Cena 15,00 zł
(VAT 8%)

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

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

LISTOPAD / NOVEMBER
ROCZNIK / VOLUME 69

2024 | 11

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

POLSKIE TOWARZYSTWO STATYSTYCZNE
POLISH STATISTICAL ASSOCIATION

GUS

TPS

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

LISTOPAD / NOVEMBER
ROCZNIK / VOLUME 69

2024 | 11 (762)

ADRES REDAKCJI I KONTAKT / EDITORIAL OFFICE ADDRESS AND CONTACT

sekretarz redakcji / Editorial Secretary: Małgorzata Zygmont (Główny Urząd Statystyczny, Polska)
Główny Urząd Statystyczny / Statistics Poland, al. Niepodległości 208, 00-925 Warszawa
ws.stat.gov.pl, e-mail: redakcja.ws@stat.gov.pl, tel./phone +48 22 608 32 25

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

Redakcja techniczna, skład i łamanie, opracowanie materiałów graficznych i korekta:
Zakład Wydawnictw Statystycznych – zespół pod kierunkiem Macieja Adamowicza
Technical editing, typesetting, preparation of graphic materials and proofreading:
Statistical Publishing Establishment – team supervised by Maciej Adamowicz

Zakład Wydawnictw Statystycznych

Druk i oprawa / Printed and bound by:
Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment
al. Niepodległości 208, 00-925 Warszawa, zws.stat.gov.pl

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

Informacje w sprawie sprzedaży i prenumeraty czasopisma / Sales and subscription of the journal:
Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment
e-mail: zws-sprzedaz@stat.gov.pl, tel./phone +48 22 608 32 10, +48 22 608 38 10

Zam. 361/2024 – nakład 275 egz.

SPIS TREŚCI
CONTENTS

OD REDAKCJI

W listopadowym numerze „Wiadomości Statystycznych. The Polish Statistician" polecamy Państwu lekturę czterech artykułów. Pierwszy dotyczy nowatorskich rozwiązań metodologicznych, dwa kolejne – różnorodnych zastosowań metod statystycznych w analizach społeczno-gospodarczych, a ostatni traktuje o istotnym zagadnieniu z zakresu terminologii statystycznej.

Dr hab. Joanna Dębicka, prof. UEW, dr hab. inż. Edyta Mazurek, prof. UEW, i dr hab. inż. Katarzyna Ostasiewicz, prof. UEW, w pracy *Proposed solutions to increase the reliability and goodness of fit of the Thurstone method* wykazują, że metoda Thurstone'a, pozwalająca na agregację preferencji i porządkowanie pod tym kątem obiektów na skali interwałowej, jest bardzo podatna na zależność od nieistotnych alternatyw (względna kolejność dwóch obiektów może zależeć od obecności w badanym zbiorze innego, niezależnego obiektu), a zatem w praktyce bywa niewiarygodna. Autorki proponują dwa ulepszenia rozpatrywanego podejścia. Pierwszym jest zastosowanie zgrubnej reguły dotyczącej progowej odległości między obiektami (w wielokrotnościach odchylenia standardowego), poniżej której uporządkowanie obiektów nie może być uważane za rzetelne, oraz empiryczne badanie stabilności porządku obiektów. Drugie polega na minimalizowaniu innego wyrażenia niż wskazane przez Thurstone'a, zwłaszcza gdy wśród częstości empirycznych występują bardzo wysokie wartości. Pozwala to na lepsze dopasowanie częstości przewidywanych do obserwowanych.

Artykuł *Assessment of the dynamics of the diffusion of mobile technologies in the Visegrad Group* dr. hab. Marcina Salamagi, prof. UEK, jest poświęcony ocenie dynamiki dyfuzji technologii mobilnych w przemyśle przetwórczym w Grupie Wyszehradzkiej (V4). Autor bada to zjawisko za pomocą funkcji Gompertza. Określa także charakter i siłę wpływu specjalizacji inwestycyjnej (atrakcyjności inwestycyjnej) oraz specjalizacji eksportowej na dyfuzję innowacji. W tym celu wykorzystuje model ekonometryczny z mechanizmem korekty błędem (ECM), oszacowany dla wybranych sektorów przetwórstwa przemysłowego na podstawie danych dotyczących handlu zagranicznego, innowacyjności i bezpośrednich inwestycji zagranicznych za lata 2010–2021. Uzyskane wyniki pozwalają na stwierdzenie, że dyfuzja technologii mobilnych w krajach V4 występuje w różnych branżach, a jej tempo jest zróżnicowane w zależności od stopnia zaawansowania technologicznego branży i struktury własnościowej kapitału przedsiębiorstwa.

W pracy *Ocena sytuacji ekonomiczno-finansowej przemysłu spożywczego z wykorzystaniem metody TOPSIS* mgr Jadwiga Drożdż dokonuje wieloaspektowej oceny ekonomiczno-finansowej sytuacji przemysłu spożywczego w Polsce w ujęciu międzybranżowym i dynamicznym. Do obliczeń wykorzystuje publikowane i niepublikowane dane Głównego Urzędu Statystycznego dotyczące wyników osiągniętych przez producentów żywności i napojów w 2022 r., które następnie porównuje z danymi za 2019 r. Na podstawie wartości miernika syntetycznego skonstruowanego za pomocą metody TOPSIS (czyli będącego funkcją odległości euklidesowych zarówno od wzorca, jak i antywzorca rozwoju) stwierdza, że sytuacja ekonomiczno-finansowa poszczególnych branż przemysłu spożywczego w 2022 r. była dobra i stabilna, co świadczy o dużej odporności sektora na niekorzystne warunki makroekonomiczne i o jego zdolności dostosowywania się do zmiennego i nieprzewidywalnego otoczenia rynkowego. Dotyczy to zwłaszcza branż ukierunkowanych na produkcję niektórych używek i przetwórstwo wtórne.

Prof. dr hab. Mirosław Szreder w artykule *Próba reprezentatywna – potrzeba i propozycja definicji* dokonuje krytycznego omówienia roli próby reprezentatywnej we wnioskowaniu naukowym i statystycznym. Proponuje ponadto definicję *próby reprezentatywnej* i wskazuje jej specyfikę w przypadku prób nielosowych, których reprezentatywność jest ograniczona jedynie do ściśle wskazanych zmiennych. W proponowanej definicji kładzie nacisk na zgodność struktury próby ze strukturą populacji oraz na możliwości wnioskowania, w którym ocenom uzyskanym z próby są przypisane odpowiednio oszacowane liczbowe wartości błędów. Postuluje także, aby próbę, która wykazuje zbieżność rozkładów tylko niektórych cech z ich rozkładami populacyjnymi, nazywać *próbą reprezentatywną ze względu na wyszczególnione zmienne*. Podkreśla przy tym, że jednoznaczne rozumienie pojęcia *próby reprezentatywnej* jest ważne przede wszystkim dlatego, że pozwala na pełną i poprawną interpretację zarówno wyników badania próbkowego, jak i błędów, którymi te wyniki mogą być obciążone.

Ponadto w numerze zamieszczamy sprawozdanie z jedenastej międzynarodowej konferencji „Privacy in Statistical Databases 2024" (PSD), która odbyła się w dniach 25–27 września 2024 r. w Antibes Juan-les-Pins (Francja), opracowane przez dr. hab. Andrzeja Młodaka, prof. UK, i dr. Tomasza Józefowskiego. Konferencje PSD, organizowane co dwa lata w różnych krajach Europy, służą wymianie doświadczeń w zakresie ogólnoświatowych badań dotyczących ochrony prywatności w statystycznych bazach danych i wytyczaniu nowych kierunków w tej dziedzinie, w tym z wykorzystaniem narzędzi sztucznej inteligencji.

Wydanie tradycyjnie zawiera także prezentację najnowszych publikacji GUS przygotowaną przez Joannę Sadowy.

Życzymy miłej lektury.

FROM THE EDITORIAL TEAM

The November issue of *Wiadomości Statystyczne. The Polish Statistician* features four articles. The first paper presents innovative methodological solutions, the following two propose various applications of statistical methods to social and economic analyses, and the last one focuses on an important issue from the field of statistical terminology.

The demonstration that the Thurstone method, allowing the aggregation of preferences and ordering objects on an interval scale, is highly susceptible to being dependent on irrelevant alternatives (where the relative order of two objects might depend on the presence of a another totally independent object in the studied collectivity), which in practice often renders this method unreliable, is the main contribution of the paper entitled *Proposed solutions to increase the reliability and goodness of fit of the Thurstone method* by Joanna Dębicka, PhD, DSc, Edyta Mazurek, BEng, PhD, DSc, and Katarzyna Ostasiewicz, BEng, PhD, DSc, Professors at Wroclaw University of Economics and Business. The authors propose two improvements to the Thurstone method. The first one consists in the application of the rule of thumb to the minimal spacing (in terms of standard deviations), below which the ordering of objects cannot be regarded as reliable, and the empirical examination of the stability of the objects' order. The other improvement involves minimising a formula different than Thurstone's original one, especially in cases where empirical frequencies reach very high values. This allows a better fit of the observed frequencies to the predicted ones.

The assessment of the diffusion of mobile technologies in the processing industry of the Visegrad Group (V4) is the main theme in the paper *Assessment of the dynamics of the diffusion of mobile technologies in the Visegrad Group* by Marcin Salamaga, PhD, DSc, Professor at Krakow University of Economics. The author researches this phenomenon by means of the Gompertz function.

The study determines the nature and strength of the impact of investment specialisation (investment attractiveness) and industry export specialisation on the diffusion of innovation. To this end, the study employs an econometric model with an error correction mechanism (ECM), estimated for selected sectors of industrial processing on the basis of data on foreign trade, innovativeness and direct foreign investment in the period of 2010–2021. The obtained results allow a conclusion that the phenomenon of mobile technologies diffusion occurs in several sectors of the economy of the V4 countries, and its rate varies depending on the degree of a branch's technological advancement and the ownership structure of an enterprise's capital.

The study presented in the article *Assessment of the economic and financial situation of the food industry by means of the TOPSIS method* by Jadwiga Drożdż, MSc, focuses on a multi-aspect assessment of the economic and financial situation of the food and drink industry in Poland from an intersectoral and dynamic perspective. The calculations are based on both published and unpublished data from Statistics Poland concerning the results of food and drink producers in 2022. These data are then compared to the analogous data for 2019. On the basis of the values of a synthetic indicator constructed by means of the TOPSIS method (i.e. the function of Euclidean distances from both the ideal and anti-ideal development patterns), the author concludes that the economic and financial situation of the particular branches of the food and drink industry in 2022 was relatively good and stable, which testifies to the strong resistance of the whole sector to unfavourable macroeconomic conditions and to its ability to adapt to changeable and unpredictable market conditions. This conclusion is especially valid in the case of producers of some stimulants and secondary-processing branches.

An objective discussion on the role of a representative sample in scientific and statistical inference is the main subject of the article by Mirosław Szreder, PhD, DSc, ProfTit, entitled *Representative sample – the need for and the proposal of a definition*. The author proposes a definition of a 'representative sample' and shows its specificity in the case where at the same time it is a non-random sample, whose representativeness is limited to some specified variables only. In the proposed definition, the author emphasises the need for the consistency of the sample structure with the structure of the population, and the possibility of making inference in which estimates obtained from the sample have correctly-estimated numerical error values assigned to them. It is also argued that samples which show convergence of the distributions of only some characteristics with their population distributions should be called 'representative samples in terms of specified variables'. The author emphasises that the unambiguous understanding of the 'representative sample' term is important mostly because it allows a full and correct interpretation of both the results of sample surveys and errors with which they might be charged.

The November issue of our journal also features a report from the 11th edition of the *Privacy in Statistical Databases 2024 international conference,* which was held on 25th–27th September 2024 in Antibes Juan-les-Pins, France. The report has been prepared by Andrzej Młodak, PhD, DSc, Professor at Kalisz University, and Tomasz Józefowski, PhD. PSD conferences, organised every two years in different European countries, serve the purpose of exchanging experiences in the field of global research on the protection of the privacy of statistical databases, and determining new directions in this area, including doing so by means of AI tools.

The issue concludes with the presentation of Statistics Poland's most recent publications prepared by Joanna Sadowy.

We wish you pleasant reading.

Wiadomości Statystyczne. The Polish Statistician, 2024, *69*(11), 1–17 https://doi.org/10.59139/ws.2024.11.1
Studia metodologiczne / Methodological studies
Zgłoszony/submitted: 20.08.2023, poprawiony/revised: 21.07.2024, zaakceptowany/accepted: 2.09.2024

Proposed solutions to increase the reliability and goodness of fit of the Thurstone method

Joanna Dębicka,[a] Edyta Mazurek,[b] Katarzyna Ostasiewicz[c]

Abstract. The Thurstone method is a method of aggregation of preferences that leads to ordering objects at an interval scale, in contrast to other methods of aggregation, the application of which results in obtaining the ordinal scale only. However, we demonstrate that the interval scale obtained by means of the Thurstone method is not fully reliable, as it is strongly susceptible to the problem of irrelevant alternatives, i.e., the order of two objects might be dependent on whether there is or not a third, completely independent object in the studied collectivity.

The other issue examined in the paper is the formula that is to be minimised in the Thurstone method. Thurstone proposed a formula easy to be minimised with tabularised values of normal distribution. However, today's calculation powers of computers make it possible to minimise another formula, which allows a better fit of observed frequencies to the predicted ones, and makes it possible to avoid the problem with empirical frequencies close to 1.

The aim of the paper is to propose two improvements to the Thurstone method. The first one involves the application of the rule of thumb to the minimal spacing (in terms of standard deviations), below which the ordering of objects cannot be regarded as reliable, and the use of the method of empirical examining of the stability of order. The second of the proposed improvements consists in minimising a formula other than the original one, especially in the cases where empirical frequencies reach very high values.

Keywords: the Thurstone method, dependence on irrelevant alternatives, aggregation of preferences, ordering of objects

JEL: C18, C83, C46

Propozycja rozwiązań zwiększających wiarygodność i dobroć dopasowania w metodzie Thurstone'a

Streszczenie. Metoda Thurstone'a jest metodą agregacji preferencji, pozwalającą na porządkowanie obiektów na skali interwałowej, w odróżnieniu od innych metod agregacji, za pomocą których otrzymuje się jedynie skalę porządkową. Trudno jednak traktować skalę interwałową

[a] Uniwersytet Ekonomiczny we Wrocławiu, Wydział Ekonomii i Finansów, Polska / Wroclaw University of Economics and Business, Faculty of Economics and Finance, Poland.
ORCID: https://orcid.org/0000-0002-8905-2565. E-mail: joanna.debicka@ue.wroc.pl.

[b] Uniwersytet Ekonomiczny we Wrocławiu, Wydział Ekonomii i Finansów, Polska / Wroclaw University of Economics and Business, Faculty of Economics and Finance, Poland.
ORCID: https://orcid.org/0000-0001-7410-1638. E-mail: edyta.mazurek@ue.wroc.pl.

[c] Uniwersytet Ekonomiczny we Wrocławiu, Wydział Ekonomii i Finansów, Polska / Wroclaw University of Economics and Business, Faculty of Economics and Finance, Poland.
ORCID: https://orcid.org/0000-0002-0115-3696. Autor korespondencyjny / Corresponding author, e-mail: katarzyna.ostasiewicz@ue.wroc.pl.

uzyskaną przy użyciu metody Thurstone'a jako wiarygodną, ponieważ – jak pokazano w artykule – jest ona bardzo podatna na zależność od nieistotnych alternatyw: względna kolejność dwóch obiektów może zależeć od obecności innego, zupełnie niezależnego obiektu w badanym zbiorze.

W artykule poruszono również kwestię wyrażenia, które ma podlegać minimalizacji. Thurstone wskazał wyrażenie, które można łatwo zminimalizować za pomocą stablicowanych wartości rozkładu normalnego. Jednak wobec obecnych mocy obliczeniowych komputerów możliwe jest minimalizowanie innego wyrażenia, które daje lepszą dobroć dopasowania częstości obserwowanych do oczekiwanych oraz pozwala uniknąć problemu dotyczącego częstości empirycznych bliskich 1.

Artykuł ma na celu zaproponowanie dwóch ulepszeń metody Thurstone'a. Pierwszym jest stosowanie zgrubnej reguły dotyczącej progowej odległości między obiektami (w wielokrotnościach odchylenia standardowego), poniżej której uporządkowanie obiektów nie może być uważane za rzetelne, oraz metody empirycznego badania stabilności porządku obiektów. Drugie polega na minimalizowaniu wyrażenia innego niż oryginalne, w szczególności gdy wśród częstości empirycznych występują bardzo wysokie wartości.

Słowa kluczowe: metoda Thurstone'a, zależność od nieistotnych alternatyw, agregacja preferencji, porządkowanie obiektów

1. Introduction

The problem of ordering objects is an important issue in many fields of the public life; for example, public elections involve ordering 'objects', which in this case are the candidates. There are several other situations where ordering 'objects' is necessary, such as projects to be financed by the government, etc.

Many different methods of ordering objects have been investigated. One of them is the Thurstone method (Thurstone, 1927; Thurstone & Chave, 1929; Thurstone & Jones, 1957), which is regarded useful not only for ordering objects, but also for finding relative spaces between them. While in e.g. a presidential election it is only important who wins (i.e. who obtains the 'first rank'), in other situations – for instance, distributing funds – it might be necessary to know which subsequent results are very close to one another and which are largely distanced.

The Thurstone method has been widely used in various contexts (e.g. Dragonetti et al., 2017; Heldsinger & Humphry, 2010; Kanda, 2002; Krabbe, 2008; Orbán-Mihálykó et al., 2022; Temesi et al., 2023; van Wijk & Hoogstraten, 2004; Woods et al., 2010), and its properties have been investigated intensively (e.g. Handley, 2001; Lipovetsky, 2007; Lipovetsky & Conklin, 2004; Vojnovic & Yun, 2016; Yellott, Jr., 1980). However, there seems to be few studies with respect to empirical results of applying the Thurstone method to real data.

We already discussed the subject of testing the Thurstone method (Dębicka et al., 2023). We showed that although a very smart test was developed for the Thurstone model (Mosteller, 1951), if it is applied to a sample consisting of 100 or more items, almost always the results will turn out statistically insignificant. In this paper, we show

that although the Thurstone method is touted as finding the interval scale, this scale is prone to change when one or more objects are removed from the set. This is called 'dependence on irrelevant alternatives', and obviously it is not a desirable property of a method.

We begin by having a closer look at the historical conditioning of the Thurstone method, i.e. the formulae to be minimised in order to obtain the estimates of modes of distributions. These formulae are still used nowadays, even though these days, aided by advanced computational powers of contemporary computers, scientists could come up with far more effective solutions.

We continue with the investigation of the stability of the scale obtained with the Thurstone method on the basis of two sets of data, i.e. the original Thurstone data (1927) and the data from our own survey (Dębicka et al., 2022). We are empirically checking the stability of the ordering of two items when changes in the set of objects and in the set of respondents occur.

The aim of our paper is twofold. Firstly, we propose the rule of thumb relating to the minimal spacing (in terms of standard deviations), below which the ordering of objects cannot be regarded as reliable, and the methods of empirical examining stability of order. Secondly, we indicate the possibility of minimising a formula other than the original one, especially in the cases where very high empirical frequencies occur.

2. The Thurstone method

According to the most common and simplest presentation of the Thurstone method, we have *n* objects, and each of them has some objective position. However, individuals can make errors comparing any pair of these objects and those errors are distributed normally (in the simplest case, with the same standard deviation σ for each pair). The greater the difference between the objects (given in units of σ), the smaller the probability that the respondent will misidentify a smaller object for a greater one. Let us express position (modes), l_k, in terms of multiplicity, x_k, of a common standard deviation σ: $l_k = x_k \cdot \sigma$. For objects i and j with positions equal $l_i = x_i \cdot \sigma$ and $l_j = x_j \cdot \sigma$, $x_j > x_i$, this probability of misidentification is equal to:

$$1 - p_{ji} = P(l_i > l_j) = F(x_i - x_j) = 1 - F(x_j - x_i),$$

where p_{ji} denotes the probability of proper identification of the order of objects *i* and *j*, and *F* is the cumulative distribution of the standard normal random variable.

If we had an infinite number of observations, the observed frequency of respondents who correctly identified the order of objects would be equal exactly to the theoretical probability:

$$\omega_{ji} = p_{ji},$$

where ω_{ji} denotes the observed frequency of respondents stating that object j is greater than i.

However, due to the inevitable limitations of the empirical study, in general, the observed frequency will deviate from real probabilities. Thus, our task is to estimate the real probabilities and real positions of the objects on the basis of the observations. To this end, the expression:

$$\mathcal{F}_1(\{x_i\}) = \sum_{j>i}\left[F^{-1}(\omega_{ji}) - (x_j - x_i)\right]^2, \tag{1}$$

is minimised with respect to x_i.

As zero on the obtained scale is arbitrary, one may conveniently rescale $\{x_i\}$ into a $[0, 1]$ interval:

$$x_i' = \frac{x_i - \min\{x_i\}}{\max\{x_i\} - \min\{x_i\}}.$$

In this way, we obtain not only the order of objects, but also their relative spacing, that is, not an ordinal scale, but an interval one. This is the advantage of the Thurstone method over simpler ones, e.g. average ranks method.

3. Problem with infinite inverse distribution

In the original Thurstone method, the object of minimisation is formula (1), and all the existing programs that have implemented the Thurstone method apply this approach. However, some problems with infinities may appear, as $F^{-1}(1) = \infty$ and $F^{-1}(0) = -\infty$. It is not highly improbable to obtain such values, if one of the objects in the set is clearly greater than all the others. In such a case, one may remove such an object from the whole set. Another way is to treat the problem purely numerically, e.g. to convert the empirical frequency 1 into 0.9999 or some other quantity close to 1, but not generating infinities. For example, the Statistica program replaces frequency 1 with 0.9999, but this value is completely arbitrary – why 0.9999 and not 0.999 or 0.99999? In other programs, e.g. SPSS or Stata, the Thurstone method

is not implemented – the user does the analysis employing the general tools available in these programs, and has to manage him- or herself the problem with frequencies 0% or 100%. However, such a technical 'trick' seems to be artificial and arbitrary, and the final results for the whole scale may depend on the actual value adopted, as it is the very essence of the Thurstone method (i.e. that the position of object i may depend on frequency $\omega_{jk}, j, k \neq i$). Let us say, we have a matrix of empirical frequencies as given in Table 1.

Table 1. Sample values illustrating the problem of approximating the 100% frequency

j \ i	O1	O2	O3
O1	.	1	0.35
O2	0	.	0.99
O3	0.65	0.01	.

Note. The quantities in table (ω_{ij}) denote empirical frequencies of respondents preferring object i over object j.
Source: authors' calculations based on fictious data.

If one converted 1 into e.g. 0.999999, the scaled result would be 0, 1, 0.93 (for O1, O2 and O3, respectively). If one decided for two decimal parts less, i.e. 0.9999, it would be 0, 0.90, 1, i.e. O2 and O3 would change their relative positions. If one used inverse PDF equal to 3.05 (as Krabbe, 2008, did), it would be 0, 0.75, 1. This obviously is an artificial example, but it shows the problem with empirical frequencies equal to 1 or 0.

However, the formula that Thurstone decided to minimise in his method might have a historical background. An alternative formula to be minimised, which still complies with the general idea of minimising the deviations of the empirical data from the predicted values is:

$$\mathcal{F}_2 = \sum_{j>i}\left[\omega_{ji} - F\left(x_j - x_i\right)\right]^2. \tag{2}$$

Please note that the same authors who minimise the original Thurstone expression use the following quantity as a criterion of goodness of fit (cf. e.g. Krabbe, 2008; Lipovetsky, 2007; Mosteller, 1951) and the same expressions underly the test (Mosteller, 1951):

$$\chi^2 = \frac{\sum_{ij}\left(p_{ij} - \omega_{ij}\right)^2}{\sigma_{ij}^2}. \tag{3}$$

It is obvious that by minimising (2), one would obtain a better fit. So, why use (1)?

