780 > 1,354 = \lambda_{\alpha}$, to hipotezę zerową należy odrzucić na korzyść hipotezy alternatywnej. Nie można więc twierdzić, że wartości APE_{Et} klasycznego modelu regresyjnego (PW_GDM_0/1_0/1) są takie same jak w przypadku regresyjnego modelu hybrydowego (PWE_GDM_0/1_HM43).

Uwzględniając powyższą decyzję weryfikacyjną oraz wcześniejsze rozważania dotyczące poziomu błędów prognoz ekstrapolacyjnych uzyskanych dla modelu klasycznego i modelu hybrydowego, można przyjąć, że błędy prognoz ekstrapolacyjnych w przypadku modelu klasycznego są znacząco większe od błędów modelu hybrydowego.

Empiryczne rozkłady modułów wartości APE_{Et} były najlepiej aproksymowane za pomocą rozkładu gamma (wykr. 2). Linią ciągłą oznaczono błędy w modelu hybrydowym PWE_GDM_0/1_HM43, a linią przerywaną – w modelu klasycznym.

Wykr. 2. Rozkład modułów względnych błędów prognoz ekstrapolacyjnych otrzymanych na podstawie modeli PWE_GDM_0/1_HM43 i PWE_GDM_0/1_0/1

liczba obserwacji



Źródło: opracowanie własne.

W tabl. 5 zestawiono najlepsze modele hybrydowe szeregu czasowego i modele regresyjne charakteryzujące się najmniejszymi wartościami oceny błędów $MAPE_I$ i $MAPE_E$. Ze względu na porównywanie własności predyktywnych i prognoz otrzymanych dla modeli hybrydowych i modeli klasycznych pominięto oceny błędów $MAPE_{I,E}$ oraz $MAPE_{E,I}$. Tablica 5 zawiera także zestawienie tych wielkości dla modeli klasycznych charakteryzujących się lepszymi własnościami predyktywnymi i niższymi ocenami błędów prognoz, co pozwala na porównanie struktury modeli oraz dokładności obu rodzajów prognoz. W ostatnich dwóch wierszach podano charakterystyki własności predyktywnych i błędów prognoz otrzymane na podstawie obu rodzajów modeli klasycznych (z trzema zespołami zmiennych zero-jedynkowych) jako punkt odniesienia dla modeli hybrydowych.

Tabl. 5. Oceny parametrów struktury stochastycznej i błędów prognoz hybrydowych modeli szeregu czasowego i regresyjnych przy zastosowaniu kryteriów $MAPE_I$ oraz $MAPE_E$

Modele	R^2	S_e w MW	V_{se}	$MAPE_I$	$MAPE_E$
			w %		
WI_GDM_HG122_0/1	0,8106	32,49	8,823	6,584	.
PLI_GDM_HG122_0/1	0,8454	28,66	7,847	5,452	.
WE_GDM_0/1_HM43	0,7530	37,11	10,075	.	9,257
PWE_GDM_0/1_HM43	0,8364	29,96	8,202	.	6,535
L_GDM_0/1_0/1	0,8227	31,45	8,540	6,308	9,129
PW_GDM_0/1_0/1	0,8553	28,18	7,772	5,406	6,983

Uwaga. Jak przy tabl. 1.

Źródło: opracowanie własne.

W pracy Szmuksta-Zawadzkiej i Zawadzkiego (2014a) przedstawiono wyniki prognozowania brakujących danych na podstawie hybrydowych modeli przyczynowo-skutkowych dla luk niesystematycznych występujących w tych samych dniach i godzinach, ale w kwietniu i październiku. Z pracy Szmuksta-Zawadzkiej i Zawadzkiego (2015, tabl. 1, s. 151) wynika, że różnica wskaźników sezonowości w przypadku tych miesięcy wynosiła 4,7 p.proc., a w przypadku sierpnia i października – 12,5 p.proc., czyli była o ok. 1,7 raza większa⁴. Pozwala to zbadać wpływ dużego wzrostu natężenia sezonowości na dokładność prognoz inter- i ekstrapolacyjnych. Natężenie sezonowości jest rozumiane jako różnica nasilenia lub skali wahań sezonowych. Zjawiskiem o większym nasileniu jest to, dla którego rozstęp wskaźników sezonowości jest większy.

⁴ W tej samej tablicy, w takim samym układzie, zestawiono charakterystyki najlepszych modeli oraz błędy prognoz w wariancie występowania luk w kwietniu, a nie w sierpniu. Modele zapisano w tej samej konwencji co modele omawiane w niniejszym artykule.

Z porównania modeli hybrydowych wynika, że w przypadku obu kryteriów otrzymano modele jednostopniowe, różniące się jednak strukturą. Przy zastosowaniu kryterium $MAPE_I$ za pomocą modeli hierarchicznych $WI_GDM_HG122_0/1$ i $PLI_GDM_HG122_0/1$ opisywano wahania o cyklu dobowym o podzielnikach 12 i 2. Różniły się one postacią analityczną – w przypadku modelu szeregu czasowego była to postać wykładnicza, a w przypadku modelu regresyjnego – postać liniowa. Przy zastosowaniu kryterium $MAPE_E$ za pomocą wykładniczych modeli hierarchicznych $WE_GDM_0/1_HM43$ i $PWE_GDM_0/1_HM43$ opisywano wahania o cyklu 12-miesięcznym.

Z porównania własności predyktywnych hybrydowych równań szeregu czasowego i równań regresyjnych wynika, że lepszymi własnościami charakteryzują się modele regresyjne. W przypadku kryteriów $MAPE_I$ i $MAPE_E$ współczynnik determinacji przyjął wartości większe odpowiednio o 4,47 i 8,11 p.proc. Modele regresyjne charakteryzowały się także mniejszymi wartościami odchylenia standardowego składników losowych i współczynnika zmienności losowej. Różnice te wynosiły odpowiednio 3,83 MW i 7,15 MW oraz 0,976 p.proc. i 1,273 p.proc. Podane liczby jednoznacznie wskazują na lepsze własności predyktywne modeli regresyjnych. Przy zastosowaniu obu kryteriów otrzymano także niższe oceny błędów prognoz. W przypadku prognoz interpolacyjnych różnica wyniosła 1,132 p.proc. (17,2%), a w przypadku prognoz ekstrapolacyjnych – 2,722 p.proc. (29,4%).

W wykładniczych modelach klasycznych zarówno własności predyktywne, jak i błędy prognoz przyjęły korzystniejsze wartości także dla predyktora regresyjnego. Współczynnik determinacji w modelach hybrydowych modeli regresyjnych osiągnął wartości nieznacznie mniejsze niż w modelu klasycznym; dla kryteriów $MAPE_I$ i $MAPE_E$ różnice te wyniosły odpowiednio 0,99 p.proc. i 1,89 p.proc., a w przypadku odchylenia standardowego składników losowych i współczynnika zmienności losowej – odpowiednio 0,48 MW i 1,78 MW oraz 0,075 p.proc. i 0,430 p.proc. Błąd prognoz interpolacyjnych dla regresyjnego modelu hybrydowego wyniósł 5,452% i był większy o 0,046 p.proc. od błędu dla modelu klasycznego.

Z porównania własności predyktywnych i dokładności prognoz dotyczących modeli różniących się natężeniem sezonowości w przypadku miesięcy, w których wystąpiły luki w danych, wynika, że wartości współczynnika determinacji, parametrów struktury stochastycznej i przeciętnego błędu względnego prognoz są zbliżone. Istotny okazał się natomiast wpływ różnicy ocen wskaźników sezonowości na dokładność prognoz ekstrapolacyjnych. W przypadku modelu o mniejszej różnicy wartości tych wskaźników otrzymano błędy mniejsze o ok. 1,7 p.proc. (35%), a w przypadku modeli klasycznych różnica ta wyniosła ok. 1,4 p.proc. (26%).

5. Porównanie wyników modelowania i prognozowania dla wariantów luk G i GDM

Porównanie dokładności opisu modeli i błędów prognoz otrzymanych dla najlepszych modeli z lukami występującymi w okresach godzinnych (G) i rozpatrywanymi wcześniej modelami z lukami występującymi w każdym z trzech rodzajów wahań (GDM) pozwala na dokonanie oceny wpływu zwiększenia liczby luk – w omawianym badaniu o ok. 80% – na dokładność opisu oraz prognoz.

Analizę przeprowadzono osobno dla modeli szeregu czasowego i modeli regresyjnych; uwzględniono także modele klasyczne. Ze względu na porównywanie własności predykcyjnych i prognoz otrzymanych dla modeli hybrydowych i modeli klasycznych pominięto (podobnie jak w tabl. 5) oceny błędów $MAPE_{I,E}$ oraz $MAPE_{E,I}$. Informacje dotyczące wariantu luk G zostały zaczerpnięte z pracy Zawadzkiego (2018).

Tabl. 6. Porównanie modeli szeregu czasowego w wariantach luk G i GDM przy zastosowaniu kryteriów $MAPE_I$ oraz $MAPE_E$

Modele	R^2	S_e w MW	V_{se}	$MAPE_I$	$MAPE_E$
			w %		
WI_G_HG432_0/1	0,7661	36,04	9,785	6,391	.
WI_GDM_HG122_0/1	0,8106	32,49	8,823	6,584	.
WE_G_HG122_0/1	0,8106	32,44	8,808	.	8,775
WE_GDM_0/1_HM43	0,7530	37,11	10,075	.	9,257
W_G_0/1_0/1	0,8207	31,57	8,572	6,194	8,554
W_GDM_0/1_0/1	0,8227	31,45	8,540	6,308	9,129

Uwaga. Jak przy tabl. 1.

Źródło: opracowanie własne.

Przy zastosowaniu kryterium $MAPE_I$ model hybrydowy z mniejszą liczbą luk (WI_G_HG432_0/1) otrzymał niższą o ponad 4,45 p.proc. ocenę R^2 oraz wyższe oceny S_e i V_{se} , odpowiednio o 3,55 MW i 0,962 p.proc., ale niższą o ok. 0,2 p.proc. ocenę błędu prognoz interpolacyjnych niż model WI_GDM_HG122_0/1. Z kolei przy zastosowaniu kryterium $MAPE_E$ model WE_G_HG122_0/1 charakteryzował się zarówno korzystniejszymi własnościami predykcyjnymi, jak i niższymi wartościami błędu prognoz. Wartość współczynnika R^2 była o 5,76 p.proc. mniejsza od otrzymanej dla modelu WE_GDM_0/1_HM43, a wartości współczynników S_e i V_{se} były mniejsze odpowiednio o 4,67 MW i 1,266 p.proc. Dla modelu o mniejszej liczbie luk otrzymano także niższą – o 0,482 p.proc. (5,5%) – ocenę błędu prognoz ekstrapolacyjnych. Różnicę tę w przypadku liczby luk większej o ponad 80% należy uznać za niewielką. W prognozach interpolacyjnych wynosiła ona 1,03%.

Z powyższych rozważań można wyprowadzić wniosek, że dla celów prognozowania pod uwagę należy brać dwa kryteria wyboru równań: $MAPE_I$ i $MAPE_E$.

Z informacji zawartych w tabl. 6 wynika ponadto, że modele klasyczne, podobnie jak wszystkie modele hybrydowe, są modelami wykładniczymi. Nieznacznie lepszymi własnościami predykcyjnymi charakteryzował się model $W_GDM_0/1_0/1$, ale niższe oceny błędów obu rodzajów prognoz otrzymano dla modelu $W_G_0/1_0/1$. W przypadku prognoz interpolacyjnych błąd był mniejszy o 1,08 p.proc., a w przypadku prognoz ekstrapolacyjnych – o 6,72 p.proc. Oznacza to, że podobnie jak w modelach hybrydowych, wzrost liczby luk w prognozach ekstrapolacyjnych był nieco silniejszy. Oceny błędów prognoz interpolacyjnych otrzymane dla modeli klasycznych były niższe nie więcej niż o 0,40 p.proc., a prognoz ekstrapolacyjnych – co najwyżej o 0,70 p.proc.

W tabl. 7 zestawiono parametry charakteryzujące własności predyktywne i błędy prognoz najlepszych modeli regresyjnych z lukami występującymi w okresach godzinnych (G) i modeli z lukami w każdym z trzech rodzajów wahań (GDM).

Tabl. 7. Porównanie hybrydowych modeli regresyjnych dla wariantów luk G i GDM przy zastosowaniu kryteriów $MAPE_I$ i $MAPE_E$

Modele	R^2	S_e w MW	V_{Se}	$MAPE_I$	$MAPE_E$
			w %		
PLI_G_HG432_0/1	0,8344	29,63	8,107	5,406	.
PLI_GDM_HG122_0/1	0,8454	28,66	7,847	5,452	.
PWE_G_HG432_0/1	0,8379	29,75	8,145	.	6,622
PWE_GDM_0/1_HM43	0,8364	29,96	8,202	.	6,535
PL_G_0/1_0/1	0,8485	28,36	7,76	5,374	6,906
PW_GDM_0/1_0/1	0,8553	28,18	7,716	5,406	6,983

Uwaga. Jak przy tabl. 1.