The historical reason is not likely that in the 1920s, minimising the latter was completely impossible, as computers were not available. Using (1), one might apply tables (as Thurstone did). Please note that by today's standards of accuracy, there are some errors in Thurstone's results (1927) due to the erroneous rounding. However, it is still tempting to use (1), as one obtains analytical formulae this way, but the problem with infinities and worse goodness of fit remains.

If the data perfectly fitted the predictions, both approaches – minimising (1) and (2) – would yield the same results, which is obvious, as there would exist such quantities as to make both (1) and (2) equal exactly 0. However, if there were some deviations, the results yielded by both methods would slightly differ.

In his seminal paper, Thurstone (1927) presented results of the analysis of 19 crimes using his own method to find ordering of these crimes. We repeated the analysis of the same data by means of formula (2), instead of the original one, to compare the results. They are presented in Table 2.

Table 2. Comparison of Thurstone's original results with the scale obtained by minimising (2)

Type of crime	Original Thurstone (1)	Minimising (2)
Abortion	0.69326	0.74232
Adultery	0.64146	0.68019
Arson	0.61714	0.64457
Assault	0.45010	0.49320
Bootlegging	0.31527	0.35629
Burglary	0.46059	0.50698
Counterfeiting	0.49876	0.53702
Embezzlement	0.50608	0.55089
Forgery	0.47691	0.51623
Homicide	0.96322	0.99414
Kidnapping	0.67115	0.68534
Larceny	0.40448	0.44478
Libel	0.34304	0.38178
Perjury	0.51172	0.53799
Rape	1	1
Receiving stolen goods	0.30464	0.34566
Seduction	0.69368	0.72590
Smuggling	0.33652	0.37149
Vagrancy	0	0

Source: authors' calculations based on data from Thurstone (1927).

As one can see, no object changed its relative position with respect to the others. Still, spacing between the objects varied in some cases even by as much as over 10%. If we are interested in proper spacing – which is usually presented as the merit of the Thurstone method – should not it rather be these distances that fit the data better?

4. Dependence on irrelevant alternatives

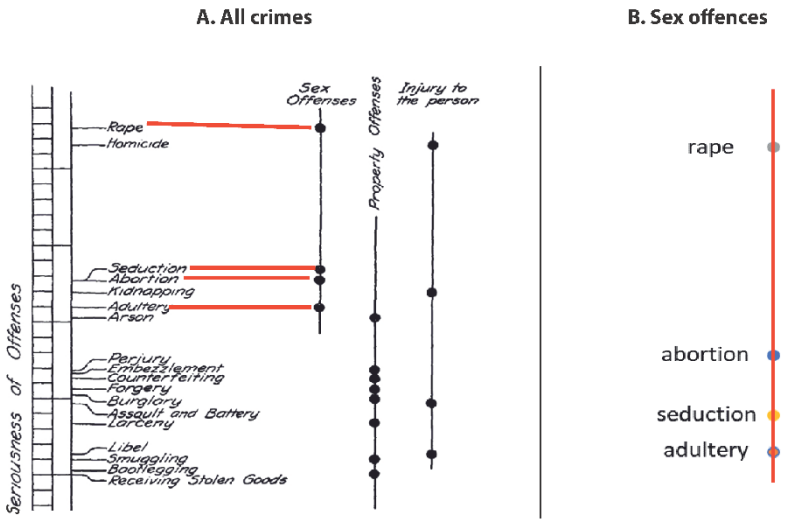
The Thurstone method is an aggregation procedure. It is rather impossible that there is some objective, set scale of crimes, and the fact there are differences between particular people's ordering of crimes results from random errors. More likely, each person has his or her own, subjective scale. Therefore, the task of researchers is to construct an aggregated scale for the whole population.

On the other hand, according to Arrow's theorem, there is no method of collective choice that would meet some reasonable assumptions (e.g. dictatorship is bad) and would be independent from irrelevant alternatives (Arrow, 1951). This applies to the Thurstone method (it being the method of preference aggregation) as well – the ordering of preferences done by means of this method would be, at least theoretically, dependent on irrelevant alternatives. However, one might ask here whether this problem is frequently encountered in empirical data. If it is, another question would have to be posed, namely how to interpret the ordering of preferences with such a shortcoming.

Thurstone (1927) investigated 19 crimes – the respondents had to compare each pair of them. Thus, by applying his method, he obtained a linear ordering of those crimes. Subsequently, he was able to present an order within the subgroups of crimes, dividing them into 'sex offences', 'property offences' and 'injury to the person'.

The disturbing fact is that if sexual crimes were just taken off the general ordering (see Figure 1a), their relative order would be quite different from the order obtained with data limited only to these pairs of crimes where both crimes belonged to the 'sexual offences' subgroup (cf. Figure 1b). Therefore, to construct 1b, we dealt with a restricted matrix of frequencies, i.e. only 4 x 4 (for those four crimes of our special interest). But we can see that in this approach, 'seduction' and 'abortion' changed their relative positions and the spacing between them changed significantly. As the respondents were supposed to compare each pair independently, this result is troublesome to say the least. If we were to decide on the basis of this result which of the sexual crimes to punish most harshly or on preventing which of them to spend most money or effort, we would have trouble deciding which result should be treated as the reliable one. The fact that it is possible the result of aggregating preferences depends on irrelevant objects which might (but do not have to) be taken into consideration in the whole procedure is indeed troublesome from the scientific point of view.

Figure 1. Ordering all crimes and sex offences



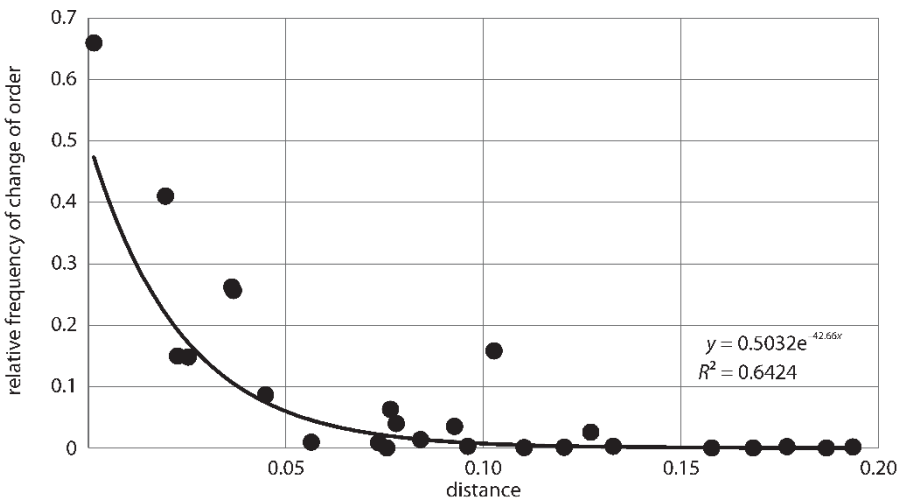
Source: A – Thurstone (1927), B – authors' work.

To investigate the deepness of the problem, we performed a detailed and complete analysis of the dataset presented by Thurstone. We examined every possible combination of crimes, i.e. starting from the set of two crimes, we took into account all the possible pairs, finally reaching all combinations for the 18 out of all 19 crimes. The research question was: are there any further changes of order among the crimes, like this change which we observed within the set of sexual crimes? Which crimes are most prone to this 'instability' of position?

Table 3 (p. 10) shows all the obtained inconsistencies. We treated the ordering of all the 19 crimes as a reference point, and listed all the deviations from the relative order in all possible subsets (combinations of crimes). Table 3 presents the relative frequencies of these changes (each pair of objects appears together in the total number of 131,070 subsets). The objects are ordered according to the overall Thurstone scale, which means that the closer to the diagonal a given cell is, the closer the crimes from a given pair are to each other on the Thurstone scale (constructed for the whole set of crimes). The most frequently observed alteration was the relative change of positions of the crimes neighbouring on the overall scale – nearly all these pairs changed at least in some subsets – but also in relation to the second nearest neighbours and even third nearest neighbours. In fact, the propensity to changing the position depends not on the relative ranks of objects, but rather on the distance between them.

Table 4 (p. 11) presents the distances between all the pairs of objects (measured in units of standard deviation), with those cells which show distances lower than 0.2 highlighted. It is striking how these highlighted cells coincide with non-zero values in Table 3. Moreover, if we gather all non-zero change-of-order relative frequencies and corresponding distances, the relationship between those quantities is strong, as pictured in Figure 2.

Figure 2. Relative frequencies of change of order of pairs of objects vs the distances between them as obtained by the Thurstone procedure



Source: authors' work based on data from Thurstone (1927) and authors' own calculations.

We repeated the same procedure for data obtained in our own research (Dębicka et al., 2022), with the set of crimes different in some parts from Thurstone's original set, in order to fit our reality better. The crimes were arranged in the following order: violent rape, assault with a severe body injury, pedophilic acts, domestic violence, threats, murder, defamation (slander), harassment, kidnapping for ransom and identity theft. Our results are presented in Tables 5 and 6 (for frequencies of changes and distances, respectively, p. 12). One can see that these results were much more stable than Thurstone's original results when we analysed their dependence on irrelevant alternatives. There are only two pairs whose relative order might be questioned – and these are the same two pairs which are separated by the smallest distances on the overall Thurstone scale.

Table 3. Relative frequencies of exchanging order based on Thurstone's data

Type of crime	O19	O16	O5	O18	O13	O12	O4	O6	O9	O7	O8	O14	O3	O2	O11	O1	O17	O10	O15
O19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O16	0	0	0.256	0.001	0.003	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O5	0	0.256	0	0.009	0.003	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O18	0	0.001	0.009	0	0.150	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O13	0	0.003	0.003	0.150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O12	0	0	0	0	0	0	4E–5	0	0	0	0	0	0	0	0	0	0	0	0
O4	0	0	0	0	0	4E–5	0	0.262	0.035	E–04	0.002	0	0	0	0	0	0	0	0
O6	0	0	0	0	0	0	0.262	0	0.010	0	0	0.002	0	0	0	0	0	0	0
O9	0	0	0	0	0	0	0.035	0.009	0	2E–4	0	0.001	0	0	0	0	0	0	0
O7	0	0	0	0	0	0	E–4	0	2E–4	0	0.148	0.086	0	0	0	0	0	0	0
O8	0	0	0	0	0	0	0.002	0	0	0.148	0	0.410	0	0	0	0	0	0	0
O14	0	0	0	0	0	0	0	0.002	0.001	0.086	0.410	0	0	0	0	0	0	0	0
O3	0	0	0	0	0	0	0	0	0	0	0	0	0	0.014	2E–5	0	0	0	0
O2	0	0	0	0	0	0	0	0	0	0	0	0	0.014	0	0.158	0	0	0	0
O11	0	0	0	0	0	0	0	0	0	0	0	0	2E–5	0.158183	0	0.063	0.040	0	0
O1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.063	0	0.660	0	0
O17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.040	0.660	0	0	0
O10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.026
O15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.026	0

Note. O1–O19 stand for the 19 crimes. Values greater than 0 are highlighted.

Source: authors' calculations based on data from Thurstone (1927).

Table 4. Distances between crimes based on Thurstone's data

Type of crime	O19	O16	O5	O18	O13	O12	O4	O6	O9	O7	O8	O14	O3	O2	O11	O1	O17	O10	O15
O19	0.00	1.05	1.09	1.16	1.19	1.40	1.56	1.59	1.65	1.72	1.75	1.77	2.13	2.22	2.32	2.40	2.40	3.33	3.46
O16	1.05	0.00	0.04	0.11	0.13	0.35	0.50	0.54	0.60	0.67	0.70	0.72	1.08	1.16	1.27	1.34	1.35	2.28	2.40
O5	1.09	0.04	0.00	0.07	0.10	0.31	0.47	0.50	0.56	0.63	0.66	0.68	1.04	1.13	1.23	1.31	1.31	2.24	2.37
O18	1.16	0.11	0.07	0.00	0.02	0.23	0.39	0.43	0.49	0.56	0.59	0.61	0.97	1.05	1.16	1.23	1.23	2.17	2.29
O13	1.19	0.13	0.10	0.02	0.00	0.21	0.37	0.41	0.46	0.54	0.56	0.58	0.95	1.03	1.13	1.21	1.21	2.14	2.27
O12	1.40	0.35	0.31	0.23	0.21	0.00	0.16	0.19	0.25	0.33	0.35	0.37	0.74	0.82	0.92	1.00	1.00	1.93	2.06
O4	1.56	0.50	0.47	0.39	0.37	0.16	0.00	0.04	0.09	0.17	0.19	0.21	0.58	0.66	0.76	0.84	0.84	1.77	1.90
O6	1.59	0.54	0.50	0.43	0.41	0.19	0.04	0.00	0.06	0.13	0.16	0.18	0.54	0.63	0.73	0.80	0.81	1.74	1.87
O9	1.65	0.60	0.56	0.49	0.46	0.25	0.09	0.06	0.00	0.08	0.10	0.12	0.48	0.57	0.67	0.75	0.75	1.68	1.81
O7	1.72	0.67	0.63	0.56	0.54	0.33	0.17	0.13	0.08	0.00	0.03	0.04	0.41	0.49	0.60	0.67	0.67	1.61	1.73
O8	1.75	0.70	0.66	0.59	0.56	0.35	0.19	0.16	0.10	0.03	0.00	0.02	0.38	0.47	0.57	0.65	0.65	1.58	1.71
O14	1.77	0.72	0.68	0.61	0.58	0.37	0.21	0.18	0.12	0.04	0.02	0.00	0.36	0.45	0.55	0.63	0.63	1.56	1.69
O3	2.13	1.08	1.04	0.97	0.95	0.74	0.58	0.54	0.48	0.41	0.38	0.36	0.00	0.08	0.19	0.26	0.26	1.20	1.32
O2	2.22	1.16	1.13	1.05	1.03	0.82	0.66	0.63	0.57	0.49	0.47	0.45	0.08	0.00	0.10	0.18	0.18	1.11	1.24
O11	2.32	1.27	1.23	1.16	1.13	0.92	0.76	0.73	0.67	0.60	0.57	0.55	0.19	0.10	0.00	0.08	0.08	1.01	1.14
O1	2.40	1.34	1.31	1.23	1.21	1.00	0.84	0.80	0.75	0.67	0.65	0.63	0.26	0.18	0.08	0.00	0.00	0.93	1.06
O17	2.40	1.35	1.31	1.23	1.21	1.00	0.84	0.81	0.75	0.67	0.65	0.63	0.26	0.18	0.08	0.00	0.00	0.93	1.06
O10	3.33	2.28	2.24	2.17	2.14	1.93	1.77	1.74	1.68	1.61	1.58	1.56	1.20	1.11	1.01	0.93	0.93	0.00	0.13
O15	3.46	2.40	2.37	2.29	2.27	2.06	1.90	1.87	1.81	1.73	1.71	1.69	1.32	1.24	1.14	1.06	1.06	0.13	0.00

Note. O1–O19 stand for the 19 crimes. Values smaller than 0.2 are highlighted.

Source: authors' calculations based on data from Thurstone (1927).

Table 5. Relative frequencies of exchanging order for the Thurstone method applied to 10 crimes from the authors' survey

Type of crime	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
O1	0.000	0.000	0.753	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O3	0.753	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000
O5	0.000	0.000	0.000	0.000	0.000	0.000	0.110	0.000	0.000	0.000
O6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O7	0.000	0.000	0.000	0.000	0.110	0.000	0.000	0.000	0.000	0.000
O8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O9	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000
O10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note. O1–O10 stand for the 10 crimes. Values on the diagonal and greater than 0.1 are highlighted.
Source: authors' calculations based on the data from their own survey.

Table 6. Distances between the 10 crimes, obtained by the Thurstone method applied to data from authors' survey

Type of crime	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
O1	0.00	0.37	0.02	0.82	2.41	0.25	2.47	1.95	1.05	1.59
O2	0.37	0.00	0.35	0.45	2.04	0.63	2.09	1.57	0.68	1.22
O3	0.02	0.35	0.00	0.80	2.39	0.28	2.44	1.92	1.03	1.57
O4	0.82	0.45	0.80	0.00	1.59	1.08	1.64	1.12	0.23	0.77
O5	2.41	2.04	2.39	1.59	0.00	2.67	0.05	0.47	1.36	0.82
O6	0.25	0.63	0.28	1.08	2.67	0.00	2.72	2.20	1.30	1.85
O7	2.47	2.09	2.44	1.64	0.05	2.72	0.00	0.52	1.42	0.87
O8	1.95	1.57	1.92	1.12	0.47	2.20	0.52	0.00	0.90	0.35
O9	1.05	0.68	1.03	0.23	1.36	1.30	1.42	0.90	0.00	0.54
O10	1.59	1.22	1.57	0.77	0.82	1.85	0.87	0.35	0.54	0.00

Note. O1–O10 stand for the 10 crimes. Values that are equal or smaller than 0.05 are highlighted.
Source: authors' calculations based on data from their own survey.

It is obvious that the smaller the distance between the objects, the higher the probability of misidentifying them (i.e. the probability that the order of two objects will be reversed). If the distance between the expected values is x (in terms of standard deviation), the probability of a single person misidentifying them amounts to $p(x) = 1 - F(x)$. For a large number of respondents, n, the probability that the whole group would misidentify the objects can be approximated by

$$1 - F\left(\frac{0.5 - p(x)}{\sqrt{\frac{p(x)\left(1 - p(x)\right)}{n}}}\right).$$

To continue the example from Tables 5 and 6, one may calculate those probabilities for the most problematic pairs of objects, as presented in Table 7.

Table 7. Probabilities of the misidentifications of objects for: a single person, 220 respondents and 1,000 respondents

Specification	O1, O6	O3, O1	O2, O3	O4, O2	O9, O4	O10, O9	O8, O10	O5, O8	O7, O5
Distance	0.253	0.025	0.350	0.449	0.227	0.544	0.353	0.467	0.054
$N = 1$	0.400	0.490	0.363	0.327	0.410	0.293	0.362	0.320	0.479
$N = 220$	0.001	0.385	10^{-5}	10^{-8}	0.003	10^{-12}	10^{-5}	10^{-9}	0.262
$N = 1{,}000$	10^{-11}	0.267	0	0	10^{-9}	0	0	0	0.087

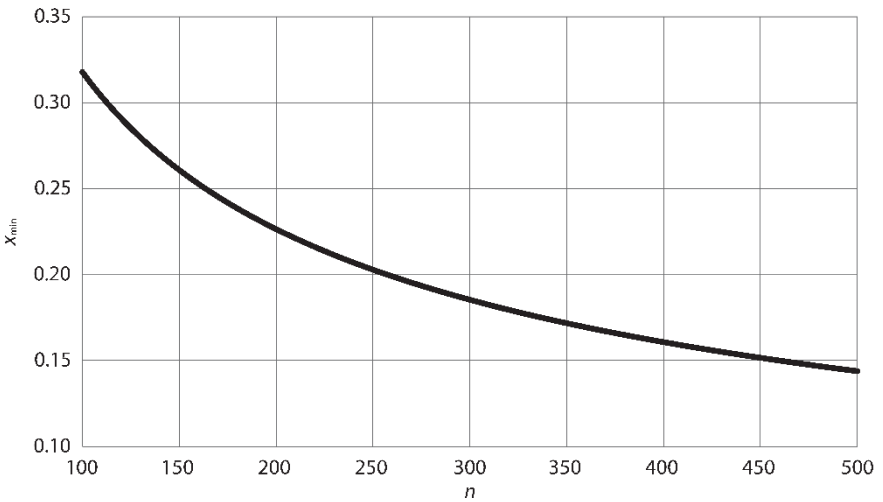
Note. The numbers of objects in the caption of the table correspond to O1–O10 in Tables 5–6.
Source: authors' calculations based on data from their own survey.

In the '$N = 1$' row of Table 7, the probabilities of an object misidentification for a single person are presented, in the '$N = 220$' row, the probabilities that the number of misidentifications will be greater than half of the trials for $N = 220$ (the actual number of respondents in our survey), and in the '$N = 1{,}000$' row, the probabilities of misidentification for a much greater number of respondents, i.e. $N = 1{,}000$. One might observe that for $N = 220$, the probability of a group misidentification of such close objects is very high, namely 39% for the first pair and 26% for the second one. And even if we increase the number of respondents to 1,000, this probability will still be high: 27% and 9%, respectively. For comparison, in the case of $N = 220$, the remaining probabilities were all less than 0.35% (the greatest one), and for $N = 1{,}000$, they were of order 10^{-9} at most.

Thus, if we require relatively small probability of misidentification, e.g. 0.005, we can calculate the limiting distance between the objects below which we could not distinguish between them on the basis of an n-element sample (see Figure 3).

Figure 3 presents the limiting value of a distance between the objects for a given n (between 100 and 500), below which the probability of misidentification will be higher than 0.005. This limiting value coincides to a large degree with the values for which we can observe changes of relative positions depending on the subset. Thus, the probability of misidentification below 0.005 may need the rule of thumb to decide which objects cannot be regard as properly ordered within the scope of the Thurstone method.

Figure 3. The limiting value of the distance between the objects for a given *n*, below which the probability of misidentification will be higher than 0.005



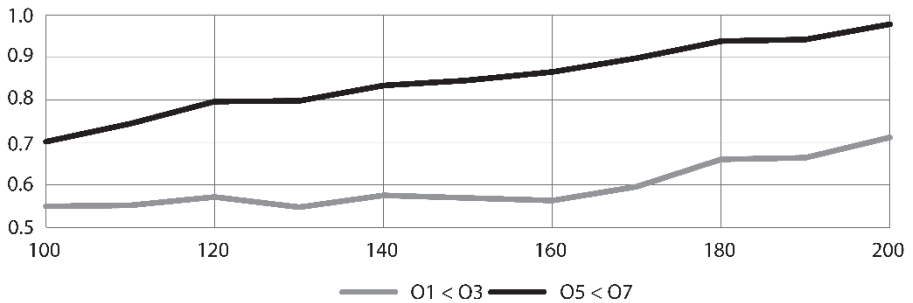
Source: authors' work based on their calculations.

5. Respondents-subsets analysis

In this part of the sensitivity analysis, we checked what the results would be if not all respondents were surveyed. To this end, we chose at random 1,000 different 100-element sets of respondents (each set chosen at random) and found the resulting order of objects. As for some pairs of objects the answers were almost uniform (the relative frequencies approaching 1 or 0), in order to avoid the problem with infinite inverse cumulative distribution function we introduced one 'artificial person', whose answers were always against the general tendency, and who was always included in the set of other respondents, chosen at random. It turned out that the pairs whose order could be changed (as compared to the set of all respondents) were O1, O3 and O5, O7 ones. In 54% of the rankings, O1 was lower than O3, and in 74% pairs, O5 was lower than O7. Thus, it appears that a 100-element sample is not sufficient to fix the order of O1, O3 and O5, O7. However, the remaining objects in all the 1,000 cases had exactly the same order.

One may suggest that a 100-element sample is not large enough. So the question is whether by increasing the sample size (but keeping it below 219, to be able to choose many different samples of a given size), we would obtain more stable results. Figure 4 presents the fractions of orders O1 < O3 and O5 < O7 for sample sizes between 100 to 200.

Figure 4. The fractions of orders O1 < O3 and O5 < O7 for sample sizes from 100 to 200 based on data from authors' survey



Source: authors' work.

What can be seen is that although the pair O5, O7 is more quickly dominated by the order O5 < O7 the order still happens to be reversed for the 200-element sample. As regards the O1, O3 pair, this phenomenon is even more striking.

The conclusion of these two kinds of a 'sensitivity analysis' might be the following. Although we do not necessarily assume that the 'objective' position of objects form a linear order, we should expect that the result of aggregating preferences would give some stable results, not depending on slight changes in the procedure, e.g. adding one or two other objects, or removing a few from the sample. Thus, the relative order of O1, O3 and O5, O7 cannot be regarded as reliable within this research. It might be the case that the expansion of the study to larger groups of respondents would yield more stable results. On the other hand, we cannot be sure that adding some other objects to the analysed set would not destabilise the order of remaining objects (which turned out to be stable in this study). Also, we cannot be sure that expanding the sample by more respondents would cause the results to remain unchanged. However, the detailed sensitivity analysis at least corroborated the presumption that the order of O2, O4, O6, O8, O9 and O10 would remain stable (as was the case in our study).

6. Conclusions

The Thurstone method is often believed to be capable of establishing an interval scale of a set of objects, in contrast to e.g. the average rank method.

However, it is hard to consider the distances between the objects that are smaller than approximately 0.2 of the standard deviation of normal distributions (underlying the Thurstone method) as valid. Most of such pairs of objects are unstable when the set of other objects they are examined with changes. This is called 'dependence on irrelevant alternatives', and obviously it is not a desirable property for any method.

Although we know that in general, none of the methods of aggregating preferences is free of this flaw, in practice some of them are more prone to be dependent on irrelevant alternatives than others. Our empirical study which used Thurstone's original data and the data we obtained by repeating Thurstone's procedure a century later demonstrates that the Thurstone method is prone to the dependency on irrelevant alternatives to a relatively large degree. On the other hand, no empirical study can prove or disprove a general property. Therefore what we suggest is not to relay too much on a method and to check its resilience every time. Pairs of objects which frequently exchange their relative ranks should be thus regarded as undistinguishable for all practical purposes. Treating one of the objects from such a pair as larger solely on the basis of it being larger in a given set of objects depends too much on the arbitrary decision which objects were and which were not included in the study.

Another observation related to the Thurstone method regards the choice of the formula to be minimised. The original formula was chosen by Thurstone in the 1920s most probably due its relative calculational simplicity. However, this historical formula is still used nowadays, which causes problems with high empirical frequencies close to 1. We suggest that researchers nowadays might apply formula (2) to the minimising procedure, which would automatically ensure a better fit of the predicted frequencies to the observed ones.

References

Arrow, K. J. (1951). *Social Choice and Individual Values*. Wiley.

Dębicka, J., Mazurek, E., & Ostasiewicz, K. (2022). Methodological Aspects of Measuring Preferences Using the Rank and Thurstone Scale. *Statistika*, *102*(3), 236–248. https://doi.org/10.54694/stat.2022.5.

Dębicka, J., Mazurek, E., & Ostasiewicz, K. (2023). *Thurstone model – a model or a procedure?* [manuscript submitted for publication].

Dragonetti, R., Ponticorvo, M., Dolce, P., Di Filippo, S., & Mercogliano, F. (2017). Pairwise comparison psychoacoustic test on the noise emitted by DC electrical motors. *Applied Acoustics*, *119*, 108–118. https://doi.org/10.1016/j.apacoust.2016.12.016.

Handley, J. C. (2001). *Comparative Analysis of Bradley-Terry and Thurstone-Mosteller Paired Comparison Models for Image Quality Assessment*. PICS: Image Processing, Image Quality, Image Capture, Systems Conference, Montreal. https://www.imaging.org/common/uploaded%20files/pdfs/Papers/2001/PICS-0-251/4604.pdf.

Heldsinger, S., & Humphry, S. (2010). Using the Method of Pairwise Comparison to Obtain Reliable Teacher Assessments. *The Australian Educational Researcher*, *37*(2), 1–19. https://doi.org/10.1007/BF03216919.

Kanda, T. (2002). Classification of menus on home dining tables based upon human meal Kansei. *Kansei Engineering International*, *3*(4), 9–14. https://doi.org/10.5057/kei.3.4_9.

Krabbe, P. F. (2008). Thurstone Scaling as a Measurement Method to Quantify Subjective Health Outcomes. *Medical Care*, *46*(4), 357–365. https://doi.org/10.1097/MLR.0b013e31815ceca9.

Lipovetsky, S. (2007). Thurstone scaling in order statistics. *Mathematical and Computer Modelling*, *45*(7–8), 917–926. https://doi.org/10.1016/j.m cm.2006.09.009.

Lipovetsky, S., & Conklin, W. M. (2004). Thurstone scaling via binary response regression. *Statistical Methodology*, *1*(1–2), 93–104. https://doi.org/10.1016/j.statmet.2004.04.001.

Mosteller, F. (1951). Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. *Psychometrika*, *16*(2), 207–218. https://doi.org/10.1007/BF02289116.

Orbán-Mihálykó, É., Mihálykó, C., & Gyarmati, L. (2022). Application of the Generalized Thurstone Method for Evaluations of Sports Tournaments' Results. *Knowledge*, *2*(1), 157–166. https://doi.org/10.3390/knowledge2010009.

Temesi, J., Szádoczki, Z., & Bozóki, S. (2023). Incomplete pairwise comparison matrices: Ranking top women tennis players. *Journal of the Operational Research Society*, *75*(1), 145–157. https://doi.org/10.1080/01605682.2023.2180447.

Thurstone, L. L. (1927). The method of paired comparisons for social values. *The Journal of Abnormal and Social Psychology*, *21*(4), 384–400. https://psycnet.apa.org/doi/10.1037/h0065439.

Thurstone, L. L., & Chave, E. J. (1929). *The measurement of attitude*. The University of Chicago Press.

Thurstone, L. L., & Jones, L. V. (1957). The Rational Origin for Measuring Subjective Values. *Journal of the American Statistical Association*, *52*(280), 458–471. https://psycnet.apa.org/doi /10.2307/2281694.

Vojnovic, M., & Yun, S.-Y. (2016). Parameter Estimation for Generalized Thurstone Choice Models. In M. F. Balcan & K. Q. Weinberger (Eds.), *Proceedings of the 33rd International Conference on Machine Learning* (vol. 48, pp. 498–506). JMLR.org.

van Wijk, A., & Hoogstraten, J. (2004). Paired comparisons of sensory pain adjectives. *European Journal of Pain*, *8*(4), 293–297. https://doi.org/10.1016/j.ejpain.2003.10.002.

Woods, R. L., Satgunam, P., Bronstad, M., & Peli, E. (2010). Statistical analysis of subjective preferences for video enhancement. In B. E. Rogowitz & T. N. Pappas (Eds.), *Human Vision and Electronic Imaging XV* (vol. 7527, pp. 98–107). SPIE.

Yellott, Jr., J. I. (1980). Generalized Thurstone models for ranking: equivalence and reversibility. *Journal of Mathematical Psychology*, *22*(1), 48–69. https://doi.org/10.1016/0022-2496(80)90046-2.

Wiadomości Statystyczne. The Polish Statistician, 2024, *69*(11), 18–35 https://doi.org/10.59139/ws.2024.11.2
Statystyka w praktyce / Statistics in practice
Zgłoszony/submitted: 7.03.2023, poprawiony/revised: 7.07.2024 zaakceptowany/accepted: 7.10.2024

Assessment of the dynamics of the diffusion of mobile technologies in the Visegrad Group[1]

Marcin Salamaga[a]

Abstract. The diffusion of innovation involves the spread of new solutions in the field of technology, organisation of production and marketing, as well as knowledge among enterprises-imitators, the society, etc. The aim of the research described in the paper is to assess the volume and rate of the diffusion of mobile technologies in the processing industry and to determine the nature and strength of the impact of investment specialisation (investment attractiveness) and industry export specialisation on this phenomenon in the Visegrad Group (V4) countries as a whole. The occurrence of innovation diffusion and its dynamics were examined using the econometric model of the Gompertz function. An error correction model (ECM) was used to determine the impact of V4 countries' investment and export specialisation on the process of innovation diffusion. The ECM was estimated for selected industrial processing sectors. The study is based on data on foreign trade, innovation and foreign direct investment for the years 2010–2021 obtained from Eurostat and the national statistical institutions of Poland, Czechia, Hungary and Slovakia. The research results show that the diffusion of mobile technologies in the V4 countries occurs in various industries, but its rate varies depending on the degree of the industry's technological advancement and the ownership structure of the enterprise capital.

Keywords: innovation, diffusion of innovation, mobile technologies, Gompertz function, error correction model, ECM, Visegrad Group, V4

JEL: O30, C22

Ocena dynamiki dyfuzji technologii mobilnych w Grupie Wyszehradzkiej

Streszczenie. Dyfuzja innowacji polega na rozprzestrzenianiu się nowych rozwiązań w zakresie technologii, organizacji produkcji i marketingu oraz wiedzy wśród przedsiębiorstw naśladowców, w społeczeństwie itd. Celem badania omawianego w artykule jest ocena wielkości i tempa dyfuzji technologii mobilnych w przemyśle przetwórczym oraz określenie charakteru i siły wpływu specjalizacji inwestycyjnej (atrakcyjności inwestycyjnej) oraz specjalizacji eksportowej na to zjawisko w krajach Grupy Wyszehradzkiej (V4) jako całości. Występowanie zjawiska

[1] Artykuł został opracowany na podstawie referatu wygłoszonego na 28. Polsko-Słowacko-Ukraińskim Seminarium Naukowym „Metody ilościowe w analizach społeczno-gospodarczych – teoria i zastosowania", które odbyło się w dniach 20–21 października 2022 r. w Bratysławie (Słowacja). / The article is based on a paper delivered at 28th Polish-Slovak-Ukrainian Scientific Seminar 'Quantitative Methods in Socio-Economic Analysis – Theory and Applications' held on 20–21 October 2022 in Bratislava, Slovakia.

[a] Uniwersytet Ekonomiczny w Krakowie, Kolegium Ekonomii, Finansów i Prawa, Instytut Metod Ilościowych w Naukach Społecznych, Polska / Krakow University of Economics, College of Economics, Finance and Law, Institute of Quantitative Methods in Social Sciences, Poland.
ORCID: https://orcid.org/0000-0003-0225-6651. E-mail: salamaga@uek.krakow.pl.

dyfuzji innowacji oraz jego dynamikę badano za pomocą funkcji Gompertza. Aby określić wpływ atrakcyjności inwestycyjnej oraz specjalizacji eksportowej krajów V4 na proces dyfuzji innowacji, zastosowano model ekonometryczny z mechanizmem korekty błędem (ECM), który oszacowano dla wybranych sektorów przetwórstwa przemysłowego. W badaniu wykorzystano dane dotyczące handlu zagranicznego, innowacyjności i bezpośrednich inwestycji zagranicznych za lata 2010–2021, zaczerpnięte z Eurostatu i krajowych urzędów statystycznych: Polski, Czech, Słowacji i Węgier. Na podstawie uzyskanych wyników można stwierdzić, że dyfuzja technologii mobilnych w krajach V4 występuje w różnych branżach, ale jej tempo jest zróżnicowane w zależności od stopnia zaawansowania technologicznego branży oraz struktury własnościowej kapitału przedsiębiorstwa.

Słowa kluczowe: innowacje, dyfuzja innowacji, technologie mobilne, funkcja Gompertza, mechanizm korekty błędem, ECM, Grupa Wyszehradzka, V4

1. Introduction

Innovations combined with the knowledge-based economy play a pivotal role in shaping the competitive advantage of the modern economy, both on macro- and microeconomic scale. Digital innovations developed in parallel with mobile technologies play a particularly important role in this respect. Recently, these technologies have become a common solution, an inseparable element of the lives of both individual users and the functioning of entire enterprises. The development of mobile technologies is the effect of numerous inventions and improvements in terms of the portable nature of new devices, their intuitive and simple operation and the possibility of connecting them to the Internet, as well as the low cost of buying and using this type of equipment.

Mobile technologies are a key element of the digital economy and play an important role in increasing the innovation and competitiveness of organisations. Such technologies ensure quick access to information and an improved communication process itself. Mobile technologies enable the implementation of many digital innovations of a social and business nature. A large number of companies widely use or even base their existence on the available mobile technologies. They not only change the way traditional services such as transport and logistics function, but they constitute an invaluable improvement in such areas as production organisation, services provision or relationships with customers. As a result, completely new business models are created, for example within the sharing economy. Creating new mobile technologies is generally expensive and requires a large scientific and research base.

In general, new mobile technologies reach Central and Eastern Europe through innovation diffusion. This term can be understood as the dissemination of innovation/technology through market and non-market channels, starting from their initial implementation anywhere in the world, which later spreads to other countries

and regions and to other markets and companies (Organisation for Economic Co-operation and Development & Eurostat, 2018). Therefore, it is the spread of new technological, organisational, and marketing solutions, as well as of knowledge among enterprises-imitators and copycats (the *spillover effect*). An important feature of technological innovation diffusion is that it spreads at a variable rate which usually changes according to the S-shaped curve (e.g. logistic curve). This means that the rate of diffusion in its initial phase is slow, then it increases more than proportionally and decreases again in the final phase. The logistics curve is the most commonly used tool in the study of the diffusion of technological innovations. In addition to it, other approaches and models are used such as the Bass model (1969), the Rogers model (1983), and the wave and hierarchical model (Morrill & Manninen, 1975). Technology diffusion occurs in both developed and developing economies and is the subject of numerous empirical analyses. While the international literature is abundant in research results on the diffusion of mobile technology based on the aforementioned models, the study of this phenomenon with the use of econometric tools in Poland and other Visegrad Group (V4) countries (Czechia, Hungary and Slovakia) is relatively rare. The existing research gap in this area makes it difficult to accurately assess the actual relationships between spillover processes and selected macroeconomic variables. This article attempts to fill this gap to a certain extent. The main objective of the undertaken analysis is to assess the volume and rate of the diffusion of mobile technologies in the processing industry and to determine the nature and strength of the impact of investment specialisation (investment attractiveness) and industry export specialisation on this phenomenon in the V4. The occurrence of mobile technology diffusion and its rate was studied using the Gompertz function. Dynamic error correction models (ECM) were used to determine the impact of the V4 investment attractiveness and specialisation in foreign trade on diffusion processes. The ECMs were estimated for industrial processing industries at various degrees of technological advancement as well as for companies with and without foreign capital. This research approach has been applied to determine how the degree of technological advancement of the industry and the ownership structure of capital affect the rate of innovation diffusion, and how the intensity of this diffusion is influenced by the revealed innovation advantage and investment attractiveness of the entire V4.

2. Literature review

The phenomenon of innovation diffusion has a long tradition of theoretical considerations, analyses and research. The classic approaches to this issue can be found in the works of Hägerstrand (1967), Jones (1967), Rogers (1962), Tarde (1895),

and others. Researcher interest in this phenomenon results from its key role in supporting technical progress of many economies. Diffusion effects are observed when the majority of potential buyers start using a given product or implement a specific technology.

Innovation diffusion describes a process associated with bringing innovations to market. The process involves the knowledge of psychology, strategy creation, market analysis, marketing and technical achievements. According to Rogers (1983), one of the creators of this concept, diffusion is the process by which innovation is communicated through time-bound channels among members of the social system. This is a special type of communication where the message concerns new ideas. Communication is a process during which the participants create and share information with each other in order to achieve mutual understanding. Issues related to innovation diffusion were also explored by Bass (1969), who argued that diffusion is a communication process in which innovation spontaneously spreads due to external influences (direct information stimuli) prompting innovative behaviour among the adopters and internal sources of information resulting from social interaction. The phenomenon of innovation diffusion has been studied in literature in various aspects. One of them involves analyses related to the profitability of business ventures, the pricing policy, the distribution channel or, in a broader sense, the company's strategy of introducing products to other markets. Subsequently, studies of this type started to look at the diffusion of digital innovation in general (e.g. Akçura & Altınkemer, 2010). Then, as mobile technologies developed and devices such as iPhones and iPads emerged, researchers also began to analyse the diffusion of innovation in this area. Another important aspect of research focusing on innovation diffusion involves the analysis of the spread of innovations by modelling their life cycles (Peres et al., 2010). The results are mathematical models describing innovation growth curves with estimated parameters such as the inflection point and saturation level. The aim of this research is, in general, to provide accurate pictures of diffusion processes over time so that managers can forecast sales, develop appropriate strategies and act accordingly (Mahajan et al., 1990). Researchers have proposed many models based on a variety of assumptions that work well in specific socio-economic conditions (Meade & Islam, 1995). Much of the research on diffusion concerns innovations related to high-tech products such as mobile phones, mobile and digital technologies, the Internet. One of the ground-breaking models dedicated to the study of this type of issues is the Bass model. However, some recent studies suggest that this model fails to work properly in the diffusion of certain high-tech products, while basic growth curve models such as the modified exponential model, the logistic model or the Gompertz model reflect these phenomena quite well (Singh, 2008). There are quite a few examples of modelling the diffusion of mobile and digital innovations

in the recent years. These include Dos Santos and Peffers (1998), who compared the effectiveness of three diffusion models in explaining the spread of e-commerce applications in banking. The diffusion of mobile telephony in the Indian market was studied by Singh (2008). Using logistic curve models and the Gompertz function, the author showed that the latter function better reflects the issue in question.

Some researchers show that the process of innovation diffusion is conditioned by the interdependencies between certain technologies (Bayus, 1987). Wareham et al. (2004) argue that the proliferation of mobile phones and the Internet was strongly linked to people's income and occupations. In addition, the authors pointed out that mobile phone adoption was faster in some ethnic groups than in others. The relationship between the price of an innovation and the income of the society was also studied by Horsky (1990), and Kalish (1983). The authors of many publications argue that the effectiveness of diffusion of innovation depends on a sufficiently large number of entities open to new solutions and accepting them (Mahler & Rogers, 1999).

In the Polish literature, studies on innovation diffusion in various aspects can be found in the works of Firlej and Żmija (2014), Gwarda-Gruszczyńska (2017), Klincewicz (2011) and Wiśniewska (2004). Empirical studies of the process of diffusion of innovation in the world literature have been conducted for years using econometric modelling (Bemmaor & Lee, 2002; Desiraju et al., 2004; Van den Bulte & Stremersch, 2004), but in the Polish literature, the use of such tools in this context is not popular, apart from a few exceptions (Kolarz, 2006). Similarly, the research deficit in this area is also visible in relation to the markets of other Central and Eastern European countries (including the other V4 nations). Meanwhile, econometric modelling is already a well-established canon in the global research stream of innovation diffusion. Models of this type allow for a real assessment of the cause-and-effect relationships between spillover processes and macroeconomic variables, as well as the forecasting of the diffusion process itself (Guidolin, 2023; Zolait, 2013).

Thus, a research niche exists in the area of innovation diffusion calling for in-depth analyses especially of innovations in the dynamically developing mobile technologies, both in Poland and in other countries of Central and Eastern Europe. The research subject of this study is therefore the proliferation of portable electronic devices enabling mobile access to the Internet, with which employers in enterprises of the processing industry equip their employees in the V4 countries. In the modelling of innovation diffusion, the Gompertz function was used and the obtained results then served in the construction of the ECMs. The ECM describes the impact of investment attractiveness and industry export specialisation on the diffusion rate of mobile innovations.

3. Research method

The process of innovation diffusion is in most cases characterised by an initially slow growth rate, followed by a rapid increase, after which a decrease is observed and then the level of innovation finally stabilises (the growth dynamics of innovation diffusion gradually fades). Many studies have confirmed that changes in the innovation diffusion rate follow an S-shaped curve (Desiraju et al., 2004). It is for this reason that research on the dynamics of innovation diffusion is based on models that allow the behaviour of the phenomenon described here to be reproduced. These include, above all, the logistics model and the Gompertz model. This article applies the Gompertz model, which has worked well in similar analyses. Unlike the logistic curve, it is an asymmetric curve, i.e. the rate of change increases at an exponential rate, resulting in an asymmetric sigmoid path around the inflection point. This type of asymmetry is primarily suitable for describing cases in which the maximum growth occurs relatively early (Meade & Islam, 1995). The Gompertz function is represented by the formula:

$$f(t) = Ae^{-exp(-B(t-C))}, \tag{1}$$

where:

A, B, C are the parameters of the Gompertz function: A is the supremum of the values of the function, B sets the displacement of the Gompertz curve along the x-axis, C is the scale parameter,

t is the time variable.

The rate of change in the Gompertz function (Growth Rate of Gompertz function – GRG) is provided by the following formula:

$$GRG = Be^{-B(t-C)}. \tag{2}$$

In the course of the Gompertz function, several ranges can be distinguished, including an area where it has a clearly increasing growth rate and an area of a decreasing growth rate, aiming at the saturation level expressed by asymptote $y = A$. The point separating the rapid growth rate of the curve from the decreasing growth rate is the inflection point of the following coordinates:

$$\frac{dy}{dt}\frac{1}{y}(C) = B. \tag{3}$$

The key objective of this paper is to examine whether and how the competitiveness of foreign trade and the investment attractiveness of the entire V4 affect the dynamics of changes in the diffusion of innovation. The competitiveness of foreign trade was measured using the Revealed Comparative Advantage (RCA) indicator (Balassa, 1965):

$$RCA_i = \frac{Ex_{ij}}{Ex_j} : \frac{Ex_i^R}{Ex^R}, \tag{4}$$

where:
Ex_{ij} is the export value of the i-th industry in the j-th economy,
Ex_j is the total value of exports of the j-th economy,
Ex_i^R is the value of exports of the i-th industry in the reference countries,
Ex^R is the total value of exports in the reference countries.

Indicator (4) is used to assess the RCA of one economy over others in exporting a specific commodity group. The higher the value it takes, the greater the advantage in exports of the analysed economy. Values greater than 1 indicate the existence of a comparative advantage, and values of less than 1 indicate the absence of such an advantage. OECD countries were adopted as a reference group in the calculations of the RCA index, which resulted from their trade and investment relations with the V4 and their technological advancement, thanks to which a seamless transmission of innovation to the V4 is possible.

In a similar way to indicator (4), the indicator of disclosed investment advantage, i.e. Relative Investment Attractivenes (RIA), is defined. It allows for taking into account the investment attractiveness of the economy. This measure is given by the formula below:

$$RIA_i = \frac{FDI_{ij}}{FDI_j} : \frac{FDI_i^R}{FDI^R}, \tag{5}$$

where:
FDI_{ij} is the value of the Foreign Direct Investment (FDI) stocks in the i-th industry in the j-th economy,
FDI_j is the total value of the FDI stocks in the j-th economy,
FDI_i^R is the value of FDI stocks in the i-th sector in the reference countries,
FDI^R is the total FDI stocks in the reference countries.

The higher the investment attractiveness of an economy, the higher the RIA index. Due to the fact that non-OECD countries are also FDI suppliers to the V4, all countries in the world were assumed as reference countries when calculating the RIA index. The relationships between indicators (3), (4) and (5) were analysed using a one-equation model with ECM. The short-term relationship in the ECM model between logarithmic variables is described by the following equation:

$$d\_lnGRG_t = \alpha_0 + \alpha_1 d\_lnRCA_t + \alpha_2 d\_lnRCAI_t + \alpha_3 ECM_{t-1} + \varepsilon_t. \quad (6)$$

The cointegrating equation representing the long-term relationship is:

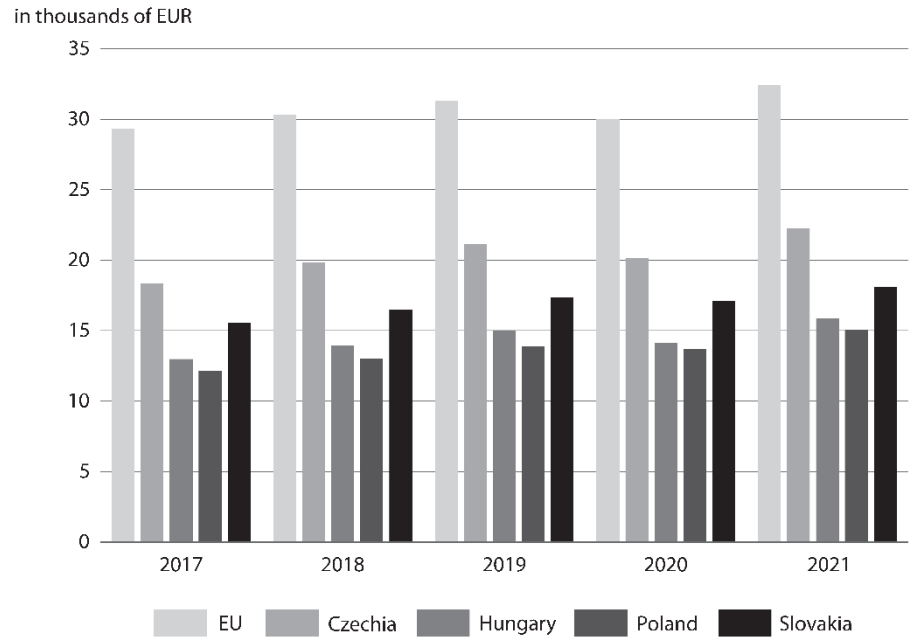
$$lnGRG_t = \gamma_0 + \gamma_1 lnRCA_t + \gamma_2 lnRCAI_t + u_t. \quad (7)$$

Lagged variables were also introduced in case autocorrelation occurred in model (6). Before estimating the model, the internal structure of time series was studied and their order of cointegration and autocorrelation was tested. In the study of time series stationarity, the augmented Dickey-Fuller test (ADF test) was used, while the Engle-Granger test in the time series cointegration study.