Źródło: opracowanie własne.

Z informacji zawartych w tabl. 7 wynika, że przy zastosowaniu kryterium $MAPE_I$ najlepszymi modelami regresyjnymi okazały się jednostopniowe hybrydowe modele liniowe z różnymi podzielnikami dla 24-godzinnego cyklu wahań. W wariancie G był to hierarchiczny model trójstopniowy z podzielnikami wynoszącymi odpowiednio 4, 3 i 2. W wariancie ze znacznie większą liczbą luk (GDM) był to model dwustopniowy o podzielnikach 12 i 2. Przy zastosowaniu kryterium $MAPE_E$ oba predyktory przyjęły postać wykładniczą, przy czym w przypadku mniejszej liczby luk był to model o tych samych podzielnikach co przy zastosowaniu kryterium $MAPE_I$, ale

w postaci wykładniczej, a w wariancie GDM – model dwustopniowy z podzielnikami cyklu rocznego wynoszącymi 4 i 3.

W modelach klasycznych otrzymano predyktory różniące się postacią analityczną – w wariancie G lepsza okazała się postać liniowa, a w wariancie GDM – wykładnicza.

Porównanie dokładności opisu i błędów prognoz najlepszych modeli regresyjnych, zarówno hybrydowych, jak i klasycznych, w obu wariantach luk pokazuje, że są one bardzo zbliżone. Różnica wartości R^2 między wartością maksymalną, otrzymaną dla modelu klasycznego PW_GDM_0/1_0/1, a wartością minimalną, otrzymaną dla hybrydowego modelu liniowego PLI_G_HG432_0/1, wyniosła 2,1 p.proc, a w przypadku modeli hybrydowych przy zastosowaniu kryterium $MAPE_I$ – 1,1 p.proc.

Dla modeli klasycznych różnica ocen R^2 wynosi 0,68 p.proc., przy czym większą wartością charakteryzuje się wariant GDM. Podobnie wygląda sytuacja w przypadku modeli hybrydowych, przy zastosowaniu kryterium $MAPE_I$. Relacje te potwierdzają niższe oceny S_e i V_{Se} .

Przy zastosowaniu kryterium $MAPE_E$ różnice ocen parametrów struktury stochastycznej modeli hybrydowych w wariantach G i GDM są niewielkie. W przypadku R^2 niższą – o 0,15 p.proc. – ocenę otrzymano dla wariantu G. Ocena S_e dla modelu z mniejszą liczbą luk jest niższa o 0,21 MW, a ocena V_{Se} – o 0,57 p.proc.

Różnica ocen parametrów R^2 , S_e i V_{Se} między lepszym z modeli klasycznych (PW_GDM_0/1_0/1) a modelem PWE_G_HG432_0/1 wynosi odpowiednio 1,06 p.proc., 1,39 MW oraz 0,429 p.proc.

Prawidłowości dotyczące zbliżonych własności predyktywnych znajdują pełne odzwierciedlenie w dokładności prognoz. Przy zastosowaniu kryterium $MAPE_I$ model o mniejszej liczbie luk uzyskał ocenę niższą o 0,163 p.proc.

Bardzo zbliżone są także oceny błędów $MAPE_E$ otrzymanych na podstawie najlepszych modeli regresyjnych w wariantach G (PWE_G_HG432_0/1) i GDM (PWE_GDM_0/1_HM43). W wariancie GDM (z liczbą luk większą o ok. 80%) błąd ten był nawet o 0,087 p.proc. (1,33%) mniejszy niż w wariancie G.

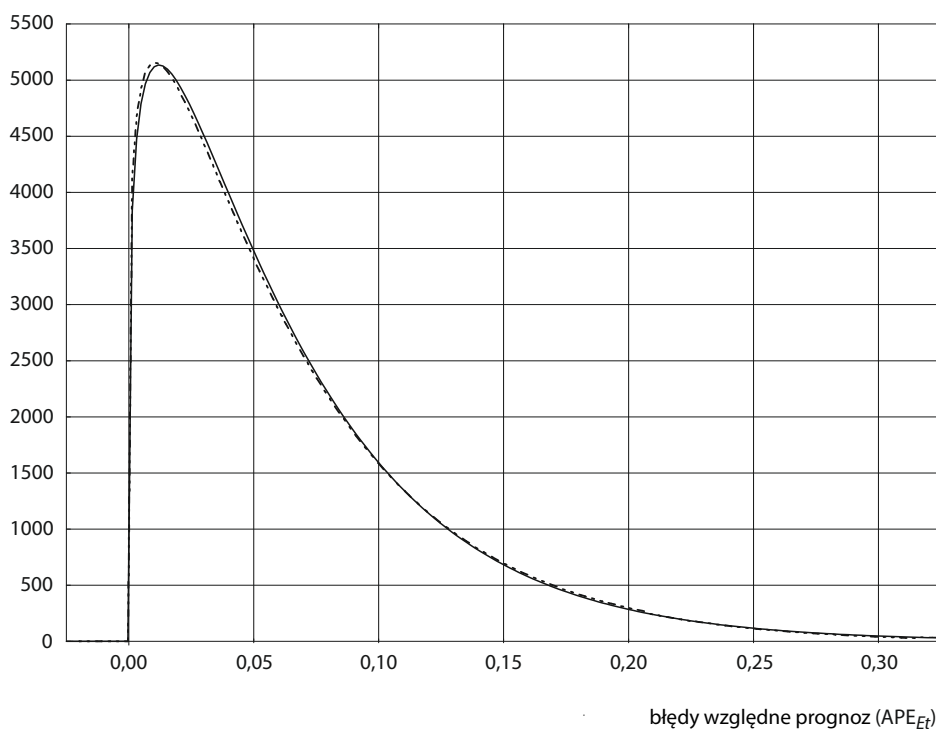
Do zbadania identyczności dystrybucji obu rozkładów empirycznych modułów względnych błędów APE_{Et} wykorzystano, podobnie jak poprzednio, test λ Kolmogorowa-Smirnowa, przy przyjęciu tego samego poziomu istotności ($\alpha = 0,05$). Statystyka empiryczna λ tego testu, dla 8760 obserwacji w każdym modelu, wynosi $\lambda = D_{nm}\sqrt{n} = 0,006164 \cdot 66,182 = 0,4080$. Z tablicy rozkładu granicznego λ odczytujemy wartość krytyczną $\lambda_{0,05} = 1,354$. Ponieważ $\lambda = 0,4080 < 1,354 =$

$= \lambda_{0,05}$, to nie ma podstaw do odrzucenia hipotezy zerowej, w której założono identyczność badanych rozkładów.

Wykres 3 przedstawia aproksymanty teoretyczne rozkładów empirycznych modułów APE_{Et} . Najlepiej były one aproksymowane, podobnie jak poprzednio, za pomocą rozkładu gamma.

Wykr. 3. Aproksymanty teoretyczne rozkładu empirycznego modułów APE_{Et} otrzymanych na podstawie modeli PWE_GDM_0/1_HM43 i PWE_G_HG432_0/1

liczba obserwacji



Źródło: opracowanie własne.

Z porównania rozkładów błędów prognoz ekstrapolacyjnych w wariantach G i GDM wynika, że są one bardzo zbliżone.

Ogólny wniosek z powyższych porównań jest taki, że modele w wariancie GDM, o liczbie luk większej o 80% niż w wariancie G, charakteryzowały się zarówno bardzo zbliżonymi własnościami predyktywnymi, jak i ocenami błędów prognoz inter- i ekstrapolacyjnych.

6. Wnioski

Z analiz przeprowadzonych na omawianym szeregu czasowym można wyprowadzić kilka istotnych wniosków szczegółowych.

W przypadku podwójnych modeli hybrydowych szeregu czasowego i modeli regresyjnych podstawą wyboru powinny być kryteria dotyczące minimalizacji błędów prognoz interpolacyjnych ($MAPE_I$) i ekstrapolacyjnych ($MAPE_E$).

Regresyjne modele hybrydowe charakteryzowały się lepszymi własnościami prognostycznymi i niższymi ocenami błędów prognoz inter- i ekstrapolacyjnych niż modele szeregu czasowego, przy czym różnice w dokładności były większe w przypadku prognoz ekstrapolacyjnych.

We wszystkich przypadkach najlepszymi modelami hybrydowymi okazały się pojedyncze modele hierarchiczne opisujące wahania o cyklu 24-godzinnym. Otrzymano dla nich mniejsze błędy prognoz niż dla modeli pojedynczych opisujących wahania o cyklu 12-miesięcznym i dla modeli podwójnych.

Własności predyktywne i błędy prognoz interpolacyjnych otrzymanych na podstawie obu klas modeli oraz prognoz ekstrapolacyjnych w wariancie, w którym luki występowały tylko w ośmiu okresach godzinnych (G), były tylko nieznacznie mniejsze niż w wariancie z lukami w ośmiu okresach godzinnych, dwóch dniach tygodnia i dwóch miesiącach (GDM), mimo że w drugim wariancie liczba luk była większa o ponad 80%.

Ocena przeciętnego względnego błędu prognoz ekstrapolacyjnych otrzymana dla modelu regresyjnego w wariancie GDM była o ponad 5 p.proc. niższa od otrzymanej dla modelu klasycznego w wariancie G. Oceny błędów prognoz ekstrapolacyjnych otrzymane dla modeli regresyjnych w obu wariantach luk były zbliżone i jednocześnie o kilka punktów procentowych niższe od oceny błędów otrzymanych dla modeli klasycznych. Niższe oceny błędów prognoz ekstrapolacyjnych przemawiają za wykorzystaniem modeli hybrydowych w prognozowaniu zmiennych z potrójnie złożonymi wahaniami sezonowymi.

Przeprowadzone badanie prowadzi do konkluzji, że modele hybrydowe szeregu czasowego i hybrydowe modele regresyjne mogą być użytecznym narzędziem modelowania i prognozowania w warunkach braku pełnej informacji. Mimo że wariant GDM charakteryzował się większą aż o 80% liczbą luk niż wariant G, to zarówno własności predyktywne, jak i oceny błędów prognoz inter- i ekstrapolacyjnych były bardzo zbliżone.

Bibliografia

- Allison, P. D. (2001). *Missing Data*. Thousand Oaks: Sage Publications.
- Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., Petropoulos, F. (2017). Forecasting with Temporal Hierarchies. *European Journal of Operational Research*, 262(1), 60–74. DOI: 10.1016/j.ejor.2017.02.046.
- Domański, Cz. (1990). *Testy statystyczne*. Warszawa: Państwowe Wydawnictwo Ekonomiczne.

- Henn, B., Raleigh, M. S., Fisher, A., Lundquist, J. D. (2013). A Comparison of Methods for Filling Gaps in Hourly Near-Surface Air Temperature Data. *Journal of Hydrometeorology*, 14(3), 929–945. DOI: 10.1175/JHM-D-12-027.1.
- Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis*, 55(9), 2579–2589. DOI: 10.1016/j.csda.2011.03.006.
- Kufel, T. (2010). *Ekonometryczna analiza cykliczności procesów gospodarczych o wysokiej częstotliwości obserwowania*. Toruń: Wydawnictwo Naukowe UMK.
- Kwac, J., Flora, J., Rajagopal, R. (2014). Household Energy Consumption Segmentation Using Hourly Data. *IEEE Transactions on Smart Grid*, 5(1), 420–430. DOI: 10.1109/TSG.2013.2278477.
- Little, R. J. A., Rubin, D. B. (2002). *Statistical Analysis with Missing Data: Second Edition*. Hoboken, New Jersey: John Wiley & Sons. DOI: 10.1002/9781119013563.
- Raudenbush, S. W., Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, London, New Delhi: Sage Publication Inc. Pobrane z: https://books.google.pl/books?id=uyCV0CNGDLQC&pg=PR3&hl=pl&source=gbs_selected_pages&cad=2#v=onepage&q&f=false.
- Szmuksta-Zawadzka, M., Zawadzki, J. (2002). Hierarchiczne modele szeregów czasowych z wahaniami sezonowymi. Budowa. Estymacja. Prognozowanie. W: A. Zeliaś (red.), *Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych* (s. 193–204). Kraków: Wydawnictwo AE.
- Szmuksta-Zawadzka, M., Zawadzki, J. (2011). Modelowanie i prognozowanie zapotrzebowania na energię elektryczną w mikroskali. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (165), 152–161.
- Szmuksta-Zawadzka, M., Zawadzki, J. (2014a). Modele hierarchiczne w prognozowaniu zmiennych o wysokiej częstotliwości obserwowania w warunkach braku pełnej informacji. *Ekonometria. Econometrics*, 46(4), 72–84. DOI: 10.15611/ekt.2014.4.07.
- Szmuksta-Zawadzka, M., Zawadzki, J. (2014b). Wykorzystanie modeli „hybrydowych” w prognozowaniu brakujących danych w szeregach ze złożoną okresowością (sezonowością). W: W. Jurek (red.), *Matematyka i informatyka na usługach ekonomii: rozważania ogólne, modele* (s. 72–83). Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Szmuksta-Zawadzka, M., Zawadzki, J. (2015). Wykorzystanie danych oczyszczonych o wysokiej częstotliwości w prognozowaniu zmiennych ze złożoną sezonowością. *Metody Ilościowe w Badaniach Ekonomicznych*, 16(4), 147–159. Pobrane z: http://qme.sggw.pl/pdf/MIBE_T16_z4.pdf.
- Taieb, S. B., Yu, J., Barreto, M. N., Rajagopal, R. (2017). *Regularization in Hierarchical Time Series Forecasting with Application to Electricity Smart Meter Data*. Pobrane z: <https://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14173/14055>.
- Weron, R. (2006). *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Chichester: John Wiley & Sons.
- Witkowska, D., Górecka, A., Szadkowska, D., Szymczak, Z. (2000). The forecasting of the demand for electric energy: comparative analysis. *Dynamic Econometric Models*, 4, 45–59.
- Zawadzki, J. (2018). Modele hybrydowe w prognozowaniu brakujących danych w szeregach o bardzo wysokiej częstotliwości obserwowania. *Folia Pomeranae Universitatis Technologiae Stettinensis, Oeconomia*, 346(92), 81–96. Pobrane z: <http://foliaoe.zut.edu.pl/pdf/346/oe2018.92.3.07.pdf>.

Niezdolność do odpowiedniego ogrzania mieszkania w gospodarstwach domowych w Unii Europejskiej i krajach stowarzyszonych

Kornelia Przedworska^a

Streszczenie. Celem badania przedstawionego w artykule jest określenie poziomu ubóstwa energetycznego gospodarstw domowych w Unii Europejskiej (UE) i krajach stowarzyszonych, ocenianego na podstawie niezdolności do odpowiedniego ogrzania mieszkania. Zastosowano metodę systematycznego przeglądu literatury. Analiza, przeprowadzona na podstawie danych Eurostatu i Głównego Urzędu Statystycznego, objęła wyniki badania EU-SILC z lat 2009–2018. Wynika z niej, że średnia liczba gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania w Europie maleje. Jednak w niektórych krajach, takich jak Macedonia, Litwa czy Bułgaria, liczba gospodarstw domowych doświadczających tego problemu nadal pozostaje duża. W Polsce poziom omawianego zjawiska jest niższy niż średni poziom w UE. Aby pogłębić analizę dotyczącą Polski, dokonano oceny korelacji wskaźnika ubóstwa energetycznego z wysokością przeciętnego dochodu *per capita* w gospodarstwie domowym oraz z poziomem ubóstwa ekonomicznego. W tym celu zastosowano współczynnik korelacji *r*-Pearsona.

Słowa kluczowe: ubóstwo energetyczne, ubóstwo ekonomiczne, dochody gospodarstw domowych, EU-SILC

JEL: C10, Q56, O18

Inability to adequately heat a dwelling in the European Union and associated countries' households

Abstract. The purpose of the study presented in the article is to determine the level of energy poverty in households of the European Union (EU) and associated countries, pronounced on the basis of the household's inability to supply its dwellings with adequate heat levels. The method of systematic literature review was applied. An analysis based on Statistics Poland and Eurostat data included the results of the EU-SILC research conducted in the years 2009–2018. The analysis indicated that the average number of households unable to sufficiently heat their homes is decreasing in Europe. However, in some countries, such as Macedonia, Lithuania or Bulgaria, the number of households experiencing energy poverty is still high. In Poland, the scale of this problem is smaller than the EU average. In order to conduct an in-depth analysis for Poland, an assessment was made of the correlation between the energy poverty rate and the average income per person in a household and the level of economic poverty. For this purpose the *r*-Pearson correlation coefficient was applied.

Keywords: energy poverty, economic poverty, household income, EU-SILC

^a Uniwersytet Ekonomiczny w Katowicach, Kolegium Zarządzania.
ORCID: <https://orcid.org/0000-0003-2714-4066>.

1. Wprowadzenie

W dobie zmian klimatycznych kwestia nośników energii wykorzystywanych przez gospodarstwa domowe do ogrzewania mieszkań stała się kluczowa dla kształtowania polityki w zakresie ekologii. Transformacja energetyczna Polski, wdrażanie zasad Zielonego Ładu i zmieniające się w związku z tym przepisy¹ prowadzą do ograniczenia możliwych sposobów ogrzewania mieszkań oraz rodzajów surowców, które mogą być źródłem ciepła. Wszelkie decyzje podejmowane przez rządy czy samorządy lokalne w tym zakresie muszą być poprzedzone analizą sposobów, za pomocą których gospodarstwa domowe radzą sobie z ogrzewaniem mieszkań. Istnieje bowiem wiele gospodarstw domowych, w przypadku których poziom zaspokojenia potrzeby odpowiedniego ogrzania mieszkania jest za niski. Występuje wówczas dyskomfort cieplny, będący jedną ze zmiennych opisujących zjawisko ubóstwa energetycznego.

Nieodpowiednie ogrzewanie mieszkań to problem wieloaspektowy. Długotrwałe przebywanie w warunkach niskiej temperatury oddziałuje na zdrowie zarówno fizyczne (schorzenia układu oddechowego, alergie, zaburzenia hormonalne i związane z układem krążenia), jak i psychiczne (stres, niepokój i obniżenie nastroju) (Boguszewski i Herudziński, 2018, s. 4). Z kolei nieumiejętne wykorzystywanie i nieefektywność źródeł ciepła oraz zły stan techniczny budynku wpływają negatywnie na jakość powietrza i stan środowiska.

Celem badania przedstawionego w artykule jest określenie poziomu ubóstwa energetycznego gospodarstw domowych w Unii Europejskiej (UE) i krajach stowarzyszonych, ocenianego na podstawie niezdolności do odpowiedniego ogrzania mieszkania (ang. *inability to keep home adequately warm*). O wadze tego problemu może świadczyć to, że w Europie doświadcza go ponad 50 mln gospodarstw domowych (UE Energy Poverty Observatory [EPOV], 2019). Porównanie poziomu dyskomfortu cieplnego w Europie i Polsce jest istotne chociażby ze względu na istniejące różnice w wykorzystywaniu nośników energii. Dla przykładu 84% węgla kamiennego zużywanego przez gospodarstwa domowe w UE przypada na Polskę (Główny Urząd Statystyczny [GUS], 2019c, s. 188). W związku z koniecznością ograniczania emisji CO₂ podejście do tego źródła energii będzie się zmieniać.

2. Ubóstwo energetyczne a niezdolność do odpowiedniego ogrzania mieszkania

Przez ubóstwo energetyczne (ang. *energy poverty*) należy rozumieć niezdolność do odpowiedniego ogrzania mieszkania (Eurostat, 2019a) oraz trudności z opłacaniem rachunków za prąd i gaz, które wpływają na dostęp do źródeł energii (Eurostat,

¹ Na przykład: Ustawa z dnia 25 sierpnia 2006 r. o systemie monitorowania i kontrolowania jakości paliw (Dz.U. 2006 nr 169 poz. 1200), uchwały antysmogowe w Polsce czy rozporządzenie Komisji UE określające parametry dla kotłów na paliwo stałe (Rozporządzenie Komisji UE 2015/1189 z dnia 28 kwietnia 2015 r. w sprawie wykonania dyrektywy Parlamentu Europejskiego i Rady 2009/125/WE).

2019b). Badania nad tym zjawiskiem są prowadzone na świecie od lat 80. XX w., ale nie wypracowano do tej pory jego powszechnie uznawanej definicji. Instytut Badań Strukturalnych (IBS), wszechstronnie zajmujący się tym problemem w Polsce, definiuje ubóstwo energetyczne jako zjawisko polegające na doświadczaniu trudności w zaspokojeniu podstawowych potrzeb związanych z energią cieplną i elektryczną w mieszkaniu (Lis i in., 2016, s. 3). Według EPOV (2019) ubóstwo energetyczne to stan, w którym poziom zaspokojenia podstawowych potrzeb gospodarstwa domowego związanych z usługami energetycznymi, takimi jak: ogrzewanie, chłodzenie, oświetlenie i dostarczenie energii do zasilania urządzeń, jest niewystarczający. Przewodniczący komitetu sterującego EPOV zwraca uwagę, że przez długi czas rządzący nie potrafili dostrzec związku pomiędzy biedą a energią. Dopiero gdy w brytyjskim środowisku akademickim pojawiła się definicja ubóstwa energetycznego, zainicjowano debatę naukową na ten temat, ale przede wszystkim dano początek nowej polityce państwa, która miała na celu zdiagnozowanie tego zjawiska i przeciwdziałanie jego skutkom (Bouzarovski, 2018, s. 9–10).

Ubóstwo energetyczne można opisywać za pomocą różnych wskaźników. W Wielkiej Brytanii, zgodnie z pierwszą oficjalną definicją z 1991 r., stosowaną również w innych krajach, której autorką jest Brendy Boardman, przyjmuje się, że „gospodarstwo domowe jest ubogie w paliwo, jeśli musi wydać ponad 10% swoich dochodów na paliwo, aby utrzymać odpowiedni poziom ciepła” (Szamrej-Baran i Baran, 2014, s. 333). W tym podejściu bierze się pod uwagę jedynie dochód i wydatki gospodarstwa domowego na energię. Należy dodać, że według Światowej Organizacji Zdrowia (WHO) odpowiedni standardowy poziom ciepła wynosi 21°C dla głównych pomieszczeń typu salon i 18°C dla pozostałych (WHO, 2007, s. 4). Definicja ta uległa zmianie, gdy uznano, że ubóstwo energetyczne jest kombinacją wzajemnie oddziałujących na siebie czynników, takich jak: niskie dochody, nieefektywne materiały budowlane, nieefektywny system grzewczy i niedostateczny dostęp do źródeł energii. Dlatego lepszym miernikiem jest LIHC (ang. *low income high cost*) (Teli i in., 2016, s. 177), który uwzględnia dochód gospodarstwa domowego i koszty utrzymania domu, a także wprowadza pojęcia górnego limitu dochodów, dochodu ekwiwalentnego czy rachunków ekwiwalentnych (Szamrej-Baran i Baran, 2014, s. 334).

Z kolei AFCP (ang. *after fuel cost poverty*) polega na porównaniu ekwiwalentnego dochodu danego gospodarstwa domowego ze standardowym progiem 60% ekwiwalentnego dochodu narodowego. Dochód ten jest rozpatrywany po odjęciu kosztów mieszkaniowych i krajowych kosztów paliwa (netto). Zgodnie z tym podejściem gospodarstwa domowe, których ekwiwalentny dochód netto kształtuje się poniżej progu 60% ekwiwalentnej krajowej mediany dochodu netto, klasyfikuje się jako ubogie energetycznie (Charlier i Legendre, 2019, s. 30–31). Główną zaletą AFCP jest to, że uwzględnia koszty mieszkaniowe. Z drugiej strony na jego podstawie gospodarstwa domowe o bardzo niskich dochodach będą klasyfikowane jako ubogie w paliwo, niezależnie od zapotrzebowania na nie.

Główny Urząd Statystyczny definiuje ubóstwo energetyczne jako sytuację, w której gospodarstwo domowe nie jest w stanie zapewnić sobie wystarczającego poziomu ciepła, chłodu, oświetlenia i energii do zasilania urządzeń w wyniku – łącznie – niskich dochodów, wysokich wydatków energetycznych i niskiej efektywności energetycznej budynku. W swoich badaniach GUS wykorzystuje kilka wskaźników do pomiaru poziomu ubóstwa energetycznego. Są to (GUS, 2019c, s. 73):

- LIHC – wysokie wymagane koszty energii (tj. powyżej mediany poziomu krajowego) i niskie dochody (tj. rozporządzalny dochód poniżej oficjalnie określonego progu ubóstwa);
- podwójna mediana wydatków na energię – udział faktycznych wydatków energetycznych w dochodzie kształtuje się na poziomie wyższym niż podwojona mediana tej wartości w populacji;
- zdolność do terminowego opłacania rachunków – zaległości w opłacaniu rachunków za energię lub niezdolność do ich opłacania;
- problemy ze stanem budynku – budynek z przeciekającym dachem, zawilgocnymi ścianami, podłogami, fundamentami, butwiejącymi oknami lub podłogami;
- niewystarczający komfort cieplny – zadeklarowana niezdolność do wystarczającego ogrzania domu/mieszkania.