4. The economic potential of the V4

Czechia, Hungary, Poland and Slovakia, which form the V4, are characterised not only by geographical proximity but also by having common civilizational, historical, cultural, intellectual and economic roots, which is a good starting point for conducting a comparative analysis between these countries in terms of their achievements. The initial goal of this group was to ensure mutual cooperation in the process of these countries' integration with NATO and the European Union. Having achieved this goal, the V4 continues to deepen collaboration in the economic, cultural, educational and scientific fields. The V4 countries also cooperate in the area of innovation. The economic potential and the development of innovation in these countries show certain differences resulting from many macroeconomic factors, including: the level and dynamics of the GDP, the situation on the labour market, the structure and dynamics of foreign trade, the FDI inflow, and the level of technological development. One of the most important indicators of a country's economic potential is, of course, GDP *per capita*. Figure 1 presents it in the V4 in 2017–2021.

Figure 1. GDP *per capita* in the V4 and the EU



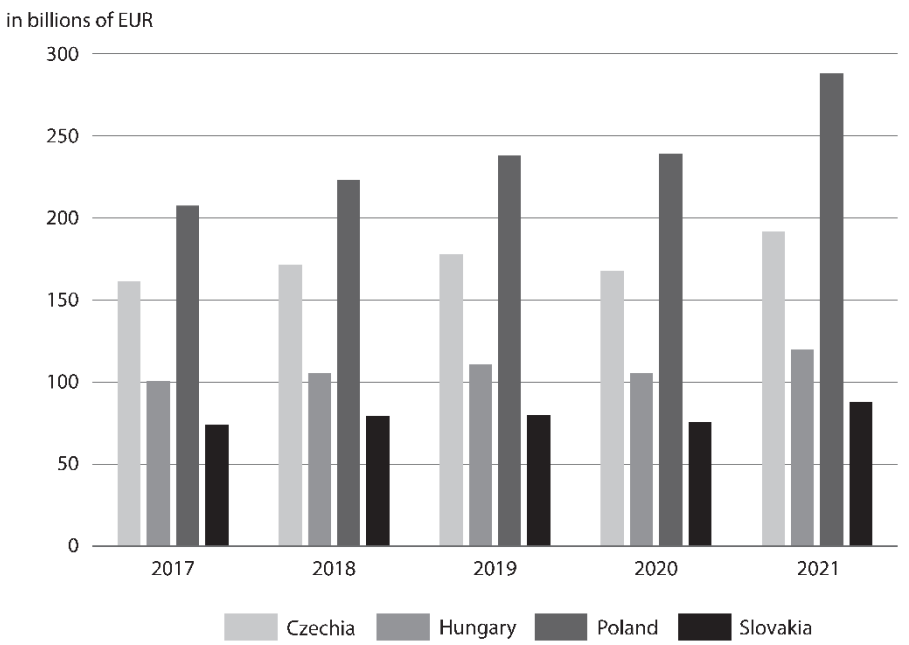
Source: author's work based on data from Eurostat.

According to Figure 1, among all V4 countries, Czechia had the highest level of GDP *per capita* in each of the studied years. The indicator increased from EUR 18,330 in 2017 to EUR 22,270 in 2021, i.e. by approximately 20.5% over the studied period. The second highest level of GDP *per capita* among the compared countries was observed in Slovakia, where the value of goods and services produced in 2021 amounted to EUR 18,100, which is an increase of over 16% compared to 2017. GDP *per capita* of Hungary in the same period was much lower and reached EUR 15,840 (increase of over 22% compared to 2017). Poland had the lowest level of GDP *per capita* in the V4: in 2022, it amounted to EUR 15,060 and was 24.3% higher than in 2017, which, on the other hand, was the highest increase in the value of produced goods and services among the compared countries. The V4 countries are significantly below the EU average which was EUR 29,320 *per capita* in 2017, and EUR 32,430 in 2021.

Figure 2 shows the change in the value of exports of the V4 in 2017–2021. Poland was the leader in exports of goods and services in all the considered years. The average value of this country's exports in 2017–2021 was approximately EUR 239,234 million, i.e. about 45% of the total export among the V4 and 11.6% of all EU export. The second largest exporter among the V4 was Czechia, where the average export value amounted

to approximately EUR 173,909 million, i.e. about 32.7% of the total export of the V4 and about 8.4% of EU export. The third largest exporter in the V4 group was Hungary, and Slovakia was the smallest exporter. The average value of Slovak export amounted to approximately EUR 79,269 million, which constituted about 14.9% of the total exports of the V4 and about 3.8% of EU export. During the analysed period, Poland experienced the fastest growth in export dynamics; its value increased from an average rate of 8.6% year on year, while in Czechia and Hungary the rate was approximately 4.4%, and 4.5% in Slovakia.

Figure 2. The value of export of the V4



Source: author's work based on data from Eurostat.

In the V4, the largest beneficiary of FDI is Poland. The average annual level of FDI resources in Poland in 2017–2021 was approximately USD 989,062 million, which constituted an average of 42% of the total FDI in the V4 countries. Czechia is the second largest beneficiary of FDI resources in the V4, with an average annual amount of about USD 731,385 million (31.1% of the total FDI in the V4). The smallest FDI resources in the examined period, with an average level of approximately USD 244,068 million, were held by Slovakia, whose contribution to the total FDI value of the V4 was 10.36%.

Currently, an important criterion applied to assess the economic potential is the level of the innovativeness of the economy. According to data from 2022 reported by the European Commission (2022), the values of the Summary Innovation Index for the V4 countries were as follows: Czechia – 92.6, Hungary – 69.8, Poland – 60.5 and Slovakia – 64.3. Therefore, the economic potential of the V4 is similar in terms of innovation and other macroeconomic indicators, although it obviously shows differences in various aspects. In this study, the economic potential of the V4 will be treated as the total of the potentials of individual member countries. This approach is justified, because the constructed models of the dependence of the dynamics of innovation diffusion based on selected macroeconomic variables did not show particularly significant differences between the V4 countries.

5. Results

5.1. Analysis of the diffusion rate of mobile technologies in the V4

The study of mobile technology diffusion was based on statistics regarding the number of mobile devices with mobile access to the Internet, such as laptops and smartphones, which companies use to equip their employees. Models of Gompertz functions of mobile technology diffusion were built and the growth rate of these functions (GRG) were calculated on the basis of data for 2010–2021 obtained from Eurostat and the national statistical institutions of the V4. Data on FDI holdings in the V4 (used to calculate the RIA index) were also taken from Eurostat and the databases of national statistical institutions. Export data used to calculate the RCA were extracted from the Comext database available in Eurostat, which contains detailed statistics on the international trade of individual EU member states. Data at a double-digit level of commodity disaggregation based on the Standard International Trade Classification (SITC) were used for the calculations. When selecting the parameters of the Gompertz function, the Gauss-Newton optimisation algorithm was applied, and the calculations were also conducted with a division into enterprises without foreign capital participation and enterprises with such participation.

The estimated Gompertz curve equations for enterprises without and with foreign capital are as follows:

$$\widehat{y_t} = 13931.377 \cdot exp\big(-exp\big(-0.478 \cdot (t - 3.009)\big)\big), \tag{8}$$

$$\widehat{y_t} = 3780.061 \cdot exp\big(-exp\big(-0.494 \cdot (t - 3.900)\big)\big). \tag{9}$$

The detailed results of the estimation of the Gompertz regression for these enterprises are presented in Table 1.

Table 1. Results of parameter estimation of the Gompertz regression function among enterprises without and with foreign capital participation for the V4

Estimation	Enterprises without foreign capital			Enterprises with foreign capital		
	parameter					
	A	B	C	A	B	C
Coefficient	13931.377	0.478	3.009	3780.061	0.494	3.900
Standard error	4551.443	0.236	1.347	685.861	0.165	0.872
p-value	0.018	0.083	0.061	0.001	0.020	0.003
R^2	0.734			0.791		

Source: author's calculations made using the Statistica 13.3 programme.

In the case of all V4 countries, it was possible to adjust the Gompertz curve to the empirical data, which confirms that the process of mobile technology diffusion takes place in the V4, both in enterprises without and with foreign capital. In the comparative analysis of the dynamics of the diffusion rate, it is important to determine the position of the inflection point of the Gompertz curve, as its coordinates reflect the length of the phase of the rapid increase in the diffusion rate and its intensity. Thus, based on the results in Table 1, it can be stated that in enterprises both without and with foreign capital the duration of rapid growth in the first group of companies was about 36 months, and in the second group approximately 47 months.

When analysing the ratings values of parameter B of the Gompertz model, it can be concluded that at the very point of the curve inflection, the diffusion of mobile technology in the sector of enterprises based on own capital had the greatest intensity in companies with foreign capital. It is noteworthy that the phase of rapid growth of innovation diffusion is longer in the sector of enterprises with foreign capital than in enterprises without such capital in all V4 countries.

Subsequently, the modelling of the mobile technology diffusion based on the Gompertz function was carried out among enterprises belonging to industries utilising different levels of technology. Two types of enterprises were taken into account: those operating in low- and medium-low-tech industries and those in high- and medium-tech sectors.

The estimated Gompertz curve equations for enterprises operating in low-tech and medium-low-tech industries, as well as in high-tech and medium-high-tech industries, respectively, are as follows:

$$\widehat{y_t} = 2725.097 \cdot exp\big(-exp\big(-0.563 \cdot (t - 3.037)\big)\big), \tag{10}$$

$$\widehat{y_t} = 4755.317 \cdot exp\big(-exp\big(-0.535 \cdot (t - 3.965)\big)\big). \tag{11}$$

The detailed results of the estimation of the Gompertz regression for these enterprises groups are presented in Table 2.

Table 2. Results of the estimation of Gompertz regression function parameters among enterprises utilising different levels of technology for the V4

Estimation	Low-tech and medium-low-tech enterprises			High-tech and medium-high-tech enterprises		
	parameter					
	A	B	C	A	B	C
Coefficient	2725.097	0.563	3.037	4755.317	0.535	3.965
Standard error	722.361	0.115	0.667	1796.711	0.213	1.864
p-value	0.007	0.002	0.003	0.033	0.041	0.071
R^2	0.697			0.735		

Source: author's calculations made using the Statistica 13.3 programme.

Table 2 shows that the phase of the rapid increase in innovation diffusion calculated on the basis of the Gompertz function was significantly longer in high- and medium-high-tech enterprises than in low- and medium-low-tech ones. The phase of rapid mobile technology diffusion dynamics among companies from the high- and medium-high-tech sectors lasted for about 48 months and about 36 months in low-tech and medium-low-tech ones.

As regards the dynamics of the growth rate of innovation diffusion calculated at the inflection points of the Gompertz function, it was lower in enterprises operating in high- and medium-high-tech industries (0.535) than in enterprises operating in low- and medium-low-tech sectors (0.563).

5.2. Modelling mobile technology diffusion using ECM

GRG values calculated for the individual years of the considered period together with RCA and RIA indicators were used to estimate dynamic econometric models. The variables were first logarithmised and the construction of the econometric models was preceded by the study of the stationarity of the time series of the variables based on the ADF test. As a result, the use of this test was confirmed by the integration of time series at significance level 0.05 in stage I(1). The ADF test, applied to residuals calculated from the corresponding integrating equations (7), showed that their time series are integrated in order I(0). Due to the cointegration of the time series of variables *lnGRG*, *lnRCA* and *lnRCAI* in order I(1), dynamic econometric models for short-term relations containing an ECM were estimated.

The results of the estimation of the parameters of mobile diffusion rate models, which show the long- and short-term relationship for enterprises in low- and medium-low-tech and high- and medium-high-tech industries are presented in Tables 3 and 4.

Table 3. ECM parameters in the V4 by industries utilising different levels of technology

Variable	Low-tech and medium-low-tech sector				High-tech and medium-high-tech sector			
	coefficient	standard error	*t*-Stat	*p*-value	coefficient	standard error	*t*-Stat	*p*-value
const	–2.829	1.160	–2.439	0.071	–0.551	0.259	–2.129	0.100
d_lnGRG_1	1.314	0.227	5.794	0.004	0.926	0.601	1.540	0.198
d_lnRCA	0.570	0.159	3.593	0.023	–0.557	0.387	–1.440	0.223
d_lnRIA	0.652	0.168	3.887	0.018	2.393	0.667	3.586	0.023
ECM_1	–0.148	0.130	–1.139	0.318	–0.058	0.016	–3.554	0.024
R^2	0.710				0.657			

Source: author's calculations made using the Statistica 13.3 programme.

Based on the parameter assessments presented in Table 3, it can be concluded that in the short term, the determinants of mobile technology diffusion in the V4 in low- and medium-low-tech industries are both export specialisation and investment attractiveness. When the rate of diffusion of innovation grows by an average of about 0.57% *ceteris paribus*, the disclosed comparative advantage in exports increases by 1% in the discussed industries. However, a 1% increase in investment advantage entails a growth in the diffusion of innovation rate by an average of about 0.652% *ceteris paribus*. Historical values of the growth rate of innovation diffusion are also a statistically significant stimulant of the innovation growth rate: an increase in the lnGRG_1 indicator implies that the rate of diffusion of innovation rises by an average of about 2.393% *ceteris paribus*. A statistically insignificant impact of the revealed comparative advantage on the GRG index in V4 in the short term in high-tech and medium-high-tech industries may be the consequence of the fact that technologically advanced goods constitute only a small part of the foreign trade of the V4 countries (a dozen or so percent at the most). In addition, having an advantage in foreign trade in high-tech industries can demotivate the local producers to innovate their goods. Only the increase of foreign competitiveness in trade stimulates domestic companies, 'forcing' them to innovate, thus activating the processes of innovation diffusion. In turn, the inflow of FDI to the V4 countries is still growing and has recently achieved significant dynamics despite the worldwide economic crisis caused by the COVID-19 pandemic. The V4 countries have effectively taken advantage of the global trend of shortening supply chains due to the pandemic and have become attractive locations to foreign investors.

The assessments of the parameters of the error correction mechanism in both models are negative, which indicates an adjustment of short-term changes to the long-term equilibrium in the entire V4.

Table 4. The parameters of cointegration equations for enterprises operating in industries utilising different levels of technology

Variable	Low-tech and medium-low-tech sector				High-tech and medium-high-tech sector			
	coefficient	standard error	*t*-Stat	*p*-value	coefficient	standard error	*t*-Stat	*p*-value
const	–5.931	1.122	–5.288	0.001	–2.114	1.141	–1.853	0.101
lnRCA	3.502	1.476	2.373	0.045	2.454	1.007	2.436	0.041
lnRIA	3.402	0.747	4.554	0.002	3.396	1.033	3.289	0.011
R^2	0.714				0.607			

Source: author's calculations made using the Statistica 13.3 programme.

As regards the long-term relationship in low and medium-low technology industries, the positive impact of RCA and RIA indicators on the rate of change of innovation diffusion as measured by the GRG indicator is also confirmed. The impact of the investment attractiveness index in high-tech and medium-high-tech sectors is stronger than that of the export competitiveness index. The situation is different in low-tech and medium-low-tech industries, where the revealed comparative advantage in exports has a stronger impact on the rate of innovation diffusion. A 1% increase in the RCA indicator in low-tech and medium-low-tech industries causes the rate of mobile technology diffusion to grow by an average of approximately 3.502% *ceteris paribus*, while the same increase in the indicator of the revealed comparative advantage in exports entails a growth in the rate of innovation diffusion by an average of approximately 3.402% *ceteris paribus*.

The strongest long-term impact of export advantage is observed in high-tech and medium-high-tech industries. A 1% increase in the RCA indicator in high-tech and medium-high-tech industries causes the rate of mobile technology diffusion to rise by an average of approximately 2.454% *ceteris paribus*. On the other hand, an increase in the RIA index by 1% in these industries results in the growth of the GRG index by an average of about 3.396% *ceteris paribus*.

Thus, it can be argued that in long-term investment in high-tech or medium-high-tech industries abroad is a more effective transmission channel for innovation diffusion than foreign trade. The situation is slightly different in low-tech and medium-low-tech industries: investment attractiveness is generally a weaker factor supporting mobile technology diffusion than competitiveness in exports.

It can be concluded that industries with low and medium-low technologies currently ensure better conditions for the absorption of mobile technology compared

to high and medium-high technology industries. One of the reasons for this situation may be that the structure of the V4 economies is still characterised by a relatively small share of industries based on very advanced technologies.

6. Conclusions

The undertaken research shows that the diffusion of mobile technologies in the V4 takes place in different industries, but its rate varies depending on the degree of technological advancement of the industry. A factor differentiating the length and rate of innovation diffusion is also the ownership structure of the company's capital. In enterprises with foreign capital, the diffusion of mobile technologies generally lasts longer than in enterprises based on domestic capital, but the conclusions about the intensity of innovation diffusion in enterprises without and with foreign capital are not quite clear.

Research has also shown that the phases when the rates of innovation diffusion are on the rise are clearly longer in high and medium-high technology industries than in their low and medium-low counterparts. A lower rate of the diffusion of mobile technology in high- and medium-high-tech industries (compared to low- and medium-low-tech industries) may indicate the existence of barriers to the diffusion process in some of those industries or the possibility of developing innovation through channels other than diffusion. The study shows that in industries with low and medium-low technology, both export specialisation and investment attractiveness support the diffusion of mobile technology in the V4. Therefore, strengthening the competitive position of companies on foreign markets as well as introducing support systems for foreign investors on the domestic market are conducive to the diffusion of innovation. In the short term in high- and medium-high-tech industries, the comparative advantage in exports is a barrier to diffusion of innovation. Thus, in these industries, the weakening position of domestic enterprises on foreign markets and their displacement by external competition may be a motivating factor for strengthening the innovation process in enterprises. The obtained results indicate the need to develop industries with high technologies and sectors with high intensity of knowledge, whose share in the economy is still giving way to less technologically-advanced industries.

Acknowledgements

The article presents the results of Project no 091/EIT/2024/POT, financed from a subsidy granted to the Krakow University of Economics.

References

Akçura, M. T., & Altınkemer, K. (2010). Digital bundling. *Information Systems and e-Business Management*, *8*(4), 337–355. https://dx.doi.org/10.1007/s10257-009-0117-5.

Balassa, B. (1965). Trade Liberalisation and "Revealed" Comparative Advantage. *The Manchester School*, *33*(2), 99–123. https://doi.org/10.1111/j.1467-9957.1965.tb00050.x.

Bass, F. M. (1969). A New Product Growth Model for Consumer Durables. *Management Science*, *15*(5), 215–227.

Bayus, B. L. (1987). Forecasting sales of new contingent products: An application to the compact disc markets. *Journal of Product Innovation Management*, *4*(4), 243–255. https://doi.org/10.1111/1540-5885.440243.

Bemmaor, A. C., & Lee, J. (2002). The Impact of Heterogeneity and Ill-Conditioning on Diffusion Model Parameter Estimates. *Marketing Science*, *21*(2), 209–220.

Desiraju, R., Nair, H., & Chintagunta, P. (2004). Diffusion of new pharmaceutical drugs in developing and developed nations. *International Journal of Research in Marketing*, *21*(4), 341–357. https://doi.org/10.1016/j.ijresmar.2004.05.001.

Dos Santos, B. L., & Peffers, K. (1998). Competitor and vendor influence on the adoption of innovative applications in electronic commerce. *Information & Management*, *34*(3), 175–184. https://doi.org/10.1016/S0378-7206(98)00053-6.

European Commission. (2022). *European Innovation Scoreboard 2022*. Publications Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/f0e0330d-534f-11ed-92ed-01aa75ed71a1/language-en.

Firlej, K., & Żmija, D. (2014). *Knowledge transfer and diffusion of innovation as a source of competitiveness of food industry enterprises in Poland*. Fundacja Uniwersytetu Ekonomicznego w Krakowie.

Guidolin, M. (2023). *Innovation Diffusion Models. Theory and Practice*. John Wiley & Sons.

Gwarda-Gruszczyńska, E. (2017). Dyfuzja innowacji – następstwo komercjalizacji nowych technologii. *Organizacja i Kierowanie. Organization and Management*, (2), 383–396. https://econjournals.sgh.waw.pl/OiK/issue/view/723.

Hägerstrand, T. (1967). *Innovation Diffusion as a Spatial Process*. University of Chicago Press.

Horsky, D. (1990). A diffusion model incorporating product benefits, price, income and information. *Marketing Science*, *9*(4), 342–365. https://doi.org/10.1287/mksc.9.4.342.

Jones, G. E. (1967). The Adoption and Diffusion of Agricultural Practices. *World Agricultural Economic and Rural Sociology*, *9*(3), 1–29.

Kalish, S. (1983). Monopolist Pricing with Dynamic Demand and Production Cost. *Marketing Science*, *2*(2), 135–159. https://doi.org/10.1287/mksc.2.2.135.

Klincewicz, K. (2011). *Dyfuzja innowacji. Jak odnieść sukces w komercjalizacji nowych produktów i usług*. Wydawnictwo Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego.

Kolarz, M. (2006). *Wpływ zagranicznych inwestycji bezpośrednich na innowacyjność przedsiębiorstw w Polsce*. Wydawnictwo Uniwersytetu Śląskiego.

Mahajan, V., Muller, E., & Bass, F. M. (1990). New product diffusion models in marketing: a review and directions for research. *Journal of Marketing*, *54*(1), 1–26. https://doi.org/10.2307/1252170.

Mahler, A., & Rogers, E. M. (1999). The diffusion of interactive communication innovations and the critical mass: The adoption of telecommunications services by German banks. *Telecommunications Policy*, *23*(10–11), 719–740. https://doi.org/10.1016/S0308-5961(99)00052-X.

Meade, N., & Islam, T. (1995). Forecasting with growth curves: An empirical comparison. *International Journal of Forecasting*, *11*(2), 199–215. https://doi.org/10.1016/0169-2070(94)00556-R.

Morrill, R. L., & Manninen, D. (1975). Critical Parameters of Spatial Diffusion Processes. *Economic Geography*, *51*(3), 269–277. https://doi.org/10.2307/143121.

Organisation for Economic Co-operation and Development, & Eurostat. (2018). *Guidelines for collecting and interpreting innovation data* (3rd edition). OECD Publishing.

Peres, R., Muller, E., & Mahajan, V. (2010). Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing*, *27*(2), 91–106. https://doi.org/10.1016/j.ijresmar.2009.12.012.

Rogers, E. M. (1962). *Diffusion of Innovations*. The Free Press.

Rogers, E. M. (1983). *Diffusion of Innovations* (3rd edition). The Free Press.

Singh, S. K. (2008). The diffusion of mobile phones in India. *Telecommunications Policy*, *32*(9–10), 642–651. https://doi.org/10.1016/j.telpol.2008.07.005.

Tarde, G. (1895). *Les lois de l'imitation*. Kimé Éditeur.

Van den Bulte, C., & Stremersch, S. (2004). Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test. *Marketing Science*, *23*(4), 530–544. https://doi.org/10.1287/mksc.1040.0054.

Wareham, J., Levy, A., & Shi, W. (2004). Wireless diffusion and mobile computing: Implications for the digital divide. *Telecommunications Policy*, *28*(5–6), 439–457. https://doi.org/10.1016/j.telpol.2003.11.005.

Wiśniewska, J. (2004). *Ekonomiczne determinanty dyfuzji innowacji produktowych i technologicznych w banku komercyjnym*. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.

Zolait, A. H. S. (Ed.). (2013). *Technology Diffusion and Adoption: Global Complexity. Global Innovation*. IGI Global. https://doi.org/10.4018/978-1-4666-2791-8.