Mierniki stosowane przez GUS wskazują pośrednio na przyczyny ubóstwa energetycznego, które można podzielić na: techniczne, ekonomiczne i związane z postawą wobec efektywnego wykorzystania energii. Przyczyny techniczne występują wtedy, gdy mieszkanie charakteryzuje się niskim poziomem efektywności energetycznej lub wadliwie działają w nim instalacje grzewcze. Przyczyny ekonomiczne wiążą się z deprawacją materialną, prowadzącą do zaległości w opłatach za ogrzewanie. Z kolei niewłaściwy sposób użytkowania posiadanych urządzeń grzewczych skutkuje tym, że wydatki na ogrzewanie są wyższe, niż wynikałoby ze specyfikacji instalacji (Boguszewski i Herudziński, 2018, s. 5).

Występowanie ubóstwa energetycznego uniemożliwia uczestnictwo w zwykłych aktywnościach i życie na poziomie przeciętnym dla danej społeczności. Przebywanie w odpowiednio ogrzonym mieszkaniu jest naturalną potrzebą organizmu, której zaspokojenie – w przypadku braku wystarczających środków finansowych – może się odbywać kosztem zaspokojenia innych potrzeb (Szamrej-Baran, 2017, s. 284). Ubóstwo energetyczne wpływa negatywnie na stan zdrowia fizycznego i psychicznego członków gospodarstwa domowego dotkniętego tym problemem. Z powodu braku zasobów materialnych takie gospodarstwo nie inwestuje w utrzymanie budynku, który stopniowo ulega degradacji. Dodatkowo niewłaściwa izolacja okien, ścian i drzwi przyczynia się do wzrostu strat ciepła i do zawilgocenia, a więc pogorszenia standardu energetycznego budynku. To z kolei powoduje jeszcze większe zużycie nośników energii i wzrost emisji CO₂ (Szamrej-Baran i Baran, 2014, s. 333).

W celu efektywnego ograniczenia problemu ubóstwa energetycznego IBS postuluje koncentrację na trzech priorytetach:

- umożliwieniu łączenia programów termomodernizacji z instrumentami pomocy społecznej;
- umożliwieniu łączenia instrumentów pomocy społecznej z inwestycjami sektora prywatnego;
- trafniejszym adresowaniu instrumentów osłonowych (Lis i in., 2016, s. 11).

3. Metoda badania

Przedmiotem badania jest niezdolność gospodarstw domowych do odpowiedniego ogrzania mieszkania. Zgodnie z art. 3 ust. 5 Ustawy z dnia 2 grudnia 1999 r. o narodowym spisie powszechnym ludności i mieszkań w 2001 r. (Dz.U. 2000 nr 1 poz. 1) gospodarstwo domowe to zespół osób mieszkających i utrzymujących się wspólnie lub jedna osoba utrzymująca się samodzielnie. Definicję tę stosuje GUS. Istotne jest tu finansowanie i decydowanie o wydatkach. Osoby mieszkające wspólnie, które utrzymują się samodzielnie, tworzą osobne gospodarstwa domowe. W Polsce w 2018 r. było 14,44 mln gospodarstw domowych. Jedno gospodarstwo składało się średnio z 2,6 osoby (GUS, 2019c, s. 21).

Analizę przeprowadzono na podstawie danych dotyczących niezdolności do odpowiedniego ogrzania mieszkania, które zostały zaczerpnięte z Europejskiego badania warunków życia ludności (ang. The European Union Statistics on Income and Living Conditions, EU-SILC) oraz publikacji GUS. Dotyczyła ona lat 2009–2018. Badanie EU-SILC jest realizowane we wszystkich krajach członkowskich UE i krajach stowarzyszonych (w Polsce od 2005 r. przez GUS – zob. GUS, 2014a). Obejmuje wszystkie gospodarstwa domowe (ang. *private households*) i ich aktualnych członków zamieszkujących na terytorium określonego kraju w czasie zbierania danych (Eurostat, 2019c).

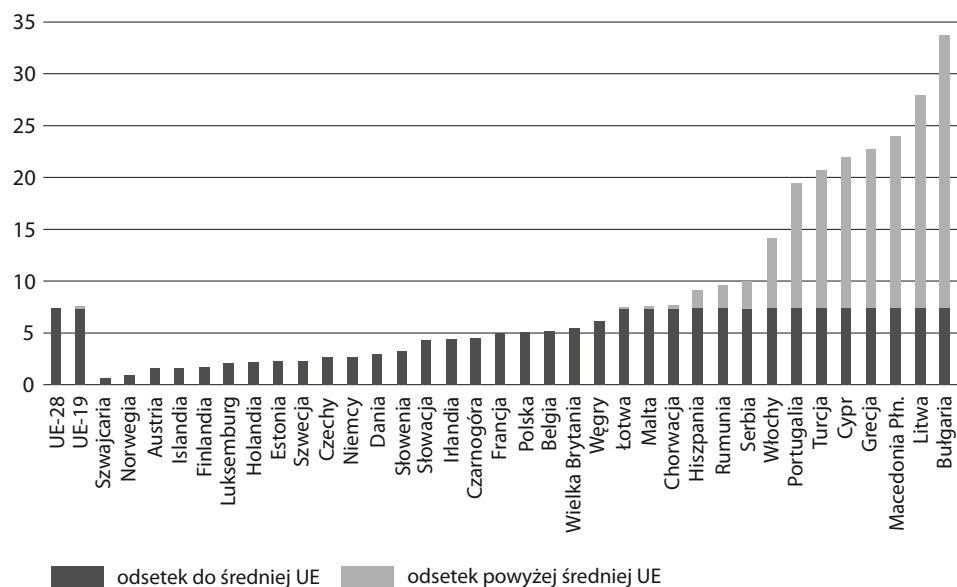
W badaniu omawianym w artykule przyjęto metodę systematycznego przeglądu literatury (Czakon, 2016, s. 119–126). Ustalono słowa kluczowe dotyczące badanego zjawiska oraz, stosując technikę kuli śnieżnej, zebrano materiały niezbędne do przeprowadzenia analiz. W celu głębszej analizy zjawiska w Polsce dokonano korelacji wskaźnika ubóstwa energetycznego z wysokością przeciętnego dochodu *per capita* w gospodarstwie domowym oraz z poziomem ubóstwa relatywnego przy zastosowaniu współczynnika korelacji liniowej *r*-Pearsona.

Ograniczeniem było prezentowanie przez Eurostat danych dotyczących wyłącznie wybranych typów gospodarstw domowych, czyli gospodarstw domowych ogółem oraz gospodarstw jednoosobowych tworzonych przez jedną osobę powyżej 65. roku życia lub gospodarstw tworzonych przez jedną osobę z dzieckiem/dziećmi na utrzymaniu. Na potrzeby badania poziomu ubóstwa energetycznego w UE w analizie wykorzystano dane za 10 kolejnych lat. W przypadku braku danych w niektórych latach przyjęto dane z roku poprzedniego, jeśli analiza danych wskazywała na duże prawdopodobieństwo uzyskania zbliżonego wyniku w danym roku.

4. Odsetek gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania

W 2018 r. w UE, do której należało 28 krajów, odsetek gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania wynosił 7,4% (wykr. 1). W strefie euro, którą tworzy 19 krajów, był on o 0,2 p.proc. wyższy. Mniej niż 1% gospodarstw domowych dotkniętych tym problemem odnotowano w Szwajcarii (0,6%) oraz Norwegii (0,9%). Najgorsza sytuacja panowała na Litwie (27,9%) oraz w Bułgarii (33,7%), przy czym w przypadku obu tych krajów korzystano z danych za 2017 r. W Polsce ponad 5% gospodarstw domowych miało problem z ogrzaniem mieszkania do odpowiedniego poziomu i jest to wynik kształtujący się poniżej średniej dla UE i strefy euro.

Wykr. 1. Odsetek gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania w UE i krajach stowarzyszonych w 2018 r.^a



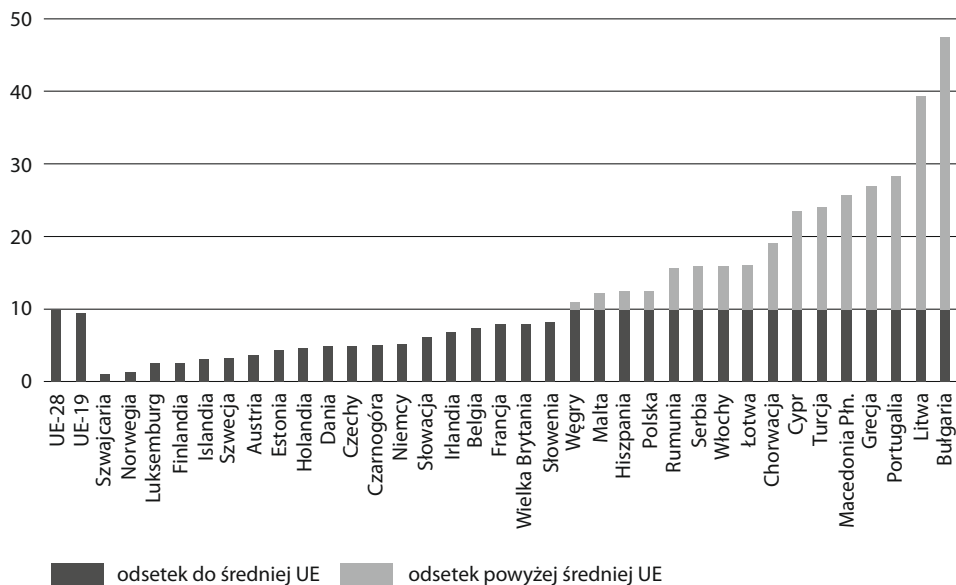
a W przypadku Bułgarii, Cypru, Litwy, Macedonii Północnej i Turcji w 2017 r., w przypadku Grecji w 2016 r. Uwaga. UE-28 – Unia Europejska (28 krajów), UE-19 – strefa euro (19 krajów).

Źródło: opracowanie własne na podstawie: Eurostat (2019a).

W przypadku jednoosobowych gospodarstw domowych najniższy odsetek gospodarstw niezdolnych do odpowiedniego ogrzania mieszkania (wykr. 2) występował w Szwajcarii (1,0%) i Norwegii (1,3%), a najwyższy – na Litwie (39,3%) i w Bułgarii (47,4%). Średnia w UE wynosiła 9,9%, a w strefie euro była o 0,5 p.proc. niższa. Od-

setek jednoosobowych gospodarstw domowych w Polsce, które dotknął dyskomfort cieplny, przewyższał średnią unijną i wynosił 12,4%.

Wykr. 2. Odsetek jednoosobowych gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania w UE i krajach stowarzyszonych w 2018 r.^a



a W przypadku Czarnogóry, Irlandii, Macedonii Północnej, Słowacji i Turcji w 2017 r., w przypadku Islandii w 2016 r.

Uwaga. Jak przy wyk. 1.

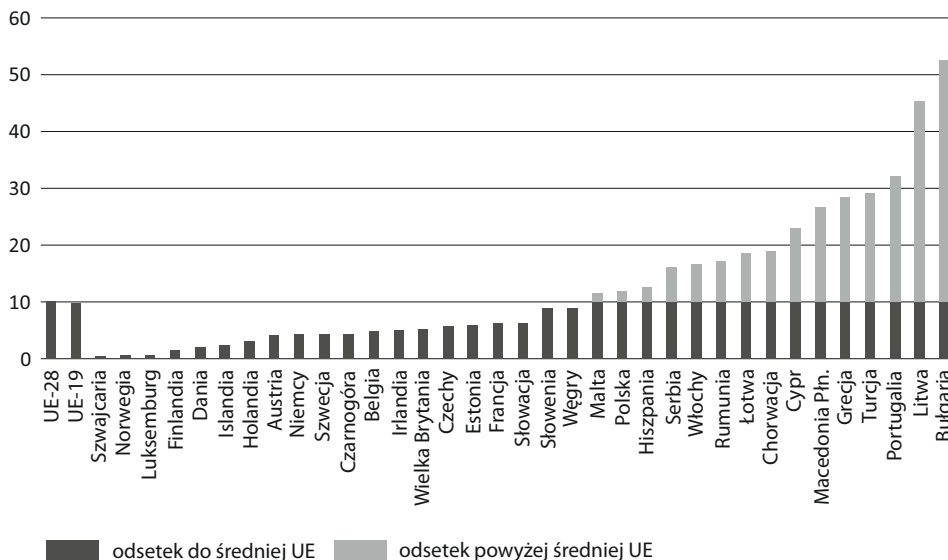
Źródło: opracowanie własne na podstawie: Eurostat (2019a).

Wśród jednoosobowych gospodarstw domowych tworzonych przez osoby w wieku 65 lat i więcej odsetek niezdolnych do odpowiedniego ogrzania mieszkania był wyższy niż dla wszystkich typów gospodarstw domowych i wynosił średnio 10,1% w całej UE oraz 9,7% w strefie euro (wykr. 3). W krajach o najniższym odsetku – czyli w Szwajcarii i Norwegii – był on niższy niż w przypadku wszystkich gospodarstw domowych, niezależnie od typu, i wynosił odpowiednio 0,4% i 0,6%. Niskie wyniki odnotowano także w Luksemburgu, Danii, Belgii, Wielkiej Brytanii, Finlandii i Czarnogórze. Natomiast w Bułgarii, Portugalii, Chorwacji oraz na Litwie i Łotwie odsetek gospodarstw niezdolnych do odpowiedniego ogrzania mieszkania był największy. W Bułgarii 52,6% gospodarstw domowych tworzonych przez osoby starsze dotykał dyskomfort cieplny. W Polsce 12,0% gospodarstw domowych nie radziło sobie z dostatecznym ogrzaniem domu – to więcej, niż wyniosła średnia w UE.

Najgorszą sytuację w 2018 r. obserwowano w gospodarstwach domowych tworzonych przez jedną osobę dorosłą z dzieckiem/dziećmi na utrzymaniu. W badaniu

EU-SILC za dziecko pozostające na utrzymaniu uważa się każdą osobę w wieku poniżej 18 lat pozostającą na utrzymaniu innego członka gospodarstwa domowego, a także osobę nieaktywną zawodowo w wieku 18–24 lat, mieszkającą z przynajmniej jednym z rodziców (Eurostat, 2019c). Najniższy odsetek osób doświadczających dyskomfortu cieplnego, które samotnie wychowują przynajmniej jedno dziecko, odnotowano w Estonii (1,9%) i Norwegii (2,2%). Najwięcej takich gospodarstw występowało w Bułgarii (41,4%) i na Cyprze (51,6%). Średnia w UE wynosiła 11,8%, a w strefie euro – 11,4%. W Polsce problem z odpowiednim ogrzaniem mieszkania miało 13,4% tego typu gospodarstw domowych (wykr. 4), przy czym należy zwrócić uwagę, że świadczenia z programu „Rodzina 500+” ukierunkowano na poprawę jakości życia dzieci. Od 2016 r. tym świadczeniem było objęte drugie i kolejne dziecko, a od 2019 r. – każde dziecko. Niemniej jednak do 2019 r. nie dotyczyło to dzieci osób wychowujących je samotnie, jeżeli na dane dziecko nie zostało ustalone świadczenie alimentacyjne na podstawie sądowego tytułu wykonawczego. Mimo że od tej reguły było kilka wyjątków, m.in. gdy nie żyło drugie z rodziców dziecka czy też ojciec dziecka był nieznanym (Ministerstwo Rodziny, Pracy i Polityki Społecznej [MRPiPS], 2019), to w tej grupie gospodarstw świadczenie nie było powszechne.

Wykr. 3. Odsetek jednoosobowych gospodarstw domowych tworzonych przez osoby w wieku 65 lat i więcej niezdolnych do odpowiedniego ogrzania mieszkania w UE i krajach stowarzyszonych w 2018 r.^a

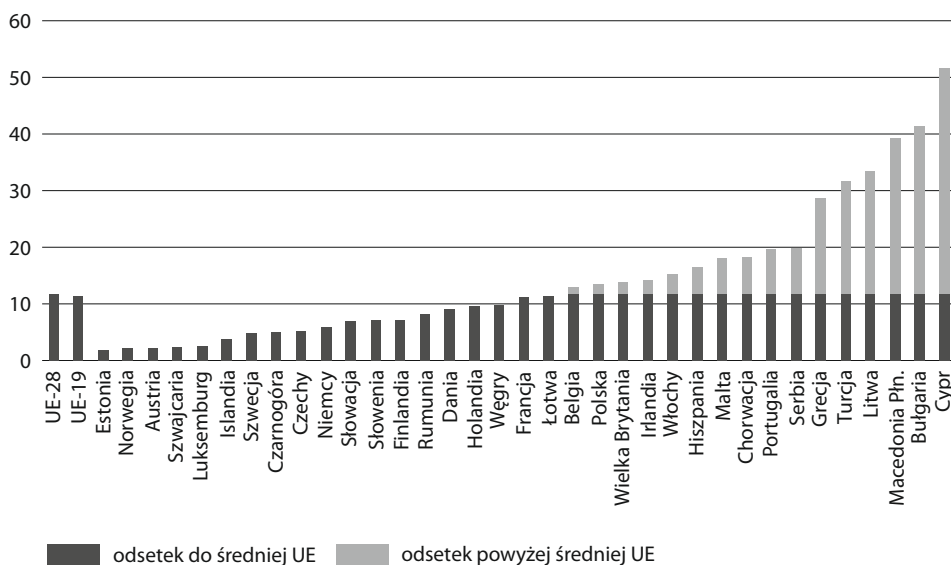


a Jak przy wykr. 2.

Uwaga. Jak przy wykr. 1.

Źródło: opracowanie własne na podstawie: Eurostat (2019a).

Wykr. 4. Odsetek gospodarstw domowych tworzonych przez osoby samotnie wychowujące dziecko/dzieci niezdolnych do odpowiedniego ogrzania mieszkania w UE i krajach stowarzyszonych w 2018 r.^a



a Jak przy wykry. 2.

Uwaga. Jak przy wykry. 1.

Źródło: opracowanie własne na podstawie: Eurostat (2019a).

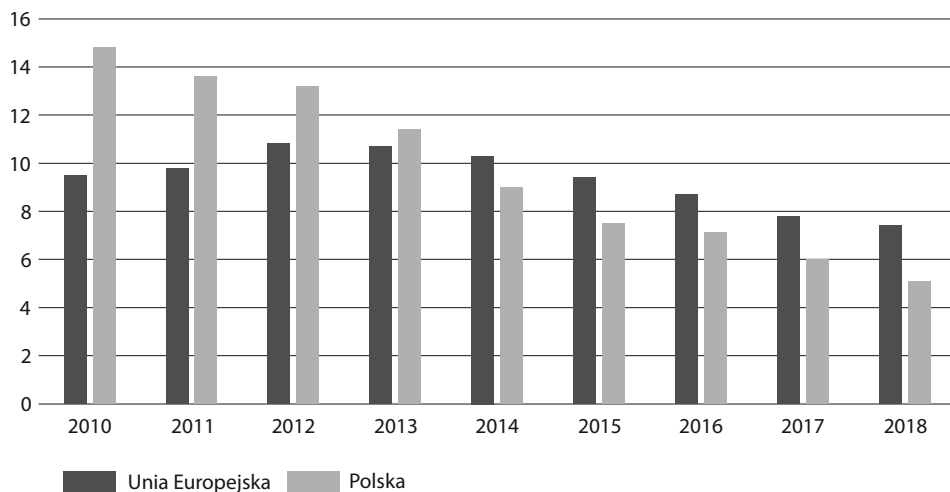
5. Zmiany skali zjawiska w latach 2009–2018

Odsetek gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania zmieniał się w okresie objętym badaniem. W UE rósł do 2012 r., a najwyższą wartość osiągnął w 2010 r. – 10,8%. W latach 2012 i 2013, kiedy w UE nastąpiło spowolnienie gospodarcze (w 2012 r. – zerowy wzrost PKB, w 2013 r. – recesja), zaobserwowano dużą liczbę takich gospodarstw (10,7%). W 2014 r. w UE odnotowano wzrost gospodarczy o 1,3% (World Bank, 2015, s. 8). Do 2018 r. odsetek gospodarstw domowych mających trudności z odpowiednim ogrzaniem mieszkania w krajach należących do UE spadał i w 2018 r. odnotowano jego najniższą wartość – 7,4%. Wykres 5 prezentuje stosowne dla UE i Polski od 2010 r. W 2009 r. Eurostat wskazał średnią jedynie dla UE; wyniosła ona 9,3%. W 2014 r. Polska po raz pierwszy odnotowała wynik niższy (9,0%) od średniej dla UE (10,3%).

Spowolnienie gospodarcze w Polsce, które nastąpiło w latach 2012 i 2013 (wzrost PKB wyniósł ok. 1,5%, przy 5,0% w 2011 r.), nie spowodowało wzrostu odsetka gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania. Taki wzrost nastąpił natomiast w przypadku średniej dla UE. W Polsce w 2009 r. 16,3% gospodarstw domowych było niezdolnych do odpowiedniego ogrzania mieszkania

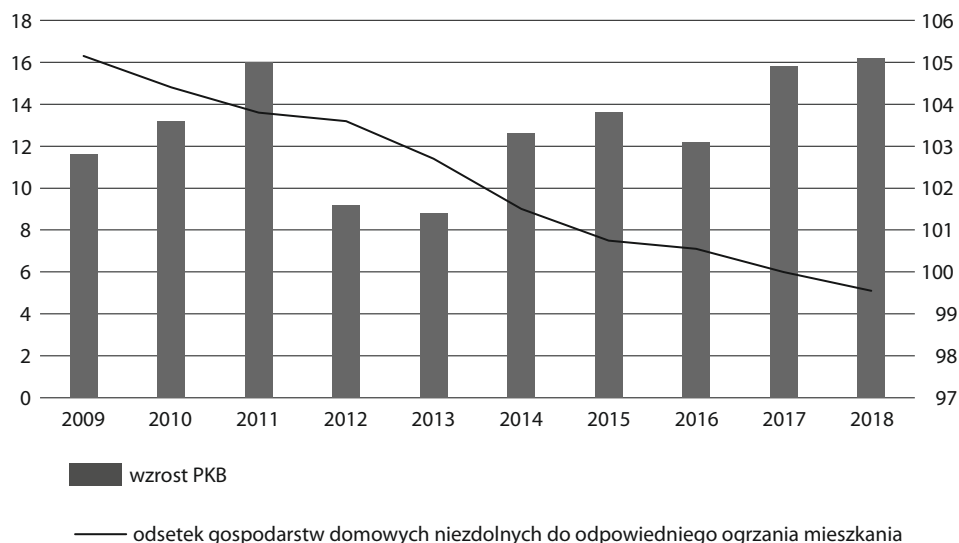
(średnio w UE – 9,3%). Do 2018 r. odsetek ten spadł do 5,1% i był niższy o 2,3 p.proc. od średniej dla UE (wykr. 6).

Wykr. 5. Odsetek gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania – porównanie wartości średniej dla UE^a i Polski



a Do 2012 r. UE-27 (bez Chorwacji).
Źródło: opracowanie własne na podstawie: Eurostat (2019a).

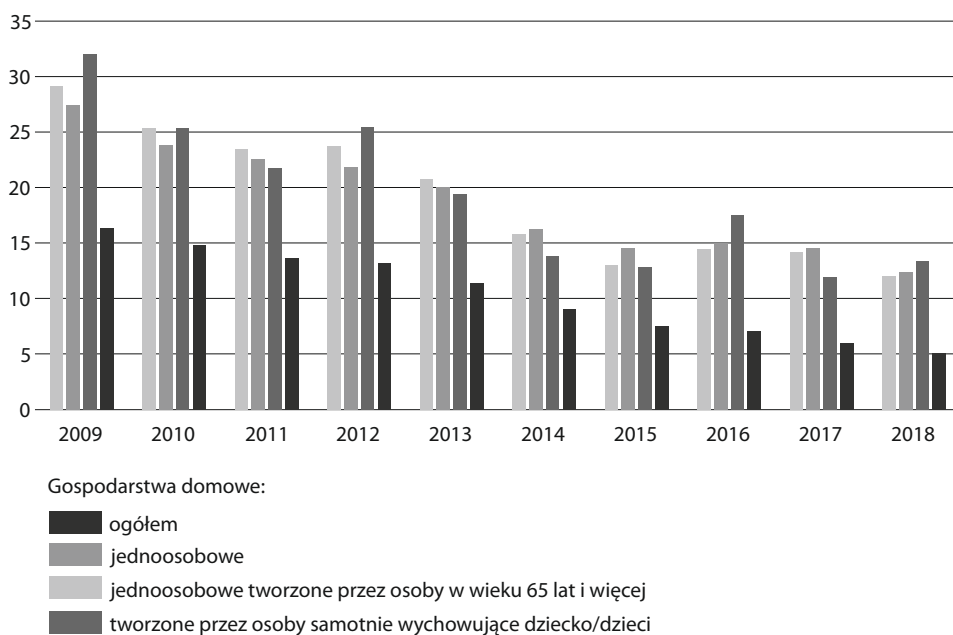
Wykr. 6. Stopa wzrostu PKB i odsetek gospodarstw domowych w Polsce niezdolnych do odpowiedniego ogrzania mieszkania



Źródło: opracowanie własne na podstawie: Eurostat (2019a) i GUS (2019b).

Najgorsza sytuacja pod względem zdolności do odpowiedniego ogrzania mieszkania wystąpiła w gospodarstwach jednoosobowych tworzonych przez osoby mające 65 lat i więcej oraz w gospodarstwach tworzonych przez osoby dorosłe z dzieckiem/dziećmi na utrzymaniu. W tych przypadkach poziom rozpatrywanego zjawiska w każdym roku objętym badaniem był wyższy niż średni dla Polski (wykr. 7).