Wiadomości Statystyczne. The Polish Statistician, 2024, *69*(11), 36–53 https://doi.org/10.59139/ws.2024.11.3
Statystyka w praktyce / Statistics in practice
Zgłoszony/submitted: 13.07.2023, poprawiony/revised: 9.11.2023, zaakceptowany/accepted: 24.04.2024

Ocena sytuacji ekonomiczno-finansowej przemysłu spożywczego z wykorzystaniem metody TOPSIS

Jadwiga Drożdż[a]

Streszczenie. Sytuacja ekonomiczno-finansowa przemysłu spożywczego to złożone zjawisko, którego nie można zmierzyć za pomocą jednego wskaźnika finansowego. Jego ocena wymaga uwzględnienia wielu aspektów, co jest możliwe dzięki konstruowaniu mierników syntetycznych. Celem badania omawianego w artykule jest ocena sytuacji ekonomiczno-finansowej przemysłu spożywczego w Polsce w ujęciu międzybranżowym i dynamicznym. Do porównań i obliczeń wykorzystano publikowane i niepublikowane dane Głównego Urzędu Statystycznego. Przeanalizowano wyniki osiągnięte przez producentów żywności i napojów w 2022 r. i porównano je z danymi za 2019 r. Następnie wyliczono wartości miernika syntetycznego skonstruowanego za pomocą metody TOPSIS, czyli będącego funkcją odległości euklidesowych zarówno od wzorca, jak i antywzorca rozwoju.

Na podstawie uzyskanych wyników można stwierdzić, że w 2022 r. sytuacja ekonomiczno-finansowa poszczególnych branż przemysłu spożywczego była dobra. W stosunku do 2019 r. nastąpiły tylko niewielkie zmiany. Świadczy to o dużej odporności sektora na niekorzystne warunki makroekonomiczne i o jego zdolności dostosowywania się do zmiennego i nieprzewidywalnego otoczenia rynkowego. Rezultaty badań potwierdzają również utrzymywanie się znacznego międzybranżowego zróżnicowania sytuacji ekonomiczno-finansowej.

Słowa kluczowe: przemysł spożywczy, sytuacja ekonomiczno-finansowa, metoda TOPSIS, wskaźniki ekonomiczno-finansowe

JEL: O12, O14, Q13

Assessment of the economic and financial situation of the food industry by means of the TOPSIS method

Abstract. The economic and financial situation of the food industry is a complex phenomenon that cannot be measured by a single financial indicator. Its assessment requires the consideration of multiple aspects, which is made possible through the construction of synthetic indicators. The aim of the study discussed in the article is to assess the economic and financial situation of the food industry in Poland from an intersectoral and dynamic perspective. Published and unpublished data provided by Statistics Poland were used to make the comparisons and calculations. Data reflecting the performance of food and beverage manufacturers in 2022 were analysed and compared to those from 2019. Subsequently, the values of the synthetic indicator were calculated. The synthetic

[a] Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy, Zakład Ekonomiki Agrobiznesu i Biogospodarki, Polska / The Institute of Agricultural and Food Economics – National Research Institute, Department of Agribusiness and Bioeconomy, Poland.
ORCID: https://orcid.org/0000-0003-3096-3177. E-mail: jadwiga.drozdz@ierigz.waw.pl.

indicator was constructed using the TOPSIS method, which is based on Euclidean distances from both the ideal and anti-ideal development patterns.

The results indicate that the economic and financial situation of individual branches of the food industry in 2022 was favourable. Compared to 2019, only minor changes occurred. This reflects the sector's high resistance to adverse macroeconomic conditions and its adaptability to operate in a volatile and unpredictable market environment. The findings also confirm the persistence of significant intersectoral variation in the economic and financial situation.

Keywords: food industry, economic and financial situation, TOPSIS method, economic and financial indicators

1. Wprowadzenie

Przemysł spożywczy jest jednym z kluczowych ogniw zrównoważonego łańcucha żywnościowego. Pełni funkcję integratora procesu produkcji żywności, łącząc pozostałe ogniwa, w szczególności rolnictwo oraz dystrybucję i konsumpcję produktów żywnościowych. Ponadto determinuje wzrost zapotrzebowania na surowce rolne, rozwija rynek zbytu produktów przetworzonych i wpływa na sferę konsumpcji żywności. Kształtuje również modele spożycia żywności i odpowiada na potrzeby odbiorców. A ponieważ produkty żywnościowe są niezbędne dla konsumentów (Szczepaniak i Firlej, 2015), zapewnienie społeczeństwu dobrej jakościowo i bezpiecznej żywności jest priorytetem.

Przemysł spożywczy w Polsce to ważny, szybko rozwijający się dział gospodarki narodowej, mający duże znaczenie gospodarcze, społeczne i środowiskowe (Szczepaniak i in., 2022). Stanowi również najważniejszą część sektora agrobiznesu, ponieważ jego zadaniem jest zabezpieczanie nietrwałych surowców roślinnych i zwierzęcych oraz przetwarzanie ich w trwałe i bezpieczne produkty nadające się do spożycia (Firlej, 2008). Znaczenie przemysłu spożywczego potwierdzają następujące dane:

- realna wartość produkcji globalnej przemysłu spożywczego w 2021 r. stanowiła ponad 6% krajowej produkcji globalnej oraz 15,8% globalnej produkcji całego przemysłu (Główny Urząd Statystyczny [GUS], 2022), a w latach 2010–2021 zwiększyła się o blisko 50%;
- realna wartość produkcji sprzedanej tego sektora stanowiła 18,5% produkcji sprzedanej przetwórstwa przemysłowego (GUS, 2022) i w latach 2010–2021 zwiększyła się o blisko 53%;
- wartość dodana brutto (WDB) przemysłu spożywczego to 10,1% WDB przemysłu ogółem i 13,3% WDB przetwórstwa przemysłowego (GUS, 2022); w latach 2010–2021 wzrosła realnie o 17,9%;
- przemysł spożywczy ma duży wkład w wytwarzanie produktu krajowego brutto (PKB), mierzonego za pomocą WDB; WDB wytworzona w tym sektorze stanowiła od 2,2% do 3% WDB wytworzonej w całej gospodarce narodowej (GUS, 2022);

- w 2021 r. w sektorze spożywczym było zatrudnionych 458,5 tys. osób, co stanowiło 14,1% zatrudnionych w przemyśle ogółem oraz 16,2% – w przetwórstwie przemysłowym (GUS, 2022);
- polski przemysł spożywczy już w pierwszych latach członkostwa w Unii Europejskiej był konkurencyjny wobec przemysłu spożywczego w większości krajów członkowskich, a jego pozycja na rynkach międzynarodowym była znacząca; w 2012 r. zajmował szóstą pozycję (z udziałem ok. 9%, w cenach porównywalnych) pod względem wartości produkcji sprzedanej i ósmą (z udziałem ok. 5%) w eksporcie produktów spożywczych (Wijnands i Verhoog, 2016); obecnie, po opuszczeniu Jednolitego Rynku Europejskiego przez Wielką Brytanię, plasuje się odpowiednio na pozycji piątej (z udziałem 10,5% w produkcji sprzedanej w 2020 r.) oraz siódmej (6,4% w eksporcie w 2021 r.; Szczepaniak i in., 2022);
- polscy producenci żywności są konkurencyjni na rynkach zagranicznych, na których lokują nadwyżkę produkcji nad popytem krajowym, o czym świadczy nominalny wzrost 2,8 razy wolumenu eksportu artykułów żywnościowych w latach 2010–2021 oraz przekraczający 40% udział tego eksportu w produkcji sprzedanej;
- wymiana handlowa produktami przemysłu spożywczego korzystnie oddziałuje na poprawę krajowego bilansu handlowego, generując wysoką nadwyżkę eksportu nad importem; w latach 2010–2021 zwiększyła się ona prawie trzyipółkrotnie (do 14,2 mld euro), a w 2022 r. wyniosła ponad 17 mld euro.

Przytoczone dane świadczą o znaczeniu i sile polskiego przemysłu spożywczego. W tym kontekście ważna jest więc kondycja ekonomiczno-finansowa jego podmiotów, tj. zdolność do generowania zysków, regulowania zobowiązań, obsługi zadłużenia, a także ekonomiczna efektywność wykorzystania czynników produkcji. Kowalczyk i Kusak (2006) definiują *sytuację ekonomiczno-finansową* przedsiębiorstwa jako jego stan finansowy wyrażający wypłacalność i zdolność generowania zysków oraz powiększania zasobów majątkowych i kapitałowych, czyli jako zjawisko wielowymiarowe, którego nie można opisać za pomocą jednego miernika. Pełny obraz sytuacji ekonomiczno-finansowej wymaga uwzględnienia wielu aspektów, ponieważ część wskaźników może świadczyć o dobrej sytuacji, podczas gdy inne mogą sygnalizować problemy finansowe. Dlatego do oceny kondycji finansowej sektora, w tym poszczególnych jego branż, wskazane jest wykorzystywanie różnych wskaźników finansowych. Użyteczny może być miernik syntetyczny skonstruowany za pomocą metody TOPSIS (ang. *technique for order preference by similarity to an ideal solution*; Hwang i Yoon, 1981), która polega na obliczeniu odległości euklidesowych ocenianego obiektu zarówno od wzorca, jak i antywzorca rozwoju, inaczej niż w przypadku metody Hellwiga, która uwzględnia tylko odległości od wzorca rozwoju (Wysocki, 2010). W literaturze przedmiotu opisano wiele przykładów zastosowania metod wielokryterialnych do oceny sytuacji ekonomiczno-finansowej przemysłu spożywczego lub jego branż.

Celem badania omawianego w artykule jest ocena sytuacji ekonomiczno-finansowej przemysłu spożywczego w Polsce w ujęciu międzybranżowym i dynamicznym.

2. Metoda badania

W badaniu wykorzystano publikowane i niepublikowane dane statystyczne GUS za 2022 r. i porównawczo za 2019 r., pochodzące z przedsiębiorstw przemysłowych zatrudniających ponad dziewięć osób, które złożyły sprawozdania finansowe F-01/I-01. Informacje te dotyczyły danych finansowych w układzie branżowym według Polskiej Klasyfikacji Działalności (PKD) 2007, zgodnym z NACE 2 Rev. 2 (Nomenclature statistique des activités économiques dans la Communauté européenne). Przyjęty podział umożliwił przeprowadzenie badań we wszystkich 16 branżach przemysłu spożywczego[1]: 11 z zakresu produkcji artykułów spożywczych, 4 – produkcji napojów oraz 1 – produkcji wyrobów tytoniowych. Oceniano stan w 2022 r. w odniesieniu do stanu w 2019 r., czyli w okresie przed wystąpieniem zjawisk kryzysowych (pandemii COVID-19 i wojny w Ukrainie).

Do oceny sytuacji ekonomiczno-finansowej wybranych branż przemysłu spożywczego zastosowano klasyczne narzędzia analizy wskaźnikowej: rentowność, płynność finansową, zadłużenie i produktywność. Agregacji cech prostych dokonano za pomocą metody TOPSIS. Następnie przeprowadzono porządkowanie liniowe badanych branż według zagregowanej cechy.

Miernik syntetyczny wyznaczono w kilku etapach. Najpierw na podstawie przesłanek merytorycznych wybrano cechy opisujące sytuację ekonomiczno-finansową, tak aby dotyczyły one obszarów: rentowności, płynności, zadłużenia, sprawności działania i wydajności podstawowych czynników produkcji. Spośród wskaźników określających każdy z obszarów wyeliminowano – wykorzystując analizę macierzy odwrotnej do macierzy korelacji Pearsona – te, które były ze sobą silnie skorelowane, tj. cechy, dla których odpowiadające im elementy macierzy odwrotnej do macierzy korelacji były znacznie większe od 1 (ponieważ ta wartość świadczy o silnym powiązaniu z innymi zmiennymi). Ostatecznie do sklasyfikowania branż przemysłu spożywczego wykorzystano dziewięć cech prostych, w tym trzy z zakresu rentowności, dwie – płynności finansowej, dwie – zadłużenia i dwie – produktywności czynników produkcji. Wartości

[1] Klasyfikacja działalności gospodarczej PKD 2007 jest zgodna z systemem obowiązującym w UE (NACE Rev.2) i grupuje przedsiębiorstwa przemysłu spożywczego w 33 klasach (25 w produkcji artykułów spożywczych, 7 w produkcji napojów i 1 w produkcji wyrobów tytoniowych). Badanie przeprowadzono w poszczególnych branżach sektora spożywczego, które ustalono na poziomie grup lub klas, a także pogrupowanych klas, tak jak w przypadku branż: piwowarskiej (do produkcji piwa dołączono produkcję słodu), młynarskiej (obejmującej przemiał zbóż wraz z produkcją skrobi i makaronu), winiarskiej (produkcja win gronowych, owocowych i wermutów), cukierniczej (produkcja trwałego pieczywa cukierniczego oraz czekolady i wyrobów cukierniczych) oraz koncentratów spożywczych (przetwórstwo herbaty i kawy oraz produkcja przypraw, wytwarzanie gotowych posiłków i dań, produkcja żywności homogenizowanej i dietetycznej).

tych cech obliczono według formuł podanych w zestawieniu i zweryfikowano ich zmienność. Dla żadnej z badanych cech współczynnik zmienności nie był niższy niż 10%, co eliminowałoby daną cechę ze względu na niską zdolność różnicowania badanych branż.

Zestawienie. Wskaźniki zastosowane do oceny sytuacji ekonomiczno-finansowej wyodrębnionych branż przemysłu spożywczego i ich wagi

Kategoria	Wskaźniki	Formuła	Waga
Rentowność w %	netto (RN)	(zysk netto × 100) / (przychody netto)	0,163
	aktywów (ROA)	(zysk netto × 100) / (aktywa ogółem)	0,072
	kapitału własnego (ROE)	(zysk netto × 100) /(kapitał własny)	0,112
Płynność finansowa	bieżąca (QR)	(aktywa bieżące) / (zobowiązania krótkoterminowe)	0,163
	cykl konwersji gotówki (CKG) w dniach	(cykl zapasów + cykl należności) – (cykl zobowiązań)	0,168
Zadłużenie w %	ogółem (ZO)	(zobowiązania i rezerwy na zobowiązania) / (aktywa ogółem)	0,104
	długoterminowe (ZD)	(zobowiązania długoterminowe) / (kapitał własny)	0,106
Produktywność	pracy (PP) w tys. zł/pracownika	(wartość dodana brutto) / (liczba zatrudnionych)	0,056
	środków trwałych (PŚT) w zł	(przychody w cenach bazowych) /(środki trwałe)	0,056

Źródło: opracowanie własne na podstawie: Sierpińska i Jachna (2004).

Określono ważność każdej cechy (wskaźnika ekonomiczno-finansowego) poprzez przypisanie wag za pomocą metody AHP (ang. *analytical hierarchy process*), która bazuje na koncepcji hierarchii celów i tworzeniu porównań dwójkowych między celami tego samego poziomu. Analizę ograniczono do wyrażenia preferencji globalnych – badanych wskaźników ekonomiczno-finansowych wobec siebie – pomijając lokalne, które mogłyby dotyczyć różnych wariantów w zakresie każdego wskaźnika. Dla przykładu, porównując rentowność netto z pozostałymi wskaźnikami, ustalono, że jest ona równorzędna z rentownością kapitału własnego i bieżącą płynnością finansową, w stopniu drugim dominuje nad rentownością aktywów, w stopniu trzecim – nad zadłużeniem ogółem, zadłużeniem długoterminowym, wydajnością pracy oraz produktywnością środków trwałych, ale nad rentownością netto dominuje w stopniu drugim cykl konwersji gotówki. Do oceny znaczenia poszczególnych kryteriów wykorzystano skalę Saaty'ego (Saaty i Vargas, 2006), w której przyjęto następujące znaczenie poszczególnych wartości:

- 1 – wskaźniki są równe i ważne;
- 3 – jeden ze wskaźników jest nieco ważniejszy niż inny;

- 5 – jeden ze wskaźników jest ważniejszy niż inny;
- 7 – jeden ze wskaźników jest znacznie ważniejszy niż inny;
- 9 – jeden ze wskaźników jest nieporównanie ważniejszy niż inny;
- 2, 4, 6, 8 – wartości pośrednie.

Dla każdej cechy prostej wybranej do wyznaczenia miernika syntetycznego otrzymane w ten sposób wagi zweryfikowano poprzez sprawdzenie, czy ocena preferencji przy ich użyciu została przeprowadzona poprawnie. W tym celu obliczono wskaźnik zgodności (CR), który określa, w jakim stopniu wzajemne porównania kryteriów parami są zgodne:

$$CR = \frac{CI}{R},$$

gdzie:
CI – współczynnik rozbieżności,
R – współczynnik zgodności losowych.

Wartość R określonego dla n-liczby kryteriów przyjęto według tablicy wygenerowanej przez Saaty'ego na podstawie symulacji dla 500 tys. macierzy. W przypadku omawianego badania, tj. dla dziewięciu badanych kryteriów, wyniósł on 1,45 (Politechnika Białostocka, 2022). Wartość współczynnika zgodności powinna być mniejsza od 0,1. Po sprawdzeniu zgodności, w dalszej części analizy wykorzystano obliczone w ten sposób wagi globalne.

Z wartości K cech (wskaźników ekonomiczno-finansowych) dla N jednostek statystycznych (branż) zestawiono (N x K)-wymiarową macierz danych:

$$[X] = \begin{vmatrix} x_{11} & x_{12} & \dots & x_{1K} \\ x_{21} & x_{22} & \dots & x_{2K} \\ & & & \\ \dots & \dots & \dots & \dots \\ x_{N1} & x_{N2} & & x_{NK} \end{vmatrix},$$

gdzie x_{ij} $(i = 1, \dots, N,\ j = 1, \dots, K)$ to wartość i-tej cechy w j-ej jednostce statystycznej.

Aby uwolnić wybrane cechy od miana i ujednolicić je w zakresie liczb, dokonano normalizacji cech za pomocą przekształcenia ilorazowego (Wójciak, b.r.; Wysocki i Lira, 2007), w którym punktem odniesienia są parametry: max, min, $\bar{x}_j$, $\sum_{i=1}^{N} x_{ij}$. W omawianym badaniu zastosowano przekształcenie ilorazowe macierzy danych według następującej formuły:

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{N} x_{ij}^2}},$$

gdzie:
i – branża przemysłu spożywczego ($i = 1, 2, \ldots, N$),
j – cecha prosta ($j = 1, 2, \ldots, K$).

Wybór tej metody normalizacji cech prostych wynika z charakteru badanych zmiennych i tego, że są one mierzalne.

Następnie macierz znormalizowanych wartości cech prostych opisujących sytuację ekonomiczno-finansową dla poszczególnych branż przemysłu spożywczego przemnożono przez współczynniki wagowe, tworząc w ten sposób macierz znormalizowaną ważoną.

Ustalono wektor wartości rozwiązania idealnego (wzorca) $\boldsymbol{z}^+$ i antyidealnego (antywzorca) $\boldsymbol{z}^-$:
dla wzorca rozwoju

$$\boldsymbol{z}^+ = \left(\max_i(z_{i1}), \max_i(z_{i2}), \ldots, \max_i(z_{iK})\right) = (z_1^+, z_2^+, \ldots, z_K^+)$$

i antywzorca rozwoju

$$\boldsymbol{z}^- = \left(\min_i(z_{i1}), \min_i(z_{i2}), \ldots, \min_i(z_{iK})\right) = (z_1^-, z_2^-, \ldots, z_K^-).$$

Po ustaleniu wzorca i antywzorca obliczono odległości euklidesowe dla każdej ocenianej jednostki zarówno od wzorca (L^+), jak i antywzorca (L^-) według następujących wzorów:

$$L^+ = \sqrt{\sum_{j=1}^{K}\left(\bar{x}_j - z_j^+\right)^2} \text{ i } L^- = \sqrt{\sum_{j=1}^{K}\left(\bar{x}_j - z_j^-\right)^2},$$

gdzie $\bar{x}_j$ oznacza wartość j-tej cechy uśrednionej po branżach.

Obliczono wartość syntetycznego miernika sytuacji finansowej (S_i) dla każdej badanej branży przemysłu spożywczego według wzoru:

$$S_i = \frac{L_i^-}{L_i^+ + L_i^-},$$

gdzie $0 \leq S_i \leq 1$ ($i = 1, 2, \ldots, N$).

Im mniejsza odległość danej jednostki od wzorca rozwoju – czyli tym samym większa od antywzorca – tym wartość cechy syntetycznej jest bliższa 1, co świadczy o korzystniejszej sytuacji ekonomiczno-finansowej danej branży przemysłu spożywczego.

Na ostatnim etapie przeprowadzono porządkowanie liniowe i zaklasyfikowano badane branże do grup o niskiej, średniej lub wysokiej wartości wskaźnika syntetycznego (S_i) określającego sytuację ekonomiczno-finansową. Oparto się na kryterium statystycznym wykorzystującym średnią arytmetyczną ($\bar{S}$) i odchylenie standardowe (O_s) wartości miernika syntetycznego, gdzie:

- klasa I – $S_i \geq \bar{S} + O_s$;
- klasa II – $\bar{S} + O_s > S_i \geq \bar{S}$;
- klasa III – $\bar{S} > S_i \geq \bar{S} - O_s$;
- klasa IV – $S_i < \bar{S} - O_s$.

W badaniu wykorzystano wskaźniki ekonomiczno-finansowe za lata 2022 i 2019, co umożliwiło ocenę nie tylko bieżącej sytuacji ekonomiczno-finansowej głównych branż przemysłu spożywczego, lecz także jej zmiany w warunkach dużej zmienności otoczenia rynkowego pod wpływem zjawisk kryzysowych, które wystąpiły w ostatnich latach (pandemii COVID-19 i wojny w Ukrainie).

3. Branżowe zróżnicowanie sytuacji ekonomiczno-finansowej przemysłu spożywczego w 2022 r.

Do oceny sytuacji ekonomiczno-finansowej poszczególnych branż przemysłu spożywczego w 2022 r. wykorzystano dziewięć wskaźników opisujących rentowność, płynność finansową, zadłużenie oraz produktywność czynnika pracy i środków trwałych. Ich wartości oraz zmiany w stosunku do 2019 r. przedstawiono w tabl. 1.

Tabl. 1. Wskaźniki opisujące sytuację ekonomiczno-finansową poszczególnych branż przemysłu spożywczego w 2022 r. i ich zmiany w porównaniu z 2019 r.

Branże przemysłu spożywczego	RN	ROA	ROE	QR	CKG w dniach	ZO	ZD	PP w tys. zł/pracownika	PŚT w zł
	w %					w %			
Przemysł spożywczy	**4,19 ↑**	**7,15 ↑**	**13,61 ↑**	**1,45 ↑**	**29,5 ↑**	**47,4 ↑**	**17,8 ↑**	**161,6 ↑**	**2,95 ↑**
Tytoniowa	4,55↓	4,75 ↓	9,85 ↓	1,25 ↓	20,0 ↓	51,9 ↑	25,2 ↓	258,1 ↓	1,43 ↑
Napojów bezalkoholowych	1,85 ↓	2,05 ↓	3,67 ↓	1,00 ↓	–22,5 ↑	44,2 ↓	19,3 ↓	275,1 ↑	1,53 ↑
Piwowarska	7,42 ↓	7,33 ↓	16,70 ↓	1,30 ↑	–18,1 ↓	56,1 ↑	28,4 ↑	306,7 ↓	1,67 ↑
Spirytusowa	11,10 ↑	11,50 ↑	55,15 ↑	1,31 ↓	26,1 ↓	79,2 ↓	85,9 ↓	357,4 ↑	2,86 ↑
Olejarska	3,13 ↑	10,72 ↑	28,24 ↑	1,32 ↓	31,9 ↑	62,1 ↓	10,3 ↓	471,0 ↑	9,03 ↑
Cukiernicza	4,43 ↓	5,33 ↑	8,03 ↑	1,17 ↓	30,7 ↑	33,6 ↓	8,2 ↓	153,9 ↑	1,63 ↑

Tabl. 1. Wskaźniki opisujące sytuację ekonomiczno-finansową poszczególnych branż przemysłu spożywczego w 2022 r. i ich zmiany w porównaniu z 2019 r. (dok.)