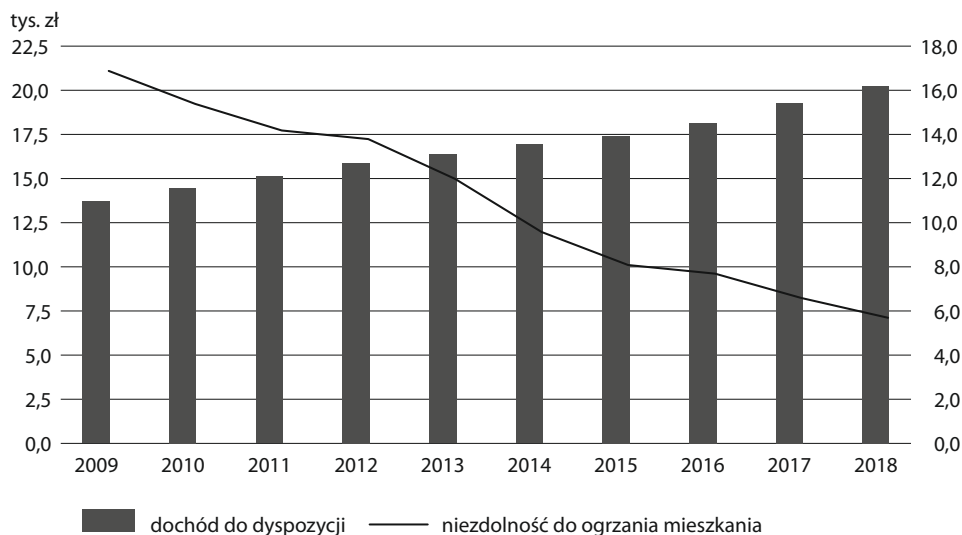
Wykr. 7. Odsetek wybranych gospodarstw domowych w Polsce niezdolnych do odpowiedniego ogrzania mieszkania



Źródło: opracowanie własne na podstawie: Eurostat (2019a).

Jak już wspomniano, ubóstwo energetyczne wynika głównie z przyczyn technicznych, ekonomicznych i związanych z postawą wobec efektywnego wykorzystywania energii. Zarówno przyczyny ekonomiczne, jak i techniczne wiążą się z niedostatecznymi zasobami finansowymi, które ma do dyspozycji gospodarstwo domowe. Dochód do dyspozycji to suma dochodów pieniężnych (w tym korzyści niepieniężnych związanych z użytkowaniem samochodu służbowego w przypadku pracy najemnej) netto (po odliczeniu zaliczek na podatek dochodowy, podatków od dochodów z własności, składek na ubezpieczenie społeczne i zdrowotne) wszystkich członków gospodarstwa domowego pomniejszona o: podatki od nieruchomości, transfery ienężne przekazane innym gospodarstwom domowym oraz rozliczenia z urzędem skarbowym (GUS, 2011, s. 44). Wykres 8 przedstawia przeciętne roczne dochody *per capita* w gospodarstwach domowych w porównaniu z odsetkiem gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania.

Wykr. 8. Przeciętne roczne dochody *per capita* w gospodarstwach domowych w Polsce i odsetek gospodarstw niezdolnych do odpowiedniego ogrzania mieszkania



Źródło: opracowanie własne na podstawie: Eurostat (2019a) i GUS (2014b, s. 144, 2014c, s. 86, 2015, s. 86, 2017a, s. 86, 2017b, s. 86, 2018, s. 204, 2019a, s. 172).

W badanym okresie widać stały wzrost przeciętnego dochodu gospodarstwa domowego oraz spadek odsetka gospodarstw niezdolnych do odpowiedniego ogrzania mieszkania. Potwierdza to obliczony dla danych z tabl. 1 współczynnik korelacji liniowej *r*-Pearsona, który wynosi 0,9733. Wynik ten wskazuje na prawie pełną ujemną zależność liniową między przeciętnym dochodem rocznym *per capita* w gospodarstwie domowym a wskaźnikiem niezdolności do odpowiedniego ogrzania mieszkania.

Tabl. 1. Przeciętne roczne dochody *per capita* w gospodarstwach domowych w Polsce i odsetek gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania

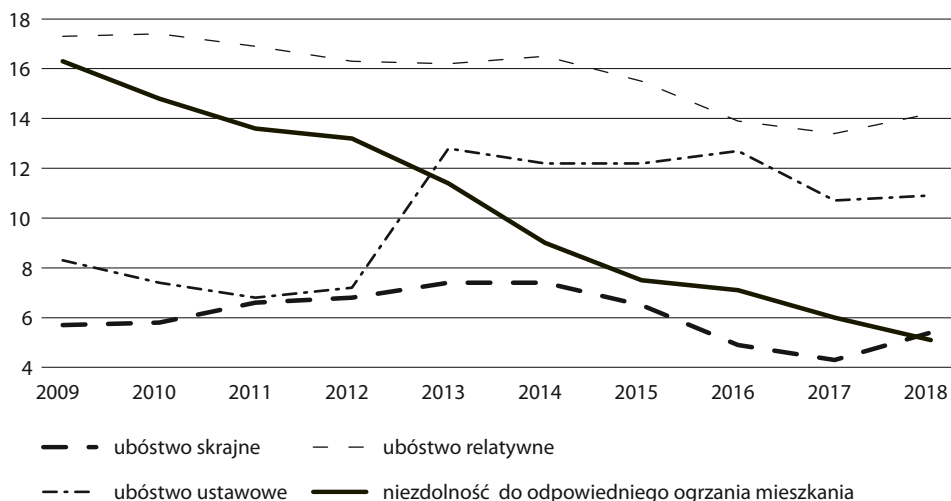
Zmienne	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Dochód do dyspozycji w zł	13681	14442	15110	15875	16349	16922	17402	18126	19275	20219
Niezdolność do odpowiedniego ogrzania mieszkania	16,3	14,8	13,6	13,2	11,4	9,0	7,5	7,1	6,0	5,1

Źródło: opracowanie własne na podstawie: Eurostat (2019a) i GUS (2014b, s. 144, 2014c, s. 86, 2015, s. 86, 2017a, s. 86, 2017b, s. 86, 2018, s. 204, 2019a, s. 172).

Z pojęciem dochodu nierozzerwalnie związane jest pojęcie ubóstwa ekonomicznego. Wykres 9 przedstawia wskaźniki ubóstwa ekonomicznego w Polsce w porównaniu z odsetkiem gospodarstw domowych niezdolnych do odpowiedniego ogrzania

mieszkania. Granicę ubóstwa skrajnego stanowi minimum egzystencji (szacowane przez Instytut Pracy i Spraw Socjalnych), które określa bardzo niski poziom zaspokojenia potrzeb. Ubóstwo ustawowe wskazuje na grupę osób potencjalnie uprawnioną do otrzymywania świadczeń pieniężnych z pomocy społecznej. Z kolei granica ubóstwa relatywnego to 50% kwoty, którą przeciętnie miesięcznie wydaje gospodarstwo domowe; wskazuje gospodarstwa, których poziom konsumpcji znacząco odbiega od przeciętnego.

Wykr. 9. Odsetek gospodarstw domowych w Polsce dotkniętych ubóstwem ekonomicznym i gospodarstw niezdolnych do odpowiedniego ogrzania mieszkania



Źródło: opracowanie własne na podstawie: Eurostat (2019a) i GUS (2020).

W związku z tym, że podstawowym wskaźnikiem ubóstwa energetycznego jest LIHC, sprawdzono zależność ubóstwa ekonomicznego i niemożności do odpowiedniego ogrzania mieszkania (tabl. 2).

Tabl. 2. Odsetek gospodarstw domowych w Polsce dotkniętych ubóstwem ekonomicznym i niezdolnych do odpowiedniego ogrzania mieszkania

Zmienne	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Ubóstwo: skrajne	5,7	5,8	6,6	6,8	7,4	7,4	6,5	4,9	4,3	5,4
relatywne	17,3	17,4	16,9	16,3	16,2	16,5	15,5	13,9	13,4	14,2
ustawowe	8,3	7,4	6,8	7,2	12,8	12,2	12,2	12,7	10,7	10,9
Niezdolność do odpowiedniego ogrzania mieszkania	16,3	14,8	13,6	13,2	11,4	9,0	7,5	7,1	6,0	5,1

Źródło: opracowanie własne na podstawie: Eurostat (2019a) i GUS (2020).

Współczynnik korelacji liniowej r -Pearsona liczony dla niezdolności do odpowiedniego ogrzania mieszkania i ubóstwa relatywnego w latach 2009–2018 wynosi 0,8954. Między zjawiskami występuje bardzo silna korelacja dodatnia – spadkowi odsetka gospodarstw dotkniętych ubóstwem relatywnym towarzyszy spadek odsetka gospodarstw domowych niezdolnych do odpowiedniego ogrzania mieszkania. Analizując współzależność między odsetkiem gospodarstw domowych niezdolnych do ogrzania mieszkania a ubóstwem skrajnym, stwierdzono przeciętną korelację wskaźników – wartość współczynnika korelacji liniowej r -Pearsona to 0,3717. Co ciekawe, w przypadku ubóstwa ustawowego wynosi ona 0,7212, co świadczy o bardzo wysokiej ujemnej korelacji między badanymi zjawiskami. Wzrostowi wskaźnika ubóstwa ustawowego towarzyszy spadek odsetka gospodarstw domowych mających problem z ogrzaniem mieszkania.

6. Podsumowanie

Przebywanie w odpowiednio ogrzonym mieszkaniu jest naturalną potrzebą, a jej niezaspokojenie negatywnie wpływa na stan zdrowia fizycznego i psychicznego członków gospodarstwa domowego, którego ten problem dotyka. Relatywnie niskie dochody gospodarstw uniemożliwiają inwestowanie w budynki mieszkalne, które ulegają stopniowej degradacji, co prowadzi do zwiększonego zapotrzebowania na nośniki energii. Z kolei większe zużycie surowców wpływa na wzrost emisji CO₂.

Rządy wielu krajów podejmują działania, aby diagnozować przyczyny ubóstwa energetycznego i mu przeciwdziałać. Punktem wyjścia jest określenie wielkości problemu. W Europie badania na szeroką skalę koordynuje Eurostat (badanie EU-SILC), a poziom ubóstwa w Polsce mierzy GUS. Podmiotem badań są gospodarstwa domowe oraz poziom zaspokojenia ich potrzeb w zakresie ogrzewania, chłodzenia, oświetlenia czy użytkowania urządzeń.

Należy rozróżnić dwie zmienne dotyczące zjawiska ubóstwa energetycznego: niezdolność do odpowiedniego ogrzania mieszkania oraz trudności z opłacaniem rachunków za prąd i gaz, które wpływają na dostęp do źródeł energii. Dopiero obie zmienne łącznie opisują ubóstwo energetyczne. Przedmiotem tego artykułu było kształtowanie się wartości zmiennej określającej ubóstwo energetyczne w aspekcie niemożności odpowiedniego ogrzania mieszkania. Z analizy wynika, że odsetek gospodarstw domowych dotkniętych tym problemem w Polsce i Europie maleje, ale są kraje (m.in. Macedonia, Litwa i Bułgaria), w których jest nadal wysoki. Niemniej jednak w Polsce od 2014 r. odnotowuje się mniejszy odsetek takich gospodarstw, niż wynosi średnia dla krajów członkowskich UE. Korelacja, mierzona współczynnikiem korelacji liniowej r -Pearsona, między niezdolnością do odpowiedniego ogrzania

mieszkania przez gospodarstwo domowe a przeciętnym dochodem *per capita* w gospodarstwie domowym wskazuje na prawie pełną ujemną zależność liniową między tymi zmiennymi. Z kolei analiza współzależności między badanym zjawiskiem a wskaźnikami ubóstwa ekonomicznego wskazuje na bardzo silną dodatnią korelację ze wskaźnikiem ubóstwa relatywnego, przeciętną dodatnią korelację ze wskaźnikiem ubóstwa skrajnego i bardzo silną ujemną korelację ze wskaźnikiem ubóstwa ustawowego.

Przeprowadzone badanie pokazało, że sytuacja ekonomiczna gospodarstw domowych w istotny sposób wpływa na odpowiednie ogrzanie mieszkania. Uzyskane wyniki mogą służyć dalszym studiom nad przyczynami i skutkami niezdolności do odpowiedniego ogrzania mieszkania oraz nad sposobami przeciwdziałania temu zjawisku.