Branże przemysłu spożywczego	RN	ROA	ROE	QR	CKG w dniach	ZO	ZD	PP w tys. zł/pracownika	PŚT w zł
	w %					w %			
Cukrownicza	11,30 ↑	8,13 ↑	10,99 ↑	3,11 ↓	127,3 ↑	26,0 ↓	0,1 ↑	452,1 ↑	1,82 ↑
Koncentratów spożywczych	4,96 ↓	6,93 ↓	12,39 ↓	1,60 ↓	42,9 ↓	44,1 ↓	12,0 ↓	162,2 ↓	2,46 ↑
Rybna	1,71 ↑	3,54 ↑	9,79 ↑	1,26 ↑	48,3 ↓	63,8 ↓	17,5 ↑	120,6 ↑	5,64 ↑
Owocowo-warzywna	4,01 ↑	4,81 ↑	11,03 ↑	1,49 ↓	78,5 ↑	56,4 ↓	31,6 ↑	152,4 ↑	2,50 ↑
Paszowa	3,91 ↓	8,11 ↑	15,14 ↑	1,56 ↓	47,0 ↑	46,4 ↑	15,1 ↑	253,9 ↑	3,78 ↑
Mleczarska	2,44 ↑	5,21 ↑	8,06 ↑	1,69 ↑	22,1 ↑	35,3 ↑	7,2 ↑	142,0 ↑	3,54 ↑
Mięsna	3,64 ↑	10,19 ↑	22,13 ↑	1,32 ↓	18,0 ↑	53,9 ↑	26,5 ↑	130,2 ↑	4,95 ↑
Młynarska	4,94 ↑	7,55 ↑	15,27 ↑	1,28 ↓	52,2 ↑	50,6 ↓	22,7 ↑	165,8 ↑	2,57 ↑
Piekarska	7,63 ↓	12,35 ↑	22,42 ↑	1,34 ↓	17,4 ↓	44,9 ↑	30,7 ↑	105,7 ↑	2,18 ↑
Winiarska	9,20 ↑	7,26 ↑	14,80 ↑	2,09 ↑	115,0 ↓	51,0 ↓	39,3 ↓	217,5 ↑	1,73 ↑

Uwaga. ↑ oznacza wzrost, a ↓ – spadek wartości wskaźnika w porównaniu z 2019 r. Symbole wskaźników objaśniono w zestawieniu.

Źródło: obliczenia własne na podstawie niepublikowanych danych GUS i GUS (2023).

Tabl. 2. Podstawowe statystyki charakteryzujące wskaźniki sytuacji ekonomiczno-finansowej badanych branż przemysłu spożywczego w 2022 r.

Wskaźniki	Min	Max	Kwartyl dolny	Mediana	Kwartyl górny	Współczynnik zmienności
RN w %	1,71	11,30	3,51	4,49	7,47	54,94
ROA w %	2,05	12,35	5,11	7,30	8,65	38,99
ROE w %	3,67	55,15	9,84	13,60	18,06	70,81
QR	1,00	3,11	1,28	1,32	1,57	31,87
CKG w dniach	–22,50	127,30	19,50	31,30	49,28	98,01
ZO w %	26,00	79,20	44,18	50,80	56,18	24,80
ZD w %	0,10	85,90	11,58	21,00	28,98	79,99
PP w tys. zł/pracownika	105,70	471,00	149,80	191,65	283,00	47,86
PŚT w zł	1,43	9,03	1,72	2,48	3,60	63,15

Źródło: opracowanie własne na podstawie tabl. 1.

Z danych zamieszczonych w tabl. 1 wynika, że rentowność netto przemysłu spożywczego w 2022 r. wyniosła 4,2% i była o 0,2 p.proc. wyższa niż w 2019 r., ale w poszczególnych branżach charakteryzowała się bardzo dużą zmiennością. Współczynnik zmienności tego wskaźnika wyniósł prawie 55%, przy rozstępie (różnicy między minimum a maksimum) na poziomie 9,6%. W 75% branż rentowność netto nie przekroczyła 7,5%, w 50% – wyniosła mniej niż 4,5%, a w 25% – mniej niż 3,5%. Najwyższą

rentowność netto osiągnięto w branżach: cukrowniczej (11,3%), spirytusowej (11,1%) i winiarskiej (9,2%); w każdej z nich wynik poprawił się w stosunku do 2019 r. Najniższy wskaźnik charakteryzował branże: rybną (1,7%), napojów bezalkoholowych (1,9%), mleczarską (2,4%) i olejarską (3,1%). Ich rentowność poprawiła się w porównaniu z 2019 r. (z wyjątkiem produkcji napojów bezalkoholowych).

Efektywność zarządzania aktywami całkowitymi mierzona stopą rentowności aktywów ogółem (ang. *return on assets* – ROA) w całym przemyśle spożywczym w 2022 r. wyniosła 7,2% i była o 1,3 p.proc. wyższa niż w 2019 r. Zróżnicowanie tego wskaźnika między wyodrębnionymi branżami przemysłu spożywczego okazało się duże (współczynnik zmienności na poziomie 39%, przy rozstępie 10,1%). Rentowność aktywów ogółem w 50% branż wyniosła co najmniej 7,3%, a w 75% – 8,7%. Do branż o najwyższej ROA należały: piekarska (12,4%), spirytusowa (11,5%), olejarska (10,7%) i mięsna (10,2%); w każdej z nich odnotowano wzrost stopy zwrotu aktywów w porównaniu z 2019 r. Najniższe wartości rozpatrywanego wskaźnika miały branża napojów bezalkoholowych (2,1%) i przetwórstwo ryb (3,5%) oraz branże: tytoniowa (4,8%), owocowo-warzywna (4,8%) i mleczarska (5,2%). Spadek wartości ROA odnotowano tylko w branżach: tytoniowej i napojów bezalkoholowych. W pozostałych branżach nastąpił wzrost, ale niższy od średniej w przemyśle spożywczym.

Rentowność kapitału własnego (ang. *return on equity* – ROE) zainwestowanego w przemyśle spożywczym w 2022 r. wyniosła 13,6% i była o 2,5 p.proc. wyższa niż w 2019 r. Oznacza to, że zyskowność środków zainwestowanych w sektorze była większa od innych, bezpiecznych możliwości inwestowania – obligacji skarbowych czy lokat bankowych. Istniało jednak bardzo duże międzybranżowe zróżnicowanie tego wskaźnika (współczynnik zmienności – 70,8%[2], rozstęp – 51,5%). W 50% branż ROE kształtowała się na poziomie równym 13,6% lub niższym (tabl. 2). Rozkład branż był w dużej mierze zbieżny z rozkładem według ROA, szczególnie w przypadku branż o najwyższej rentowności aktywów. Najwyższą stopę zwrotu kapitału osiągnęły przedsiębiorstwa branż: spirytusowej (55,2%), olejarskiej (28,2%), piekarskiej (22,4%) i mięsnej (22,1%); w każdej z nich odnotowano wzrost w relacji do 2019 r. W 25% branż ROE nie przekroczyła 9,8%, a w 75% wyniosła co najwyżej 18,1%. Najniższą stopę zwrotu odnotowano w produkcji napojów bezalkoholowych (3,7%), w której nastąpił spadek w porównaniu z 2019 r., oraz w branżach: cukierniczej (8,0%), mleczarskiej (8,1%) i rybnej (9,8%), w których wartość wskaźnika wzrosła w relacji do okresu poprzedzającego zjawiska kryzysowe, a także w tytoniowej (9,9%), w której mimo spadku wartości ROE inwestowanie w sektorze było nadal konkurencyjne wobec innych możliwości inwestowania.

[2] Wartość współczynnika zmienności (v) poniżej 10% oznacza małą zmienność, $10\% \leq v < 30\%$ – średnią, $30\% \leq v < 50\%$ – dużą, a $v \geq 50\%$ – bardzo dużą (Czerwińska-Kayzer i in., 2013).

W ocenie sytuacji ekonomiczno-finansowej przetwórstwa spożywczego ważnym kryterium jest płynność finansowa. Można ją określić m.in. za pomocą wskaźnika płynności bieżącej (ang. *quick ratio* – QR), który świadczy o zdolności do regulowania bieżących zobowiązań. W większości branż przemysłu spożywczego wartość tego wskaźnika w 2022 r. zmniejszyła się w porównaniu z 2019 r., ale tylko w produkcji napojów bezalkoholowych okazała się zbyt mała, ponieważ nie osiągnęła wartości referencyjnej na poziomie 1,2 (Czerwińska-Kayzer, 2022). QR cechuje się stosunkowo niewielkim zróżnicowaniem międzybranżowym w porównaniu ze stanem z lat wcześniejszych, co można stwierdzić na podstawie przeglądu literatury przedmiotu (Czerwińska-Kayzer i Stanisławska, 2013). W 50% branż jego wartość przekraczała 1,32, a w 25% – 1,57. Najwyższą wartość osiągnął w branżach: cukrowniczej, winiarskiej, mleczarskiej i koncentratów spożywczych. O zbyt małym udziale środków własnych w finansowaniu bieżącej działalności gospodarczej można mówić tylko w przypadku branż: cukierniczej, w której wartość aktywów obrotowych była o 17% wyższa od kwoty zobowiązań krótkoterminowych, i napojów bezalkoholowych, w której bieżąca działalność gospodarcza była w pełni finansowana środkami obcymi. Przedsiębiorstwa poszczególnych branż przemysłu spożywczego na ogół nie miały problemów z terminową obsługą bieżących zobowiązań płatniczych.

Środki własne w obrocie przekładają się na długość cyklu konwersji gotówki, który w przemyśle spożywczym przejawia wyraźną tendencję do skracania się (Gołaś i in., 2012). W 2022 r. wyniósł on średnio ok. 30 dni i cechował się bardzo dużym zróżnicowaniem międzybranżowym (współczynnik zmienności – 98%, duży rozstęp – ok. 150 dni – między minimalną a maksymalną wartością wskaźnika). Przeważały branże o długości cyklu większej niż średnia. Do branż przemysłu spożywczego z cyklem krótszym niż 19,5 dnia, który miało 25% branż, należały: piwowarska, napojów bezalkoholowych, piekarska i mięsna. W pierwszych dwóch wymienionych branżach cykl był ujemny, co oznacza, że finansowanie działalności operacyjnej odbywało się poprzez maksymalne wydłużanie terminów płatności zobowiązań wobec dostawców. Z kolei cykl dłuższy niż 49 dni, co miało miejsce w 75% branż, dotyczył gałęzi: cukrowniczej, winiarskiej, owocowo-warzywnej i młynarskiej. W pewnej mierze jego długość mogła wynikać z sezonowego charakteru zaopatrzenia surowcowego tych kierunków przetwórstwa. W 9 z 16 badanych branż cykl konwersji gotówki w 2022 r. był krótszy niż w 2019 r.; średnio w przemyśle spożywczym – o 5 dni.

Kolejna kategoria oceny sytuacji ekonomiczno-finansowej przemysłu spożywczego to zadłużenie. Poszczególne miary zadłużenia, zarówno ogółem, jak i długoterminowego, wskazują, że generalnie problem nadmiernego wykorzystania kapitału obcego do finansowania działalności nie dotyczył firm przemysłu spożywczego. Zadłużenie ogółem w tym sektorze w porównywanych latach było względnie stabilne; zobowiązania stanowiły ok. 47% aktywów ogółem. Stosunkowo niewielkie okazało się zadłużenie

międzybranżowe: w 50% branż było ono równe 50,8% lub niższe, a w 25% – większe niż 56,2%. Przedsiębiorstwa spożywcze odznaczały się na ogół optymalnym i bezpiecznym udziałem zobowiązań w finansowaniu aktywów. Najbardziej zadłużone były branże: spirytusowa, rybna i olejarska (ponad 60%), a najmniej: cukrownicza, cukiernicza i mleczarska.

Firmy spożywcze sukcesywnie ograniczały zadłużenie długoterminowe, które w 2022 r. stanowiło 17,8% wartości kapitału własnego – o 4,4 p.proc. mniej niż w 2019 r. Również w przypadku tego wskaźnika odnotowano bardzo duże zróżnicowanie międzybranżowe: w 25% branż było niewielkie (niższe niż 11,6%), w kolejnych 25% – średnie (11,6–21%), w następnych 25% – ponadprzeciętne (21–29%), w ostatnich 25% – wysokie (ponad 29%). Najbardziej zadłużone zobowiązaniami długoterminowymi były branże: spirytusowa, winiarska, owocowo-warzywna i piekarska, a najmniej – branże: cukrownicza, mleczarska i cukiernicza.

W ocenie sytuacji ekonomiczno-finansowej poszczególnych branż przemysłu spożywczego należy uwzględnić również produktywność podstawowych czynników produkcji: pracy i środków trwałych. Wydajność pracy (mierzona WDB w cenach stałych przypadającej na pracownika) w przemyśle spożywczym w 2022 r. była o 15,3% większa niż w 2019 r. i cechowała się dużym zróżnicowaniem międzybranżowym. W 50% branż przekraczała 190 tys. zł/pracownika, w tym największa była w branżach olejarskiej i cukrowniczej, a w 25% branż wynosiła poniżej 150 tys. zł/pracownika, w tym najniższa była w branżach: piekarskiej, rybnej i mięsnej.

Produktywność środków trwałych w przemyśle spożywczym w 2022 r. była o blisko 30% wyższa niż w 2019 r. Wzrost wartości tego wskaźnika odnotowano we wszystkich branżach tego sektora, przy jednoczesnym dużym zróżnicowaniu międzybranżowym. Największy przyrost produkcji (w cenach bazowych) po powiększeniu środków trwałych odnotowano w branżach: olejarskiej, rybnej i mięsnej, w których wzrost wartości środków trwałych o 1 zł powodował powiększenie wartości produkcji o 5–9 zł, a najmniejszy – w branżach: tytoniowej, napojów bezalkoholowych, cukierniczej, piwowarskiej, winiarskiej i cukrowniczej (o mniej niż 2 zł). Produktywność środków trwałych była jedynym wskaźnikiem, którego wartość dla każdej z 16 wyodrębnionych branż zwiększyła się w 2022 r. w stosunku do 2019 r.

4. Klasyfikacja branż przemysłu spożywczego

Podczas analizy sytuacji ekonomiczno-finansowej poszczególnych branż przemysłu spożywczego przy wykorzystaniu miernika syntetycznego przeprowadzono porządkowanie liniowe, a następnie zaklasyfikowano każdą branżę tego sektora do jednej z czterech klas typologicznych ze względu na sytuację ekonomiczno-finansową w 2022 r. Wyniki tej klasyfikacji porównano z wynikami za 2019 r. (tablice 3 i 4).

Tabl. 3. Klasyfikacja branż przemysłu spożywczego na podstawie wartości syntetycznego miernika oceny sytuacji ekonomiczno-finansowej (S_i)

Wartości graniczne S_i	Poziom	Klasa typologiczna	Wartość S_i dla badanych branż
		2019	
> 0,5513	wysoki	I	piwowarska (0,6439) napojów bezalkoholowych (0,5696) piekarska (0,5615)
0,5513–0,4772	średni wyższy	II	tytoniowa (0,5212) koncentratów spożywczych (0,5151) cukiernicza (0,5022) spirytusowa (0,5021)
0,4773–0,4034	średni niższy	III	mleczarska (0,4716) olejarska (0,4616) mięsna (0,4499) rybna (0,4397) cukrownicza (0,4258) winiarska (0,4245) paszowa (0,4134) młynarska (0,4095)
< 0,4034	niski	IV	owocowo-warzywna (0,3145)
		2022	
> 0,5518	wysoki	I	piwowarska (0,5983) piekarska (0,5615)
0,5518–0,4866	średni wyższy	II	spirytusowa (0,5499) olejarska (0,5476) ↑ mięsna (0,5246) ↑ napojów bezalkoholowych (0,5219) ↓ mleczarska (0,5075) ↑ cukiernicza (0,4887) koncentratów spożywczych (0,4877) tytoniowa (0,4869) młynarska (0,4868) ↑
0,4867–0,4217	średni niższy	III	paszowa (0,4746) cukrownicza (0,4729)
< 0,4217	niski	IV	rybna (0,4178) ↓ winiarska (0,3636) ↓ owocowo-warzywna (0,3421)

Uwaga. ↑ oznacza awans do klasy o wyższym poziomie miernika w stosunku do 2019 r., a ↓ – przejście do niższej klasy; w pozostałych przypadkach branże pozostały w tych samych klasach typologicznych co w 2019 r.

Źródło: opracowanie własne na podstawie niepublikowanych danych GUS i GUS (2023).

Tabl. 4. Wewnątrzklasowe wartości cech – cząstkowych mierników oceny sytuacji ekonomiczno-finansowej branż przemysłu spożywczego

Wskaźniki	2022					2019
	ogółem	klasa typologiczna				
		I	II	III	IV	
RN w %	4,49	7,53	4,43	7,61	4,01	4,73
ROA w %	7,30	9,84	6,93	8,12	4,81	5,45
ROE w %	13,60	19,56	12,39	13,07	11,03	11,12
QR	1,32	1,32	1,31	2,34	1,49	1,50
CKG w dniach	31,30	–0,35	26,10	87,15	78,50	32,80
ZO w %	50,80	50,50	50,60	36,20	56,40	48,45
ZD w %	21,00	29,55	19,30	7,60	31,60	20,70
PP w tys. zł/pracownika	191,65	206,20	165,80	353,00	152,40	178,45
PŚT w zł	2,48	1,93	2,57	2,80	2,50	1,81

Źródło: opracowanie własne na podstawie tabl. 1 i 3.

W 2022 r. do klasy I, o najkorzystniejszej sytuacji ekonomiczno-finansowej, zakwalifikowały się tylko dwie branże: piwowarska i piekarska, mimo że niektóre ich wskaźniki rentowności i płynności pogorszyły się w porównaniu z 2019 r. (tabl. 3). Klasa ta odznaczała się wysoką rentownością netto (tabl. 4) oraz najwyższą rentownością kapitału własnego (19,6%) i aktywów ogółem (9,8%) na tle pozostałych klas typologicznych. Wysoka stopa zysku w relacji do przychodów netto oraz kapitału własnego i aktywów w tej klasie w znacznym stopniu była determinowana dużym udziałem kapitału obcego w finansowaniu działalności gospodarczej. Świadczy o tym niemal najniższa mediana wskaźnika płynności finansowej oraz ujemny cykl konwersji gotówki i duży udział zobowiązań ogółem (w relacji do aktywów ogółem), w tym głównie długoterminowych (do kapitału własnego). Ujemny cykl konwersji gotówki oznacza komfortową sytuację dla firm, ponieważ ich działalność jest finansowana przez dostawców. Przedsiębiorstwa osiągały wysoką efektywność finansową dzięki właściwej polityce finansowania działalności gospodarczej oraz sprawnemu zarządzaniu majątkiem. Należy także podkreślić, że klasę I charakteryzuje z jednej strony wyższa niż średnia w przemyśle spożywczym wydajność pracy, a z drugiej – najniższa wśród wszystkich klas produktywność środków trwałych, co stanowi jedyną słabą stronę branż tej klasy i może wynikać z wysokiej majątkochłonności, głównie branży piwowarskiej.

Klasa II była najliczniejsza – znalazło się w niej dziewięć branż przemysłu spożywczego: spirytusowa, olejarska, mięsna, napojów bezalkoholowych, mleczarska, cukiernicza, koncentratów spożywczych, tytoniowa i młynarska. Część z nich (olejarska, mięsna, mleczarska i młynarska) w porównaniu z 2019 r. przesunęła się z klasy III, a branża napojów bezalkoholowych – z klasy I. Branże te cechowała minimalnie niższa od średniej dla badanych branż rentowność netto oraz nieco mniejsze ROA i ROE,

niemniej jednak poziom tych wskaźników był stosunkowo dobry. Rentowność kapitału własnego wyniosła 12,4%, co stanowiło efekt odpowiedniej polityki cen i kontroli kosztów oraz sprawnego gospodarowania majątkiem. W rezultacie płynność finansowa okazała się bezpieczna dla terminowej realizacji zobowiązań płatniczych, a długość cyklu konwersji gotówki – satysfakcjonująca. Średnia wartość wskaźnika płynności finansowej w tych branżach wyniosła 1,31, a długość cyklu obrotu gotówki nie przekroczyła 30 dni, co jest dobrym rezultatem, lepszym niż przeciętny w badanych branżach przemysłu spożywczego. Klasę tę cechował poziom zadłużenia ogółem zbliżony do średniego w przemyśle spożywczym i nieco niższy poziom zadłużenia długoterminowego, dla których mediana wyniosła odpowiednio 50,6% wobec 50,8% oraz 19,3% wobec 21%. Słabą stroną tej klasy była wydajność pracy mierzona wartością WDB/pracownika (165,8 tys. zł przy 191,65 tys. zł średnio w badanych branżach), ale już produktywność środków trwałych była nieco wyższa niż średnia w przemyśle spożywczym (odpowiednio 2,57 zł i 2,48 zł).

Klasę III tworzyły dwie branże: paszowa i cukrownicza, które utrzymały swoją pozycję z 2019 r. Na podstawie innych badań można stwierdzić, że w latach 2011–2015 pozycja branży paszowej pod względem efektywności ekonomicznej była podobna (Czerwińska-Kayzer i Florek, 2018). Poziom sytuacji ekonomiczno-finansowej klasy III można określić jako średni niski, mimo że jej rentowność netto osiągnęła najwyższą wartość ze wszystkich klas, ale stopa zwrotu z kapitału własnego była niższa od średniej dla przemysłu spożywczego i wyniosła 13,07%. Ze względu na dużą rozbieżność prezentowanych wyników stan finansowy podmiotów branż klasy III należy analizować odrębnie. W branży paszowej był on bezpieczny, a długość cyklu obrotu gotówką – dopuszczalna. Natomiast w cukrownictwie zaobserwowano nadpłynność finansową i zbyt długi cykl konwersji gotówki. Świadczy to o tym, że słabą stroną tej branży jest gospodarowanie majątkiem obrotowym, co w pewnej mierze wynika z jej kampanijnego charakteru. Przedsiębiorstwa obu branż charakteryzuje niewielkie zadłużenie – zarówno ogółem, jak i długoterminowe – co wskazuje na niewykorzystywanie efektów dźwigni finansowej (Gołębiowski i Tłaczała, 2009) i ograniczanie zaangażowania kapitału obcego w finansowanie działalności. Pozytywną cechą tej klasy jest najwyższa w porównaniu z pozostałymi klasami produktywność, zarówno pracy, jak i środków trwałych. W 2022 r. jej mediana wyniosła odpowiednio 353 tys. zł WDB/pracownika oraz 2,80 zł, tj. znacząco powyżej mediany dla przemysłu spożywczego.

Klasę IV – o najsłabszej sytuacji ekonomiczno-finansowej – tworzyły trzy branże: rybna, winiarska i owocowo-warzywna. Większość opisujących ją wskaźników miała najniższe wartości wśród wszystkich klas typologicznych i na ogół kształtowały się one poniżej mediany dla wszystkich badanych branż. Wartości niektórych wskaźników można uznać za zadowalające: mediana rentowności netto wyniosła 4%, aktywów – 4,8%, a ROE – ok. 11%, co wskazuje na konkurencyjność sektora wobec

alternatywnych możliwości inwestowania środków. Stan finansowy przedsiębiorstw należących do tej klasy był bezpieczny, ale cykl obrotu gotówką – dwukrotnie dłuższy od mediany w badanych branżach przemysłu spożywczego. To słaba strona tej klasy, ponieważ firmy zmuszone były poszukiwać środków pieniężnych na prawie 80 dni. Branże tej klasy cechowało również duże zadłużenie, nie tylko ogółem, lecz także zobowiązaniami długoterminowymi. Mediana zobowiązań ogółem stanowiła 56,4% aktywów ogółem, a zobowiązań długoterminowych – 31,6% kapitału własnego. Wysokie zaangażowanie kapitału obcego w finansowanie działalności stwarzało możliwość uzyskania większego efektu dźwigni finansowej w porównaniu z pozostałymi klasami. W rezultacie mediana wydajności pracy w tej klasie była najniższa, poniżej średniej w badanych branżach, ale produktywność środków trwałych – taka jak średnia w przemyśle spożywczym.

5. Podsumowanie

Przeprowadzona analiza sytuacji ekonomiczno-finansowej przemysłu spożywczego, w której wykorzystano syntetyczny miernik rozwoju, pozwoliła wyodrębnić cztery klasy typologiczne branż tego sektora. Uzyskane wyniki wskazują na duże zróżnicowanie międzybranżowe sytuacji ekonomiczno-finansowej zarówno w 2022 r., jak i w stanowiącym punkt odniesienia 2019 r. W najlepszej sytuacji ekonomiczno-finansowej były branże ukierunkowane na produkcję niektórych używek i przetwórstwo wtórne (klasa I). Przedsiębiorstwa tych branż osiągały wysoki stopień zwrotu z kapitału własnego, korzystna była struktura finansowania ich działalności, ponadto wykorzystywano w nich efekt dźwigni finansowej. Najsłabszą sytuację ekonomiczno-finansową zanotowano w branżach: owocowo-warzywnej, winiarskiej i rybnej (klasa IV). Problemami były ich niska rentowność i długi cykl obrotu gotówką, co przekładało się na stosunkowo niską produktywność, szczególnie pracy.

Rezultaty badania pozwalają także stwierdzić, że to nie kierunek przetwórstwa spożywczego (przetwórstwo produktów roślinnych i zwierzęcych, przerób wtórny czy produkcja używek), ale rodzaj przetwórstwa (poszczególne branże przemysłu spożywczego) decyduje o sytuacji ekonomiczno-finansowej. Branże każdego kierunku przetwórstwa spożywczego zostały sklasyfikowane w dwóch lub trzech klasach typologicznych. Wśród niewątpliwych determinant kształtujących wyniki ekonomiczno-finansowe w poszczególnych branżach należy wymienić: otoczenie makroekonomiczne, relacje kosztowo-cenowe i konkurencję w danej branży.

Bibliografia

Czerwińska-Kayzer, D. (2022). *Przemysł spożywczy i jego zróżnicowanie w zakresie płynności finansowej*. Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu.

Czerwińska-Kayzer, D., Florek, J. (2018). Zróżnicowanie efektywności w branży produkującej pasze i karmę dla zwierząt. *Przedsiębiorstwo & Finanse*, (4), 5–15. http://pif.wsfiz.edu.pl/wp-content/uploads/2019/05/PiF_2018_423.pdf.

Czerwińska-Kayzer, D., Florek, J., Stanisławska, J. (2013). Zastosowanie metody TOPSIS do oceny sytuacji finansowej przemysłu spożywczego. *Zagadnienia Ekonomiki Rolnej*, *335*(2), 22–37.

Czerwińska-Kayzer, D., Stanisławska, J. (2013). Zróżnicowanie płynności finansowej w polskim przemyśle spożywczym w 2011 roku. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, *15*(2), 52–57. https://rnseria.com/article/119648/pl.

Firlej, K. (2008). *Rozwój przemysłu rolno-spożywczego w sektorze agrobiznesu i jego determinanty*. Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.

Główny Urząd Statystyczny. (2022). *Rocznik Statystyczny Rzeczypospolitej Polskiej 2022*. https://stat.gov.pl/obszary-tematyczne/roczniki-statystyczne/roczniki-statystyczne/rocznik-statystyczny-rzeczypospolitej-polskiej-2022,2,22.html.

Główny Urząd Statystyczny. (2023). *Wyniki finansowe przedsiębiorstw w 2022 roku*. https://stat.gov.pl/obszary-tematyczne/podmioty-gospodarcze-wyniki-finansowe/przedsiebiorstwa-niefinansowe/wyniki-finansowe-przedsiebiorstw-w-2022-roku,40,1.html.

Gołaś, Z., Bieniasz, A., Łuczak, A. (2012). Zróżnicowanie kondycji ekonomiczno-finansowej przemysłu spożywczego w Polsce w latach 2005–2010. *Zagadnienia Ekonomiki Rolnej*, *333*(4), 100–122.

Gołębiowski, G., Tłaczała, A. (2009). *Analiza finansowa w teorii i w praktyce*. Difin.

Hwang, C. L., Yoon, K. (1981). *Multiple Attribute Decision Making. Methods and Applications: A State-of-the-Art Survey*. Springer-Verlag. https://doi.org/10.1007/978-3-642-48318-9.

Kowalczyk, J., Kusak, A. (2006). *Decyzje finansowe firmy. Metody analizy.* Wydawnictwo C.H. Beck.

Politechnika Białostocka. (2022). *Wielokryterialne metody optymalizacji procesów decyzyjnych.* https://navoica.pl/courses/course-v1:PolitechnikaBialostocka+PB.WIZ.05+2022_2/course/.

Saaty, T. L., Vargas, L. G. (2006). *Decision Making with the Analytic Network Process. Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks*. Springer. https://doi.org/10.1007/0-387-33987-6.

Sierpińska, M., Jachna, T. (2004). *Ocena przedsiębiorstw według standardów światowych*. Wydawnictwo Naukowe PWN.

Szczepaniak, I., Ambroziak, Ł., Bułkowska, M., Drożdż, J. (2022). Znaczenie przemysłu spożywczego w gospodarce Polski i Unii Europejskiej. *Przemysł Spożywczy*, *76*(12), 5–12. https://doi.org/10.15199/65.2022.12.1.

Szczepaniak, I., Firlej, K. (red). (2015). *Przemysł spożywczy – makrootoczenie, inwestycje, ekspansja zagraniczna*. Fundacja Uniwersytetu Ekonomicznego w Krakowie. http://ierigz.waw.pl/publikacje/poza-seria/przemysl-spozywczy-makrootoczenie-inwestycje-ekspansja-zagraniczna.

Wijnands, J. H. M., Verhoog, D. (2016). *Competitiveness of the EU food industry. Ex-post assessment of trade performance embedded in international economic theory*. LEI Wageningen UR. https://edepot.wur.nl/369980.

Wójciak, M. (b.r.). *Metodologia normalizacji kryteriów oceny ekoefektywności technologii*. https://gig.eu/sites/default/files/attachments/projekty/zad_3.2_metodologia_normalizacji_kryteriow_oceny_ekoefektywnosci_technologii.pdf.

Wysocki, F. (2010). *Metody taksonomiczne w rozpoznaniu typów ekonomicznych rolnictwa i obszarów wiejskich*. Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu.

Wysocki, F., Lira, J. (2007). *Statystyka opisowa*. Wydawnictwo Akademii Rolniczej im. Augusta Cieszkowskiego w Poznaniu.

Wiadomości Statystyczne. The Polish Statistician, 2024, *69*(11), 54–68 https://doi.org/10.59139/ws.2024.11.4
Edukacja statystyczna / Statistical education
Zgłoszony/submitted: 28.02.2024, poprawiony/revised: 11.07.2024, zaakceptowany/accepted: 2.09.2024

Próba reprezentatywna – potrzeba i propozycja definicji

Mirosław Szreder[a]

Streszczenie. „Próba reprezentatywna" należy do tych nielicznych pojęć wnioskowania statystycznego, które nie doczekały się precyzyjnej i powszechnie uznanej definicji. Potrzeba istnienia takiej definicji jest mniej odczuwalna w naukach ścisłych i przyrodniczych, w których empiryczne badania eksperymentalne, mające głównie cele eksploracyjne, rzadko wymagają uzyskania prób reprezentatywnych. Z kolei cele deskryptywne i konfirmacyjne, właściwe dla nauk społecznych i ekonomicznych, sprawiają, że jednym z najważniejszych wymogów odnoszących się do próby badawczej jest możliwość dokonywania na jej podstawie uogólnień odpowiednich oszacowań dla populacji wraz z oceną ich błędów.

Celem artykułu jest krytyczne omówienie roli próby reprezentatywnej we wnioskowaniu naukowym i statystycznym, a ponadto zaproponowanie definicji „próby reprezentatywnej" i wskazanie jej specyfiki względem prób nielosowych, których reprezentatywność jest ograniczona jedynie do ściśle wskazanych zmiennych. W proponowanej definicji autor kładzie nacisk na zgodność struktury próby ze strukturą populacji oraz na możliwości wnioskowania, w którym ocenom uzyskanym z próby są przypisane odpowiednio oszacowane liczbowe wartości błędów. Postuluje także, aby próbę, która wykazuje zbieżność rozkładów tylko niektórych cech z ich rozkładami populacyjnymi, nazywać nie „próbą reprezentatywną", lecz „próbą reprezentatywną ze względu na wyszczególnione zmienne". Uzasadniono także zbędność tworzenia we wnioskowaniu statystycznym kategorii „próby reprezentacyjnej".

Słowa kluczowe: próba reprezentatywna, wnioskowanie statystyczne, wnioskowanie naukowe, cele eksploracyjne, cele deskryptywne, cele konfirmacyjne

JEL: C10, C13, C18

Representative sample – the need for and the proposal of a definition

Abstract. A 'representative sample' is one of few terms used in statistical inference which do not have a precise and commonly recognised definition. The need for such a definition is less nagging in sciences and natural sciences, where empirical experimental research has exploratory objectives, and thus rarely requires obtaining representative samples. On the other hand, descriptive and confirmatory objectives of social sciences and economics necessitate samples on the basis of which it would be possible to make generalisations of relevant estimates for a given population and assess their errors.

The aim of this paper is to discuss the role of a representative sample in scientific and statistical inference, and additionally to propose its definition and show its specificity in the case where at the same time it is a non-random sample, whose representativeness is limited to some specified variables only.In the proposed definition, the author focuses on the consistency of the sample structure with the structure of the population and the possibility of making inference

[a] Uniwersytet Gdański, Wydział Zarządzania, Polska / University of Gdańsk, Faculty of Management, Poland. ORCID: https://orcid.org/0000-0002-7597-0816. E-mail: miroslaw.szreder@ug.edu.pl.

in which estimates obtained from the sample have numerical error values assigned to them. It is also argued that samples which show convergence of the distributions of only some characteristics with their population distributions should be called 'representative samples in terms of specified variables' instead of just 'representative samples'. Moreover, it is demonstrated that the category of a random sample which is inseparably linked to statistical inference does not need any equivalents.

Keywords: representative sample, statistical inference, scientific inference, exploratory objectives, descriptive objectives, confirmatory objectives

1. Wprowadzenie

W badaniach statystycznych – pełnych i niewyczerpujących (próbkowych) – istnieje wiele kategorii i pojęć od dawna precyzyjnie zdefiniowanych i niebudzących kontrowersji co do ich rozumienia czy stosowania. Dotyczy to zwłaszcza opisu statystycznego, charakterystycznego dla badań pełnych, czyli takich, w których pomiarem (obserwacją) objęte zostają wszystkie jednostki analizowanej populacji. Więcej niejednoznaczności, jeśli chodzi o definiowanie i rozumienie kategorii stosowanych przez statystyków, spotyka się w badaniach niewyczerpujących. Jest to spowodowane zarówno stosunkowo krótkim – bo zaledwie stuletnim – okresem rozwoju i utrwalania się w nauce matematycznego (probabilistycznego) modelu wnioskowania statystycznego, jak i ściślejszymi niż w opisie statystycznym wymogami jednoznaczności rozumienia pojęć w modelowaniu matematycznym. Dokonujący się od trzech dekad dynamiczny rozwój informatycznych technologii przetwarzania danych, któremu towarzyszy równie szybki wzrost zainteresowania badaczy niewyczerpującymi badaniami statystycznymi, pogłębia potrzebę jednoznacznego definiowania najważniejszych pojęć występujących w tego rodzaju badaniach. Wydaje się to jednym z podstawowych warunków rzetelności przekazu wyników badań statystycznych oraz poprawnej i jednoznacznej ich interpretacji.

Jedną z najważniejszych kategorii występujących w badaniach niewyczerpujących jest *próba badawcza*, nazywana też *próbą statystyczną*. Od sposobu pobierania próby i konkretnej techniki jej uzyskania zależą nie tylko możliwości uogólniania wyników na populację, z której pochodzi próba, lecz także możliwości określenia precyzji wnioskowania i wielkości błędów otrzymanych ocen czy trafności podjętych decyzji odnoszących się do postawionych wcześniej hipotez. To naturalne, że atrybuty próby, zdeterminowane sposobem jej wyboru, zostają dla wygody ujęte w krótkiej, zwykle dwuczłonowej, nazwie tworzącej nową kategorię, np. *próba losowa*, *próba kwotowa*, *próba reprezentatywna* itp. Wśród tych kategorii za jedną z najważniejszych – a jednocześnie najbardziej niejednoznacznych – należy uznać *próbę reprezentatywną* (zob. np. Analytical Methods Committee, 2016; Rudolph i in., 2023). Pojęcie *próby reprezentatywnej* nie jest ani tożsame z pojęciem *próby losowej*, ani nie jest równoważne z żadną inną popularną kategorią statystyczną.

W statystyce nie istnieje jedna, uznana przez większość badaczy definicja *próby reprezentatywnej*. Potrzeba stworzenia takiej definicji istniała od dawna, ale obecnie staje się

szczególnie pilna wobec rosnącej frakcji badań, mniej lub bardziej profesjonalnych, w których zwięzłe stwierdzenie, że wykonano je na próbie reprezentatywnej, bez wskazania metody jej uzyskania, ma stanowić wystarczające uzasadnienie dla ekstrapolacji wniosków na całą populację. Dotyczy to w szczególności niektórych badań rynkowych i społecznych, w tym części badań opinii publicznej, w których użycie kategorii *próba reprezentatywna* ma z jednej strony informować o możliwości uogólnienia wniosków z próby na całą zbiorowość, z której pochodzi próba, a z drugiej – w wielu przypadkach zniechęcić odbiorcę do wnikania w szczegółową procedurę jej wyboru. Trzeba jednak zauważyć, że brak jednoznacznej, ścisłej definicji *próby reprezentatywnej* uniemożliwia w praktyce ocenę rodzajów i skali błędów, którymi mogą być obciążone wyniki badania niewyczerpującego. Gdy termin ten jest różnie definiowany i rozumiany zarówno przez badaczy, jak i odbiorców, wówczas nie jest jasne, jakie konkretnie techniki próbkowania mogą się kryć za tym terminem, a w rezultacie – jakiego rodzaju i jakiej wielkości błędy mogą obciążać wyniki takiego badania.

Dążenie do przypisania próbie badawczej waloru reprezentatywności wiąże się oczywiście z samą istotą wnioskowania statystycznego. Próba dobrze reprezentująca populację, z której pochodzi, daje prawo do formułowania prawidłowości i opisu wzorców, które są charakterystyczne dla całej tej populacji. Z kolei próba, która nie stanowi dobrej reprezentacji populacji, jest pozbawiona prawa do uogólnień. Reprezentatywność ujawnia więc swoje podstawowe zalety we wnioskowaniu statystycznym. Czy również we wnioskowaniu naukowym jako takim? To już jest kwestia warta przedyskutowania.

Celem artykułu jest krytyczne omówienie roli próby reprezentatywnej we wnioskowaniu naukowym i statystycznym, a ponadto zaproponowanie definicji *próby reprezentatywnej* i wskazanie jej specyfiki względem prób nielosowych, których reprezentatywność jest ograniczona jedynie do ściśle wskazanych zmiennych.

2. Próba reprezentatywna we wnioskowaniu naukowym i statystycznym

Relacji między wnioskowaniem naukowym (ang. *scientific inference*) a wnioskowaniem statystycznym (ang. *statistical inference*) nie można przekonująco scharakteryzować bez odwołań do istoty poznania naukowego i celów badań naukowych. Ważną rolę w dyskusji na ten temat odgrywa kategoria próby, a ściślej możliwości dokonywania uogólnień na podstawie próby. „Esencją wiedzy jest uogólnianie”[1] – stwierdza Reichenbach w swoim słynnym dziele na temat filozofii nauki (1951, s. 5). I dodaje: „Uogólnienie ponadto jest samą naturą wyjaśnienia. Przez wyjaśnienie

[1] Oryg.: „The essence of knowledge is generalization”.

zaobserwowanego faktu rozumiemy włączenie tego faktu w ramy ogólnego prawa"[2] (Reichenbach, 1951, s. 6). Dalej wskazuje, że naukowe objaśnianie wymaga dużej liczby obserwacji i krytycznego myślenia.

Podobnie cele nauki rozumieją współcześni badacze uczestniczący w dyskusji na temat miejsca wnioskowania statystycznego w nauce. Podkreślają oni, że naukowe uogólnienie wiąże się ze znalezieniem i wskazaniem warunków (okoliczności), w odniesieniu do których uzyskane wyniki badawcze lub zaobserwowane empirycznie regularności mają zastosowanie (Rothman i in., 2013). Przy czym w naukowym poznaniu chodzi głównie o dążenie do sformułowania uniwersalnych praw natury, nieograniczonych do szczególnego miejsca i ustalonego czasu obserwacji. W tym kontekście wiele badań statystycznych – zwłaszcza społecznych lub rynkowych – w mniejszym stopniu będzie realizować tego typu cele eksploracyjne, a raczej będzie zorientowanych na cele opisowe. Badania eksploracyjne zmierzają do poznania przyczyn i mechanizmów działań, postawienia nowych pytań oraz doskonalenia metod do dalszych badań (Babbie, 2004; Tong, 2019), a badania opisowe – również dostarczające ważnych informacji – takich celów w większości nie mają. Dotyczy to zwłaszcza badań niewyczerpujących w obszarze zagadnień ekonomicznych i społecznych, które najczęściej odnoszą się do precyzyjnie zdefiniowanej, specyficznej w czasie i przestrzeni populacji. Uogólnienie wykraczające poza populację, z której pobrano próbę, wymagałoby dodatkowej wiedzy lub wsparcia informacyjnego z innych źródeł. Nie wynika z tego oczywiście żaden wniosek, który pozwalałby kwestionować wartość naukową lub poznawczą wiedzy, która jest rezultatem kumulacji osiągnięć ze zrealizowanych poprawnie niewyczerpujących badań statystycznych. Opisowe, a także konfirmacyjne znaczenie badań statystycznych (np. w naukach medycznych i przyrodniczych) jest trudne do przecenienia. Niektórzy badacze (np. Tong, 2019) podkreślają właśnie tę ich rolę (konfirmacyjną i opisową) w poznaniu naukowym. Stwierdzają, że wnioskowanie statystyczne jest najodpowiedniejsze w analizach konfirmacyjnych, w których wcześniej dokładnie określono plan badawczy i metody badawcze, a samo wnioskowanie stanowi ważny końcowy etap iteracyjnego procesu uczenia się (ang. *iterative learning process*), typowego dla poszukiwań naukowych[3].

Wydaje się, że dobrym punktem wyjścia do określenia pożądanych właściwości próby badawczej w badaniach naukowych i statystycznych mogą być dwa wyżej

[2] Oryg.: „Generalization, furthermore, is the very nature of explanation. What we mean by explaining an observed fact is incorporating that fact into a general law".

[3] Podobną opinię wyraża George Box, który stwierdza m.in.: „Nie ma innego uzasadnienia dla istnienia statystyki poza byciem katalizatorem dla dociekań naukowych, w których tylko ostatnie stadium, gdy wykonana została już cała praca twórcza, jest związane z ustalonym ostatecznym modelem i z rygorystycznym przetestowaniem wniosków" (Spiegelhalt i in., 1995, s. 77; oryg.: „Statistics has no reason for existence except as a catalyst for scientific enquiry in which only the last stage, when all the creative work has already been done, is concerned with a final fixed model and a rigorous test of conclusions").

wymienione, odrębne dla tych badań, cele: eksploracyjny i konfirmacyjny. Przesądzają one o tym, że innych właściwości oczekuje się od prób w badaniach eksploracyjnych, a innych – w badaniach konfirmacyjnych. Dotyczy to także reprezentatywności próby – atrybutu, którego rola w badaniach eksploracyjnych, zwłaszcza w naukach ścisłych i przyrodniczych, jest inna niż w badaniach konfirmacyjnych czy opisowych.

Marcia McNutt, w latach 2013–2016 redaktorka naczelna prestiżowego czasopisma „Science", obecnie prezeska amerykańskiej National Academy of Sciences, stwierdziła w wywiadzie sprzed kilku lat, że: „Badanie eksploracyjne raczej odkrywa i analizuje nowe problemy niż koncentruje się na testowaniu istniejących hipotez. Ale badacze niekiedy czują, że powinni swoje badania eksploracyjne zaprezentować w formie testowania hipotez, co jest całkowicie nierzetelne"[4] (Shell, 2016). Dla lepszego rozróżnienia celów badań eksploracyjnych i badań konfirmacyjnych niektórzy badacze sugerują nazywanie badań eksploracyjnych *badaniami generującymi nowe hipotezy* (ang. *hypothesis-generating research*), a badań konfirmacyjnych – *badaniami skoncentrowanymi na testowaniu hipotez* (ang. *hypothesis-testing research*). W szczególności statystyczna weryfikacja hipotez – mająca charakter dedukcyjny (z wyjątkiem podejścia bayesowskiego; zob. Szreder, 2019) – nie może być uważana za badanie eksploracyjne. Nie oznacza to jednak, że wnioskowanie statystyczne nie tworzy nowej wiedzy, np. poprzez estymację parametrów populacji, opis zależności między cechami lub falsyfikację hipotez. W całym przedsięwzięciu badawczym, w szczególności w naukach przyrodniczych i medycznych, kognitywna rola wnioskowania statystycznego nie jest jednak decydująca (Tong, 2019). Najważniejsze są cele i badania wyjaśniające pierwotne przyczyny i mechanizmy oddziaływania na siebie zmiennych, a więc rezultaty dociekań wywodzących się z nauk merytorycznie właściwych dla danego zagadnienia badawczego. Wnioskowanie statystyczne jest włączone w ten proces badawczy w jednej z ostatnich faz. Nie powinno więc dziwić, że inaczej na rolę wnioskowania statystycznego – i w dalszej kolejności na wymogi odnoszące się do właściwości próby – patrzą naukowcy, których domeną są badania eksploracyjne, a inaczej badacze specjalizujący się w zastosowaniach wnioskowania statystycznego.

O ile reprezentatywność próby jest uznawana za warunek *sine qua non* wiarygodności niewyczerpujących badań społecznych i ekonomicznych, o tyle w naukach medycznych i przyrodniczych kategoria ta ma inny status. Reprezentatywność nie zawsze jest w nich wskazana lub pożądana. Interesującą dyskusję na ten temat zamieszczono w 2013 r. w czasopiśmie „International Journal of Epidemiology". Rozpoczyna ją artykuł Rothmana i in. (2013) o sugestywnym i nieco zaskakującym tytule *Why representativeness should be avoided* (Dlaczego powinno się unikać reprezentatywności). Autorzy przekonują, że próba reprezentatywna stanowi dobrą podstawę do

[4] Oryg.: „An exploratory study explores new questions rather than tests an existing hypothesis. But scientists have felt that they had to disguise an exploratory study as hypothesis testing and that is totally dishonest".

wnioskowania statystycznego o ściśle zdefiniowanej populacji, ale nie upoważnia do dalej idących uogólnień, charakterystycznych dla poszukiwań stricte naukowych. Eksploracyjne cele nauki obejmują niezmienne prawa natury wykraczające poza daną zbiorowość, którą statystyk musi dokładnie określić przed pobraniem próby. „Błędem jest myślenie, że wnioskowanie statystyczne jest tym samym co wnioskowanie naukowe"[5] (Rothman i in., 2013, s. 1013). Jeżeli cel badania nie ogranicza się do opisu stanu i cech pewnej populacji (co ma miejsce w przypadku wielu badań społecznych), lecz obejmuje wykorzystanie metod eksperymentalnych, to próba reprezentatywna często nie jest w ogóle brana pod uwagę. Immunolodzy wykonujący eksperymenty na zwierzętach (łac. *in vivo*) dobierają zwykle próby silnie niereprezentatywne, homogeniczne, składające się z osobników mających podobne geny, żyjących w tych samych warunkach i korzystających z identycznej diety. Cechy immunologiczne tych zwierząt mogą nie być takie same jak u ludzi, ale poszukiwanie dróg uogólnień wyników tych badań na zbiorowości ludzkie wiąże się z odkrywaniem zależności między kontrolowanymi czynnikami i środowiskiem życia a odpornością organizmów. Decydującą rolę mają tu do odegrania nauki chemiczne i biologiczne. O takich próbach mówi się, że nie są reprezentatywne w oszacowaniach (ocenach) odnoszących się do danej populacji, lecz na podstawie szerszej wiedzy o mechanizmach funkcjonowania organizmów zwierząt i ludzi stwierdza się, że mogą być reprezentatywne, jeśli chodzi o interpretację. Dlatego w naukach eksperymentalnych stosowane jest niekiedy rozróżnienie między *uogólnianiem ocen* (ang. *generalisability of estimate*) i *uogólnianiem interpretacji wniosków* (ang. *generalisability of interpretation*; zob. Rudolph i in., 2023).