Bibliografia

- Boguszewski, R., Herudziński, T. (2018). *Ubóstwo energetyczne w Polsce*. Warszawa: Pracownia Badań Społecznych SGGW. Pobrane z: https://www.cire.pl/pliki/2/2018/ubostwo_energetyczne_w_polsce_raport_03_09_2018.pdf.
- Bouzarovski, S. (2018). *Energy Poverty: (Dis)Assembling Europe's Infrastructural Divide*. Cham: Palgrave Macmillan. DOI: 10.1007/978-3-319-69299-9.
- Charlier, D., Legendre, B. (2019). A Multidimensional Approach to Measuring Fuel Poverty. *The Energy Journal*, 40(2), 27–53. DOI: 10.5547/01956574.40.2.bleg.
- Czakon, W. (red.). (2016). *Podstawy metodologii badań w naukach o zarządzaniu*. Warszawa: Wydawnictwo Nieoczywiste.
- EPOV. (2019). *What is Energy poverty?* Pobrane z: <https://www.energypoverty.eu/about/what-energy-poverty>.
- Eurostat. (2019a) [Baza danych]. *Inability to keep home adequately warm – EU-SILC survey*. Pobrane z: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_mdcs01&lang=en (dostęp: 25.11.2019).
- Eurostat. (2019b) [Baza danych]. *Arrears on utility bills – EU-SILC survey*. Pobrane z: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_mdcs07&lang=en (dostęp: 25.11.2019).
- Eurostat. (2019c) [Baza danych]. *Income and living conditions (ilc)*. Pobrane z: https://ec.europa.eu/eurostat/cache/metadata/en/ilc_esms.htm#meta_update1508767944514 (dostęp: 25.11.2019).
- GUS. (2011). *Dochody i warunki życia ludności (raport z badania EU-SILC 2009)*. Pobrane z: https://stat.gov.pl/cps/rde/xbcr/gus/wz_dochody_i_warunki_zycia-rap_2009.pdf.
- GUS. (2014a). *Europejskie badanie dochodów i warunków życia (EU-SILC) w 2013 r.* Pobrane z: <https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/europejskie-badanie-dochodow-i-warunkow-zycia-eu-silc-w-2013-r-,7,5.html>.
- GUS. (2014b). *Dochody i warunki życia ludności Polski (raport z badania EU-SILC 2012)*. Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5486/6/5/3/wz_dochody_warunki_zycia_raport_2012.pdf.

- GUS. (2014c). *Dochody i warunki życia ludności Polski (raport z badania EU-SILC 2013)*. Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5486/6/7/1/dochody_i_warunki_zycia_2013.pdf.
- GUS. (2015). *Dochody i warunki życia ludności Polski (raport z badania EU-SILC 2014)*. Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5486/6/8/1/dochody_i_warunki_zycia_2014_internet.pdf.
- GUS. (2017a). *Dochody i warunki życia ludności Polski (raport z badania EU-SILC 2015)*. Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5486/6/10/1/dochody_i_warunki_zycia_ludnosci_polski_raport_z_badania_eu-silc_2015.pdf.
- GUS. (2017b). *Dochody i warunki życia ludności Polski (raport z badania EU-SILC 2016)*. Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5486/6/9/1/dochody_i_warunki_zycia_ludnosci_polski_raport_z_badania_eu-silc_2016.pdf.
- GUS. (2018). *Dochody i warunki życia ludności Polski (raport z badania EU-SILC 2017)*. Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5486/6/11/1/dochody_i_warunki_zycia_ludnosci_polski_raport_z_badania_eu-silc_2017.pdf.
- GUS. (2019a). *Dochody i warunki życia ludności Polski (raport z badania EU-SILC 2018)*. Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5486/6/12/1/dochody_i_warunki_zycia_ludnosci_polski_-_raport_z_badania_eu-silc_2018.pdf.
- GUS. (2019b) [Baza danych]. *Roczne wskaźniki makroekonomiczne cz. III*. Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultstronaopisowa/1772/1/5/roczne_wskazniki_makroekonomiczne_cz_iii.xlsx (dostęp: 6.12.2019).
- GUS. (2019c). *Zużycie energii w gospodarstwach domowych w 2018 r.* Warszawa: GUS. Pobrane: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5485/2/4/1/zuzycie_energii_w_gospodarstwach_domowych_w_2018.pdf.
- GUS. (2020). *Zasięg ubóstwa ekonomicznego w Polsce w 2019 r.* Pobrane z: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5487/14/7/1/zasieg_ubostwa_ekonomicznego_w_polsce_w_2019.docx.
- Lis, M., Miazga, A., Sałach, K., Szpor, A., Świącicka, K. (2016). *Ubóstwo energetyczne w Polsce – diagnoza i rekomendacje*. Instytut Badań Strukturalnych. Pobrane z: <https://ibs.org.pl/publications/ubostwo-energetyczne-w-polsce-diagnoza-i-rekomendacje/>.
- MRPiPS. (2019). *Rodzina 500+ dla rodziców samotnie wychowujących dzieci*. Ministerstwo wyjaśnia. Pobrane z: <https://www.gov.pl/web/rodzina/rodzina-500-dla-rodzicow-samotnie-wychowujacych-dzieci-ministerstwo-wyjasnia>.
- Rozporządzenie Komisji UE 2015/1189 z dnia 28 kwietnia 2015 r. w sprawie wykonania dyrektywy Parlamentu Europejskiego i Rady 2009/125/WE.
- Szamrej-Baran, I. (2017). Wielowymiarowa analiza ubóstwa energetycznego w ujęciu regionalnym w Polsce. *Studia i Prace WNEiZ US*, 47(1), 283–295. DOI: 10.18276/sip.2017.47/1-24.
- Szamrej-Baran, I., Baran, P. (2014). Subiektywne i obiektywne mierniki ubóstwa energetycznego. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu. Research Papers of the Wrocław University of Economics*, (367), 332–339. DOI: 10.15611/pn.2014.367.36.
- Teli, D., Dimitriou, T., James, P. A. B., Bahaj, A. S., Ellison, L., Waggott, A. (2016). Fuel poverty-induced ‘prebound effect’ in achieving the anticipated carbon savings from social housing

retrofit. *Building Services Engineering Research and Technology*, 37(2), 176–193. DOI: 10.1177/0143624415621028.

Ustawa z dnia 2 grudnia 1999 r. o narodowym spisie powszechnym ludności i mieszkań w 2002 r. (Dz.U. 2000 nr 1 poz. 1).

Ustawa z dnia 25 sierpnia 2006 r. o systemie monitorowania i kontrolowania jakości paliw (Dz.U. 2006 nr 169 poz. 1200).

WHO. (2007). *Housing, energy and thermal comfort: A review of 10 countries within the WHO European Region*. Pobrane z: <https://apps.who.int/iris/bitstream/handle/10665/107815/E89887;jsessionid=74C860F3D815B1D49965121863E3072D?sequence=1>.

World Bank. (2015). *EU Regular Economic Report MODEST RECOVERY, GLOBAL RISKS: Europe and Central Asia Region The World Bank: Final Report 2015*. Pobrane z: <https://www.worldbank.org/content/dam/Worldbank/document/eca/eu-rer-1-eng.pdf>.

WYDAWNICTWA GUS. WRZESIEŃ 2020 PUBLICATIONS OF STATISTICS POLAND. SEPTEMBER 2020

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

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



Tytuł: *Polska na drodze zrównoważonego rozwoju. Raport 2020*

Title: *Poland on the way to SDGs. Report 2020*

Język: polski, angielski

Language: Polish, English

Dodatkowe informacje: publikacja cyfrowa

Additional information: digital publication

Raport z realizacji globalnych Celów Zrównoważonego Rozwoju (Sustainable Development Goals, SDG) w Polsce jest pierwszym cyfrowym opracowaniem GUS. Publikacja została przygotowana z okazji piątej rocznicy podpisania przez Polskę Agendy 2030 na rzecz zrównoważonego rozwoju i stanowi podsumowanie dotychczasowych postępów w realizacji SDG. Proces przemian społecznych i gospodarczych w naszym kraju przedstawiono na tle świata oraz Unii Europejskiej, co umożliwia porównanie osiągnięć Polski z dokonaniem krajów o zbliżonych uwarunkowaniach społeczno-gospodarczych.

Raport składa się z pięciu rozdziałów (*Ludzie, Dobrobyt, Planeta, Pokój i Partnerstwo*), nawiązujących do obszarów o kluczowym znaczeniu dla ludzi i planety, które zostały zdefiniowane w preambule Agendy 2030 jako 5P (People, Prosperity, Planet, Peace, Partnership). Jest udostępniany sukcesywnie; we wrześniu ukazała się pierwsza część pt. *Ludzie*, a kolejne będą zamieszczane do końca 2020 r. Wykorzystano możliwości, jakie daje cyfrowa forma – opracowanie zawiera interaktywne wykresy, z których można przejść do tablic z bardziej szczegółowymi danymi i pobrać je w kilku formatach, a także linki prowadzące do stron powiązanych z tematem zrównoważonego rozwoju lub źródłowych baz danych.

We wrześniu br. ukazały się ponadto:

- *Bezrobocie rejestrowane. II kwartał 2020 r.*;
- „Biuletyn statystyczny” nr 8/2020;

- *Budżety gospodarstw domowych w 2019 r.;*
- *Ceny robót budowlano-montażowych i obiektów budowlanych (lipiec 2020 r.);*
- *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach 2000–2020 (wrzesień 2020);*
- *Kultura w 2019 r.;*
- *Nakłady i wyniki przemysłu w I–II kwartale 2020 r.;*
- *Pomoc społeczna i opieka nad dzieckiem i rodziną w 2019 r.;*
- *Produkcja ważniejszych wyrobów przemysłowych w sierpniu 2020 r.;*
- *Sytuacja makroekonomiczna w Polsce na tle procesów w gospodarce światowej w 2019 r.;*
- *Sytuacja społeczno-gospodarcza kraju w sierpniu 2020 r.;*
- *Sytuacja społeczno-gospodarcza województw nr 2/2020;*
- *Telekomunikacja w 2019 r.;*
- *Transport – wyniki działalności w 2019 r.;*
- *„Wiadomości Statystyczne. The Polish Statistician” nr 9/2020;*
- *Zatrudnienie i wynagrodzenia w gospodarce narodowej w pierwszym półroczu 2020 r.;*
- *Zeszyt metodologiczny. Statystyka turystycznej bazy noclegowej.*

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

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

Justyna Gustyn (Główny Urząd Statystyczny, Departament Opracowań Statystycznych)

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 Ministerstwa Nauki i Szkolnictwa Wyższego. Zgodnie z komunikatem MNiSW z dnia 31 lipca 2019 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych wraz z przypisaną liczbą punktów „WS” otrzymały 20 punktów.

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

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

1. Zgłaszanie artykułów

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

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

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

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

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

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

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

2. Przebieg prac redakcyjnych

Zgłoszony artykuł jest oceniany i opracowywany w czteroetapowym procesie:

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

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

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

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

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

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

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

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

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

3. Zasady etyki publikacyjnej COPE

Redakcja „WS” dokłada wszelkich starań, aby utrzymać najwyższe standardy etyczne, 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, zespół redakcyjny, recenzentów i wydawcę.

3.1. Odpowiedzialność autorów

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

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

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

3.2. Odpowiedzialność zespołu redakcyjnego

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

nych przejawów nierzetelności naukowej i możliwości wystąpienia konfliktu interesów. Obiektywną ocenę oryginalności tekstu zapewnia system antyplagiatowy stosowany przez redakcję.

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

3.3. Odpowiedzialność recenzentów

1. Recenzenci przyjmują artykuł do recenzji tylko wtedy, gdy uznają, że:
 - posiadają odpowiednią wiedzę w określonej dziedzinie, aby rzetelnie ocenić pracę;
 - zgodnie z ich stanem wiedzy nie istnieje konflikt interesów w odniesieniu do autorów, przedstawionych w artykule badań i instytucji je finansujących, co potwierdzają w oświadczeniu;
 - mogą wywiązać się z terminu ustalonego przez redakcję, aby nie opóźnić publikacji.
2. Recenzenci są zobligowani do zachowania obiektywności i poufności oraz powstrzymania się od osobistej krytyki. Zawsze powinni uzasadnić swoją ocenę, przedstawiając stosowną argumentację.
3. Recenzenci powinni wskazać ważne dla wyników badań opublikowane prace, które w ich ocenie powinny zostać przywołane w ocenianym artykule.
4. W razie stwierdzenia wysokiego poziomu zbieżności treści recenzowanej pracy z innymi opublikowanymi materiałami lub podejrzenia innych przejawów nierzetelności naukowej recenzenci są zobowiązani poinformować o tym redakcję.
5. Po ukończeniu recenzji przechowywanie przesłanych przez redakcję materiałów (w jakiejkolwiek formie) oraz posługiwanie się nimi przez recenzentów jest niedozwolone.

3.4. Odpowiedzialność wydawcy

1. Materiały opublikowane w „WS” są chronione prawem autorskim.
2. Wydawca udostępnia pełną treść wszystkich artykułów w internecie na zasadach otwartego dostępu, tj. bezpłatnie i bez technicznych ograniczeń. Użytkownicy mogą czytać, pobierać, kopiować, drukować i wykorzystywać do innych celów artykuły zamieszczone online, zgodnie

- z właściwymi przepisami o dozwolonym użytku, pod warunkiem wskazania źródła pochodzenia artykułu. Inne sposoby wykorzystania treści artykułów z „WS” wymagają zgody wydawcy.
3. Wydawca deklaruje gotowość do opublikowania poprawek, wyjaśnień oraz przeprosin.

4. Wymogi redakcyjne

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

Zachęcamy do przygotowania pracy z wykorzystaniem szablonu artykułu „WS” – do pobrania ze strony ws.stat.gov.pl/ForAuthors.

4.1. Struktura i zawartość artykułu

Wymagane elementy artykułu:

1. Tytuł.
2. Dane autora: imię i nazwisko, ORCID, afiliacja.
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, przedmiot i najważniejsze wnioski.
Streszczenie to podstawowe źródło informacji o artykule, warunkujące też decyzję czytelnika o zapoznaniu się z całą pracą. Dlatego powinno być przygotowane ze szczególną starannością i dbałością o umieszczenie w nim wszystkich wymaganych elementów.
4. Słowa kluczowe – najistotniejsze pojęcia lub wyrażenia użyte w pracy (nie mniej niż trzy). Powinny być zawarte w streszczeniu i/lub tytule.
5. Kod/kody z klasyfikacji Journal of Economic Literature (JEL).
6. Tłumaczenie tytułu, streszczenia i słów kluczowych (na język angielski w przypadku artykułu napisanego w języku polskim, a na język polski w przypadku artykułu napisanego w języku angielskim).
7. W artykule opisującym badanie empiryczne wymagane są następujące części:
 - wprowadzenie, zawierające: syntetyczne przedstawienie zagadnień teoretycznych, uzasadnienie podjęcia danego problemu badawczego, cel badania i krytyczne odniesienie do literatury przedmiotu. W wyjątkowych przypadkach, kiedy istotne dla podjętego tematu jest obszerniejsze przedstawienie dyskusji toczącej się w literaturze, przegląd literatury może stanowić odrębną część artykułu;
 - metoda badania, zawierająca: przedmiot i okres badania, źródła danych i zastosowane metody badawcze, w tym uzasadnienie ich wyboru;
 - wyniki badania wraz z wnioskami;
 - podsumowanie, które powinno być zwięzłe i odzwierciedlać istotę problemu badawczego przedstawionego w artykule, bez podawania danych liczbowych; końcowe wnioski powinny odnosić się do treści artykułu, a w szczególności do celu badania.Wszystkie części powinny być opatrzone numerami.

8. Bibliografia, zawierająca pełny wykaz prac i materiałów przywołanych w artykule, przygotowana zgodnie z wymogami czasopisma.

4.2. Przygotowanie artykułu

1. Artykuł powinien być utrzymany w formie bezosobowej.
2. Tekst należy zapisać alfabetem łacińskim. Nazwy własne, tytuły itp. oryginalnie zapisane innym alfabetem powinny być poddane transliteracji.
3. Nie należy stosować stylów; formatowanie należy ograniczyć do wymogów redakcyjnych.
4. Objętość artykułu łącznie ze streszczeniem, słowami kluczowymi, bibliografią, tablicami, wykresami i innymi materiałami graficznymi nie powinna być mniejsza niż 10 stron maszynopisu ani przekraczać 20 stron.
5. Edytor tekstu: Microsoft Word, format *.doc lub *.docx.
6. Krój czcionki:
 - Arial – tytuł, autor, streszczenia, słowa kluczowe, kody JEL, śródtytuły, elementy graficzne (tablice, zestawienia, wykresy, schematy), przypisy;
 - Times New Roman – tekst główny, bibliografia.
7. Wielkość czcionki:
 - 14 pkt – tytuł, autor, tytuły rozdziałów;
 - 12 pkt – tekst główny, tytuły podrozdziałów;
 - 10 pkt – pozostałe elementy.
8. Marginesy – 2,5 cm z każdej strony.
9. Interlinia – 1,5 wiersza; tablice i przypisy – 1 wiersz; przed tytułami rozdziałów i podrozdziałów oraz po nich – pusty wiersz.
10. Wcięcie akapitowe – 0,4 cm; bibliografia – bez wcięcia, wysunięcie 0,4 cm.
11. Przy wycieniach należy posłużyć się listą punktowaną z punktorami w postaci kropek (wysunięcie 0,4 cm, wcięcie 0 cm); wiersze (oprócz ostatniego) zakończone średnikiem.
12. Strony ponumerowane automatycznie.
13. Tablice i elementy graficzne (wykresy, mapy, schematy) muszą być przywołane w tekście.
14. Wykresy, mapy i schematy należy zamieścić w tekście głównym. Wykresy powinny być edytowalne (optymalnie wykonane w programie Excel; w przypadku wykonania w programie graficznym powinny mieć postać wektorową). Dane, na podstawie których opracowano wykresy, należy przekazać osobno w pliku programu Excel (lub innym edytowalnym w pakiecie Microsoft Office), ewentualnie wykresy powinny dawać możliwość odczytania z nich danych.
15. Tablice muszą być edytowalne. Nie należy stosować rastrów, cieniowania, pogrubiania czy też podwójnych linii itp.
16. Wskazówki dotyczące opracowywania map znajdują się w publikacji *Mapy statystyczne. Opracowanie i prezentacja danych*, dostępnej na stronie internetowej GUS.
17. Pod tablicami i każdym elementem graficznym należy podać źródło opracowania, a także objaśnić użyte w nich skróty i symbole.
18. Oznaczenia literowe należy zapisywać następująco: liczby i inne wielkości niezłożone – małe lub duże litery, kursywa, bez pogrubienia (np. a , A , $y(x)$, a_i); wektory – małe litery,

kursywa, pogrubione (np. *a*, *w*, $y(x)$, w_i); macierze – duże litery, proste, pogrubione (np. **A**, **M**, **Y(x)**, **M_i**).

19. Objasnienia znaków umownych w tablicach: kreska (–) – zjawisko nie wystąpiło; zero (0) – zjawisko istniało w wielkości mniejszej od 0,5; (0,0) – zjawisko istniało w wielkości mniejszej od 0,05; kropka (.) – brak informacji, konieczność zachowania tajemnicy statystycznej, wypełnienie pozycji jest niemożliwe lub niecelowe; „w tym” – oznacza, że nie podaje się wszystkich składników sumy.
20. Stosowane są następujące skróty: tablica – tabl., wykres – wyk.
21. Przypisy rzeczowe, słownikowe lub informacyjne należy umieszczać na dole strony. Przypisy bibliograficzne, zgodnie ze standardem APA (American Psychological Association), należy podawać w tekście głównym.
22. Bibliografię należy przygotować zgodnie ze standardem APA.

4.3. Zasady przywoływania publikacji w treści artykułu

1. Jeden autor: bez względu na to, ile razy przywoływana jest praca, zawsze należy podać nazwisko autora i datę publikacji pracy, a w przypadku więcej niż jednej pracy danego autora opublikowanej w tym samym roku należy dodać kolejne litery alfabetu przy dacie (np. 2001a). Przykład zapisu: Jak stwierdza Iksiński (2001)... Badania wskazują, że... (Iksiński, 2001).
2. Dwoch autorów: bez względu na to, ile razy przywoływana jest praca, zawsze należy podać nazwiska obu autorów i datę publikacji pracy, a w przypadku więcej niż jednej pracy tych autorów opublikowanej w tym samym roku należy dodać kolejne litery alfabetu przy dacie. Nazwiska autorów zawsze należy łączyć spójnikiem „i”, nawet w przypadku przywoływania publikacji obcojęzycznej. Przykład zapisu: Jak sugerują Iksiński i Nowak (1999)... Badania wskazują, że... (Iksiński i Nowak, 1999).
3. Od trzech do pięciu autorów: przywołanie po raz pierwszy – należy wymienić nazwiska wszystkich autorów, rozdzielając je przecinkami i stawiając spójnik „i” pomiędzy dwoma ostatnimi nazwiskami. Przy kolejnych powołaniach na tę samą pracę należy podać nazwisko pierwszego autora, a nazwiska pozostałych autorów zastąpić określeniem „i współpracownicy” (gdy wchodzi one w skład zasadniczej części zdania) lub „i in.” (gdy stanowią element przypisu podanego w nawiasie). Przykład zapisu z przywołaniem po raz pierwszy: Jak sugerują Nowak, Iksiński i Jankiewicz (2003)... Badania (Nowak, Iksiński i Jankiewicz, 2003) wskazują, że... Przykład zapisu z kolejnymi przywołaniami: Badania Nowaka i współpracowników (2003)... Badania te wskazują, że... (Nowak i in., 2003).
4. Sześciu i więcej autorów: należy wymienić tylko nazwisko pierwszego autora, zarówno gdy praca przywoływana jest po raz pierwszy, jak i w późniejszych przywołaniach, a nazwiska pozostałych autorów zastąpić określeniem „i współpracownicy” (gdy wchodzi one w skład zasadniczej części zdania) lub „i in.” (gdy stanowią element przypisu podanego w nawiasie). W bibliografii załącznikowej należy umieścić nazwiska wszystkich autorów pracy. Przykład zapisu: Nowakowski i współpracownicy (1997) twierdzą, że... Pierwsze badania na ten temat sugerują... (Nowakowski i in., 1997).

5. Przywoływanie jednocześnie kilku prac: należy wymienić je alfabetycznie według nazwiska pierwszego autora. Przywołania kolejnych prac muszą być oddzielone średnikiem. Lata wydania prac tego samego autora / tych samych autorów muszą być oddzielone przecinkiem. Przykład zapisu: Iksiński (2001)... Nowak i Iksiński (1999, 2005)... (Iksiński, 1997, 1999, 2004a, 2004b; Nowak i Iksiński, 1999).
6. Przywoływanie pracy za innym autorem: stosuje się w tekście, natomiast w bibliografii należy umieścić tylko pracę czytaną. Przykład zapisu: Jak wykazał Nowakowski (1990; za: Zieniecka, 2007)... Badania sugerują, że... (Nowakowski, 1990; za: Zieniecka, 2007).

4.4. Przykłady opisu bibliograficznego

1. Bibliografia powinna być zamieszczona na końcu opracowania. Prace należy podać alfabetycznie według nazwiska pierwszego autora. W przypadku dwóch lub więcej prac tego samego autora / tych samych autorów należy je uporządkować chronologicznie według roku publikacji. Jeśli kilka prac tego samego autora / tych samych autorów zostało opublikowanych w tym samym roku, należy ułożyć je alfabetycznie według tytułu i odpowiednio oznaczyć, dopisując przy roku publikacji litery a, b, c itd.
2. Artykuł w czasopiśmie, w którym każdy kolejny numer/zeszyt (issue) w ramach jednego rocznika ma osobną numerację stron (w każdym zeszycie pierwsza strona opatrzona jest numerem 1): Nazwisko, X., Nazwisko 2, X. Y., Nazwisko 3, Z. (rok). Tytuł artykułu. *Tytuł czasopisma, rocznik(zeszyt)*, strona początku–strona końca.
3. Artykuł w czasopiśmie, w którym kolejne numery/zeszyty (issues) w ramach jednego rocznika nie mają osobnej numeracji stron (pierwsza strona w kolejnym zeszycie opatrzona jest numerem kolejnym po ostatniej stronie w zeszycie poprzednim): Nazwisko, X., Nazwisko 2, X. Y., Nazwisko 3, Z. (rok). Tytuł artykułu. *Tytuł czasopisma, rocznik*, strona początku–strona końca. Jeśli artykuł ma numer DOI (Digital Object Identifier), należy podać go na końcu opisu bibliograficznego: Nazwisko, X., Nazwisko 2, X. Y. (rok). Tytuł artykułu. *Tytuł czasopisma, rocznik*, strona początku–strona końca. DOI: xxxxx.
4. Książka: Nazwisko, X., Nazwisko 2, X. Y. (rok). *Tytuł książki*. Miejsce wydania: Wydawnictwo.
5. Książka napisana pod redakcją: Nazwisko, X. (red.). (rok). *Tytuł książki*. Miejsce wydania: Wydawnictwo.
6. Rozdział w pracy zbiorowej: Nazwisko, X. (rok). Tytuł rozdziału. W: Y. Nazwisko, B. Nazwisko 2 (red.), *Tytuł książki* (s. strona początku–strona końca). Miejsce wydania: Wydawnictwo.
7. Jeśli dany tekst znajduje się na stronie internetowej i nie jest artykułem w czasopiśmie, książką ani rozdziałem w książce, należy podać autora, datę publikacji (jeśli jest znana), tytuł, a następnie zamieścić informację o stronie, z której został pobrany, oraz – jeśli są to materiały informacyjne – datę dostępu. Tekst: Nazwisko, X. (rok). *Tytuł tekstu*. Pobrane z: adres strony internetowej (dostęp: DD.MM.RRRR).

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

ZAKRES TEMATYCZNY DZIAŁÓW

THEMATIC SCOPE OF SECTIONS

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

Studia metodologiczne

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

Statystyka w praktyce

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

Studia interdyscyplinarne. Wyzwania badawcze

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

Edukacja statystyczna

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

Z dziejów statystyki

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

Dyskusje. Recenzje. Informacje

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