W badaniach medycznych na nieadekwatność postulatu reprezentatywności próby zwraca się uwagę również w nieco innym kontekście. Otóż prawidłowości i współzależności cech zaobserwowane w próbie reprezentatywnej odnoszą się do całej populacji, którą próba reprezentuje, ale w poszczególnych grupach osób (np. grupach wieku) mogą się istotnie różnić. Samo stwierdzenie, że określona terapia jest skuteczna dla zróżnicowanych jednostek zbiorowości, znajdujących swoje odzwierciedlenie w próbie reprezentatywnej, nie oznacza, że jest ona wskazana dla każdej podgrupy tej zbiorowości. Ogólny efekt danej terapii jest bowiem uśrednionym efektem, ważonym rozkładami liczebności poszczególnych podgrup jednostek zbiorowości. Ważniejsze w naukach medycznych jest odkrycie oddziaływania rozważanej terapii w wyszczególnionych podgrupach, np. u dzieci, osób dorosłych i starszych.

[5] Oryg.: „The mistake is to think that statistical inference is the same as scientific inference". Inni dyskutanci prezentują podobny, chociaż nieco bardziej wyważony punkt widzenia na kwestię reprezentatywności próby. Sprowadza się on do stwierdzenia stanowiącego tytuł jednego z tekstów w tej dyskusji: „Reprezentatywność zwykle nie jest konieczna i często powinno się jej unikać" (Richiardi i in., 2013). Wydaje się jednak, że najwłaściwsza jest konkluzja Ebrahima i Smitha (2013), że nie należy ani unikać reprezentatywności, ani jej bezkrytycznie adoptować, lecz każdorazowo oceniać jej wartość w zaprojektowanym przedsięwzięciu badawczym.

Jednak o tym, czy w określonym badaniu naukowym pożądana jest próba reprezentatywna, decyduje nie dziedzina badań, lecz cel badawczy. W podanym wyżej przykładzie z zakresu medycyny próba reprezentatywna będzie preferowana wtedy, gdy cel badań będzie miał charakter opisowy (deskryptywny), a nie eksploracyjny. Typową sytuacją tego rodzaju jest charakterystyka natężenia występowania chorób lub czynników ryzyka zachorowań na określonym terenie lub w zbiorowości określonej pewnymi cechami. Warto przy tym zauważyć, że wyniki takiego badania opisowego nie tylko wzbogacają wiedzę o rozpowszechnianiu się danej choroby, ale zidentyfikowane w nim prawidłowości statystyczne mogą stanowić podstawę formułowania nowych, oryginalnych hipotez badawczych. Nie są więc te dwa cele – opisowy i eksploracyjny – tak bardzo oddalone od siebie, jak mogłyby sugerować wskazane wcześniej klasyfikacje.

Abstrahując w dalszej części opracowania od nauk eksperymentalnych, warto zauważyć, że w naukach społecznych i ekonomicznych reprezentatywność próby bywa dość często utożsamiana z losowym wyborem próby. W rzeczywistości zaś losowanie jest jednym, ale nie jedynym sposobem uzyskania próby upoważniającej do uogólnień na całą populację. Wiele zależy od tego, jaką wstępną wiedzą o populacji dysponuje badacz, zanim podejmie decyzję o sposobie próbkowania. Wartość tej apriorycznej wiedzy o poddanej badaniu populacji ma decydujące znaczenie w rozstrzygnięciu nie tylko o tym, jaką techniką losowania należy się posłużyć, lecz czy w ogóle wybór losowy jest w świetle tej wiedzy najlepszym rozwiązaniem. I mimo że coraz więcej wiemy o populacjach poddawanych badaniom statystycznym, to losowy wybór próby był i pozostaje najpowszechniejszą metodą próbkowania.

3. Losowość a reprezentatywność próby

Z losowością mamy do czynienia w dwojakim znaczeniu. Z jednej strony otaczają nas naturalne zjawiska i procesy, na których konsekwencje nie mamy żadnego wpływu, takie jak burze, trzęsienia ziemi czy wybuchy wulkanów, których czas i miejsce wystąpienia oraz skutki uznajemy za losowe. Tak rozumiana losowość jest nam narzucona przez naturę, a dokładniej – wynika z niedostatecznej wiedzy na temat przyczynowo-skutkowych mechanizmów powstawania określonych zjawisk i zdarzeń. Z drugiej zaś strony losowość – a właściwie wybór losowy – włączamy do naszych działań intencjonalnie dla osiągnięcia określonego celu. Innymi słowy, by rozwiązać pewien problem, który sam w sobie nie musi mieć żadnych atrybutów stochastycznych, odwołujemy się niekiedy do wyboru losowego – losowania. Celem jest zwykle dokonanie takiego wyboru spośród określonego zestawu opcji, który będzie wolny od wszelkiego racjonalizmu i subiektywizmu człowieka, np. losowanie sędziów do rozpatrywania spraw napływających do sądów. Chodzi więc o taki mechanizm wyboru, który nie jest w żaden

sposób zdeterminowany przez czynniki dające się przypisać czyjemuś racjonalnemu działaniu, ludzkim sympatiom czy emocjom. Dowlen (2008) określa taki wybór mianem *aracjonalny* dla podkreślenia tego, że nie jest on ani racjonalny (oparty na przesłankach rozumowych), ani irracjonalny (podjęty z pobudek emocjonalnych).

Współcześnie można zaobserwować rosnącą popularność włączania mechanizmu losowego do różnych dziedzin życia. Losuje się nie tylko sędziów do prowadzenia poszczególnych spraw w wymiarze sprawiedliwości, recenzentów w przewodach na stopnie i tytuły naukowe, terminy rozpatrywania wiz dla obcokrajowców, lecz także przedstawicieli społeczeństwa do sformowania paneli dyskusyjnych i doradczych dla sprawujących władzę w danym kraju lub regionie. W ostatnich latach szczególnie popularne jest korzystanie z opinii losowo wybranych przedstawicieli społeczeństwa do pracy zespołowej i wypracowania stanowiska w różnych ważnych sprawach gospodarczych, społecznych i związanych z ochroną środowiska naturalnego. Korzysta z tej możliwości Parlament Europejski w Strasburgu (Pub Affairs Bruxelles, 2021) oraz niektóre kraje: Francja[6], Wielka Brytania i Niemcy, zainteresowane opiniami paneli obywatelskich (ang. *citizens' panels*). Podmioty te uznają losowanie za skuteczny mechanizm pozwalający formować panele reprezentatywne (stanowiące adekwatną reprezentację) dla zbiorowości danego kraju lub regionu. Warto jednak postawić ogólniejsze pytanie: czy losowanie jest najlepszym – a może jedynym – sposobem na uzyskanie próby reprezentatywnej z określonej zbiorowości?

Relacje między pojęciami *reprezentatywny* i *losowy* są ważne, ponieważ dotykają samych fundamentów wnioskowania statystycznego. Warto zdać sobie sprawę z tego, że formalne wnioskowanie statystyczne jest nieodłącznie związane z aktem losowania z dwóch zasadniczych powodów:

1. W wyniku losowania uzyskuje się przeciętnie – w hipotetycznym długim ciągu powtarzalnych losowań – próbę, w której rozkłady cech statystycznych są zbliżone do rozkładów tych cech w populacji, czyli próbę dobrze odzwierciedlającą strukturę populacji, bez względu na to, które z tych cech są przedmiotem zainteresowania badacza.
2. Probabilistyczny (matematyczny) model wnioskowania statystycznego, odwołujący się do częstościowej interpretacji prawdopodobieństwa, wymaga losowego generowania zdarzeń (obserwacji) w próbie, co z kolei umożliwia badaczowi wyrażenie w kategoriach probabilistycznych wielkości błędów oszacowań lub ryzyka podjęcia błędnej decyzji w odniesieniu do testowanych hipotez.

[6] We Francji jednym z pierwszych ogólnokrajowych paneli obywatelskich, stworzonych z inicjatywy prezydenta Emmanuela Macrona po protestach tzw. żółtych kamizelek w 2019 r., była 150-osobowa grupa losowo wybranych osób, której zadaniem było wypracowanie stanowiska w sprawie reform związanych ze zmianami klimatycznymi, a także z postulatami zwiększenia demokracji bezpośredniej we Francji (Trian, 2021).

Wśród podstawowych przesłanek losowego wyboru próby podkreśla się na ogół pierwszy z wyżej wymienionych powodów – gwarancję reprezentatywności próby (przeciętnie). Rzadziej wspomina się o tym, że równie ważna jest możliwość przypisania wynikom wnioskowania probabilistycznych miar dokładności uzyskanych ocen. W statystyce – tak jak w każdej innej dyscyplinie naukowej – znaczenie mają nie tylko wyniki przeprowadzonych badań, lecz także towarzyszące im błędy. Dlatego rozpatrując inne (nieprobabilistyczne) sposoby uzyskania próby będącej dobrą reprezentacją populacji, należy mieć na uwadze także to, czy umożliwiają one oszacowanie błędów wyników, które miałoby odpowiednie uzasadnienie naukowe.

W relacjach między kategoriami *losowości* i *reprezentatywności* istotne znaczenie ma to, że nie każda konkretna próba wylosowana ze zdefiniowanej wcześniej populacji będzie wykazywała taki sam stopień podobieństwa rozkładów cech co rozkłady analogicznych cech w populacji. Innymi słowy, niektóre próby będą miały strukturę bardziej, a niektóre mniej przypominającą strukturę populacji. Jedne będą bardziej, a inne mniej reprezentatywne. Przeciętnie natomiast mechanizm indywidualnego, niezależnego losowania obserwacji z populacji gwarantuje uzyskanie próby o rozkładach cech takich jak rozkłady tych cech w populacji[7]. To, że w statystyce posługujemy się różnymi schematami wyboru próby – a nie wyłącznie losowaniem prostym indywidualnym – wiąże się z dążeniem do uzyskania nie tylko przeciętnie, lecz także w konkretnym losowaniu, takiej próby, której charakterystyki będą jak najbardziej zbliżone do analogicznych charakterystyk w populacji. I mimo że matematyczny model wnioskowania oparty jest na definicji prostej próby losowej, to w praktyce, dla zwiększenia szans na otrzymanie w konkretnym losowaniu próby o znacznym stopniu reprezentatywności, korzysta się z kilku innych schematów losowania. Ich wspólną cechą jest włączenie do każdego z tych schematów informacji wstępnych (a priori), które badacz posiadł przed przystąpieniem do wyboru próby. Im bogatsza jest wstępna wiedza o populacji – np. pochodząca z wcześniejszych badań na temat tych samych lub innych jej cech albo wynikająca z informacji zawartych w rejestrach administracyjnych lub innych zbiorach danych – tym większe są szanse na wysoki stopień reprezentatywności próby, gdy wiedza ta zostanie wykorzystana w losowaniu. Dobrą tego ilustracją może być znany i z powodzeniem stosowany w wielu badaniach schemat losowania warstwowego, w którym na podstawie apriorycznych informacji o populacji jej jednostki zostają pogrupowane w homogeniczne warstwy (powarstwowane),

[7] W statystyce podejście, w którym zakłada się, że źródłem losowości ocen parametrów populacji jest przyjęty plan losowania, a wartości cech w populacji nie są losowe (czyli są ustalone), nazywa się *podejściem randomizacyjnym*. Obok niego istnieje drugie podejście, nazywane *modelowym*, którym się w tym opracowaniu nie zajmujemy. Jego najważniejszym założeniem jest to, że wektor wartości badanej cechy traktowany jest jako wektor realizacji zmiennych losowych, a źródłem losowości jest model nadpopulacji (zob. np. Bracha, 1996; Steczkowski, 1995).

a losowanie odbywa się odrębnie z każdej warstwy. Schemat ten pozwala na uzyskanie próby, której struktura ze względu na cechy warstwujące jest identyczna ze strukturą populacji. Zwłaszcza w przypadku populacji silnie niejednorodnych jest to jedna z dobrze znanych możliwości zwiększenia efektywności wnioskowania.

Wiedza o populacji, którą badacz posiadł przed wyborem jednostek do próby, jest tym źródłem informacji, które odgrywa najważniejszą rolę w dążeniu do zapewnienia możliwie największego podobieństwa struktury próby do struktury populacji. Jeżeli dla badacza wystarczające jest ograniczenie tego podobieństwa do ściśle określonych cech (np. w populacji przedsiębiorstw do rodzaju działalności według PKD, wielkości zatrudnienia lub formy organizacyjnoprawnej), to może się on zadowolić próbą, której struktura będzie zgodna ze strukturą populacji tylko ze względu na te cechy. Spełnienie postulatu identycznej lub podobnej struktury próby i populacji w odniesieniu do kilku konkretnych cech statystycznych jest zadaniem prostszym niż uzyskanie próby, która odzwierciedlałaby strukturę populacji ze względu na wszystkie cechy. Przykładem takiego podobieństwa ograniczonego do wybranych cech próby i populacji jest technika wyboru kwotowego – jedna z nieprobabilistycznych technik próbkowania. Za jej podstawową zaletę uważa się zgodność rozkładów cech (charakterystyk kontrolnych) wskazanych przez badacza na wstępie z rozkładami tych cech w populacji. Znając – zwykle na podstawie danych wtórnych – strukturę populacji ze względu na wyróżnione cechy, dokonuje się nielosowego wyboru jednostek do próby w takich proporcjach, aby udziały jednostek o określonych wariantach tych cech były takie same w próbie i w populacji. O takich próbach często mówi się, że są reprezentatywne ze względu na określony, ograniczony zestaw cech. Wynika z tego, że kategoria reprezentatywności pojawia się także w odniesieniu do nielosowych sposobów wyboru próby.

4. Propozycja zdefiniowania *próby reprezentatywnej*

W literaturze anglojęzycznej definicje *próby reprezentatywnej* przybierają raczej formę definicji realnej niż nominalnej (zob. np. Babbie, 2004). Definicje te nie objaśniają charakteru takiej próby, sposobu jej wyboru lub struktury, lecz koncentrują się na jej atrybucie praktycznej użyteczności. Na przykład Lavrakas (2008) stwierdza, że próba reprezentatywna daje możliwość wiarygodnej ekstrapolacji wyników na populację, którą reprezentuje[8]. Podobnie próbę reprezentatywną charakteryzują Rudolph i in. (2023), pisząc, że jest to próba, z której uzyskane oszacowania mogą być uogólniane

[8] Oryg.: „A representative sample is one that ensures external validity in relationship to the population of interest the sample is meant to represent".

na populację, z której pochodzi[9]. Definicje te słusznie podkreślają najważniejszy walor próby reprezentatywnej – możliwość uogólniania na jej podstawie wyników na reprezentowaną populację – ale nie są one, jak się wydaje, dostatecznie precyzyjne i jednoznaczne.

Przytoczone definicje w niewystarczający sposób charakteryzują samą próbę, ograniczając się jedynie do jej zastosowań w procesie wnioskowania statystycznego. Z tego powodu, odwołując się wyłącznie do tych definicji, trudno byłoby przekonać badacza wykorzystującego na przykład próbę kwotową, że jego próba nie spełnia warunków definicji próby reprezentatywnej. Korzystający z prób kwotowych twierdzą dość często, że posługują się próbami reprezentatywnymi. Tak jednak nie jest, ponieważ samo kontrolowanie zgodności struktury próby ze strukturą populacji jedynie w odniesieniu do wyselekcjonowanych zmiennych (charakterystyk) jest niewystarczające, aby próbę nazwać *reprezentatywną*. Można natomiast powiedzieć, że jest reprezentatywna dla danej populacji ze względu na kilka konkretnych zmiennych, uznanych za ważne dla celu badań, a które w wyborze kwotowym pełnią funkcję charakterystyk kontrolnych. Podnoszony nieraz przez zwolenników próbkowania kwotowego argument, że dla badacza kluczowa jest zgodność właśnie ze względu na zmienne najważniejsze dla celu badań, co ma czynić tę próbę reprezentatywną, jest trudny do obrony. A to dlatego, że ten rodzaj rozumowania – czyli przekonanie o tym, że badacz potrafił bezbłędnie wskazać wszystkie istotne dla celu badania cechy jednostek populacji – był w przeszłości częstą przyczyną poważnych błędów wnioskowania na podstawie prób kwotowych (Szreder, 2010, 2012). Okazywało się mianowicie, że niewystarczająco poznane i niedoceniane inne cechy, których rozkłady w populacji nie są w żaden sposób weryfikowane z rozkładami z próby (ponieważ nie należą do charakterystyk kontrolnych), powodowały niereprezentatywność próby, mimo że była ona reprezentatywna w odniesieniu do kilku wskazanych na wstępie zmiennych. W zdefiniowaniu próby reprezentatywnej trzeba więc to rozróżnienie brać pod uwagę.

Próby nieprobabilistyczne, którym przypisuje się walor reprezentatywności w odniesieniu do kilku określonych przez badacza cech populacji, stanowią w praktyce podstawę uogólnień, o których mowa w przytoczonych wyżej definicjach. Za niedostatek tych definicji trzeba więc uznać dopuszczanie powiązania próby reprezentatywnej z bliżej niesprecyzowanymi sposobami uogólnień na populację. Co prawda w niektórych definicjach akcentuje się „upoważnione, wiarygodne uogólnienia" (ang. *external validity*), ale wydaje się to wciąż za mało jak na precyzję pożądaną w każdej definicji. Przede wszystkim wnioskowanie na podstawie prób nielosowych nie daje możliwości wiarygodnego oszacowania błędów ocen. Nie istnieje bowiem matematyczny

[9] Oryg.: „We define a study sample to be representative of a well defined target population if the results estimated in that sample are generalisable to the target population".

model dla próbkowania nieprobabilistycznego, który dawałby szansę na wiarygodne liczbowe określenie błędów wnioskowania. Dlatego ogólna, uniwersalna definicja próby reprezentatywnej powinna uwzględniać także i ten aspekt. Innymi słowy, proponuję, aby definicja próby reprezentatywnej odnosiła się szczegółowo do dwóch aspektów:
- zgodności struktury próby ze strukturą populacji;
- atrybutów wnioskowania statystycznego na podstawie próby.

Oczywiście w takiej definicji nie powinno być mowy o wyróżnionych cechach lub zmiennych, których zgodności rozkładów oczekuje się w próbie i w populacji. Powinna to być definicja na tyle uniwersalna, aby poprawnie wskazywać na właściwości próby niezależnie od tego, które rozkłady zmiennych i w jakim stopniu interesują badacza lub odbiorcę wyników wnioskowania.

Proponuję następującą definicję próby reprezentatywnej:

> Próba reprezentatywna to taka, której struktura jest identyczna ze strukturą populacji lub do niej zbliżona oraz która może stanowić podstawę uogólnień wyników na populację i wiarygodnej oceny ich błędów.

Tak definiowana próba reprezentatywna odpowiada z jednej strony dotychczasowemu rozumieniu tej kategorii, a z drugiej nie pozwala na to, by krótkim terminem *próby reprezentatywne* nazywać te próby, których struktury są identyczne ze strukturą populacji w odniesieniu jedynie do niektórych cech. Te ostatnie powinny być nazywane *próbami reprezentatywnymi ze względu na wyszczególnione zmienne,* a nie *reprezentatywnymi* w ogóle (zgodnie z zaproponowaną wyżej definicją). W niczym nie deprecjonuje to takich prób, w tym m.in. kwotowych, a jedynie porządkuje terminy używane w statystyce. Próby reprezentatywne ze względu na konkretne zmienne (cechy) mogą być próbami losowymi i nielosowymi. Popularny w empirycznych zastosowaniach wnioskowania statystycznego mechanizm ważenia i kalibracji wag stanowi dobry przykład dążenia do uzyskania właśnie tego rodzaju „ograniczonej" reprezentatywności prób[10].

Warto przywołać jeszcze jeden termin zbliżony brzmieniem do *próby reprezentatywnej,* a mianowicie wyrażenie *próba reprezentacyjna*. Nie jest on co prawda często stosowany, a jego istnienie wiąże się przede wszystkim z badaniami reprezentacyjnymi, oraz z działem statystyki zwanym *metodą reprezentacyjną*. O ile termin *badanie reprezentacyjne* jest jasno określony w literaturze, o tyle *próba reprezentacyjna* jedynie zastępuje i powiela termin *próba losowa*. Jest więc terminem zbędnym, który może prowadzić do nieporozumień we wnioskowaniu statystycznym. Z samej definicji

[10] Szeroko na temat sposobów ważenia i kalibracji oraz zastosowań tych mechanizmów w statystyce piszą Kozłowski i Szreder (2020), a także Szymkowiak (2019).

badania reprezentacyjnego wynika bowiem, że jest to rodzaj badań niewyczerpujących opartych na próbach losowych. Tak rozumieją to Bracha (1998, s. 18): „Częściowe badanie statystyczne, do którego próba została wybrana zgodnie z pewnym planem losowania, nazywamy badaniem reprezentacyjnym", Steczkowski (1995, s. 11): „Metoda reprezentacyjna polega na losowym doborze próby (reprezentacji) ze skończonej zbiorowości generalnej, opisie tej próby za pomocą charakterystyk statystycznych, a następnie na uogólnieniu otrzymanych wyników na zbiorowość generalną, z której próba ta pochodzi" czy Rószkiewicz i in. (2013, s. 143): „Wnioskowaniem na podstawie prób losowych, które cechuje reprezentatywność statystyczna, zajmuje się nauka zwana metodą reprezentacyjną, zaś badanie oparte na próbie losowej określane jest mianem badania reprezentacyjnego". Wynika więc z tego jednoznacznie, że próbą wykorzystywaną w badaniach reprezentacyjnych jest próba losowa. Nie ma potrzeby nazywania jej *próbą reprezentacyjną* ani wprowadzania takiej kategorii do wnioskowania statystycznego.

5. Podsumowanie

Próba reprezentatywna jest jedną z podstawowych kategorii występujących w zastosowaniach wnioskowania statystycznego. W teorii wnioskowania – a ściślej w jego modelu matematycznym – kategoria ta nie jest obecna, ponieważ zastępuje ją kluczowe w probabilistyce pojęcie *próby losowej*. Dlatego potrzeba precyzyjnego zdefiniowania *próby reprezentatywnej* odnosi się nie do teorii, lecz do empirycznych zastosowań wnioskowania statystycznego. Potrzeba ta jest ważniejsza w zastosowaniach ekonomicznych i społecznych, w których cele wnioskowania mają głównie charakter deskryptywny lub konfirmacyjny. Mniejsze jest natomiast jej znaczenie w badaniach eksperymentalnych – w naukach ścisłych i przyrodniczych – dla których najważniejsze są cele eksploracyjne. W tych ostatnich rzadko postuluje się planowanie i realizację eksperymentów na próbach reprezentatywnych dla całej populacji.

Jednoznaczne rozumienie pojęcia *próby reprezentatywnej* jest ważne przede wszystkim dlatego, że pozwala na pełną i poprawną interpretację zarówno wyników badania próbkowego, jak i błędów, którymi te wyniki mogą być obciążone. W artykule zaproponowano definicję *próby reprezentatywnej,* która bierze pod uwagę zgodność struktury próby ze strukturą populacji oraz możliwości takiego wnioskowania na podstawie próby, w którym ocenom uzyskanym z próby towarzyszą naukowo uzasadnione oszacowania wielkości błędów. Autor postuluje, aby próby, które wykazują zbieżność rozkładów tylko niektórych cech z ich rozkładami populacyjnymi, nazywać *próbami reprezentatywnymi ze względu na wyszczególnione zmienne*.

Bibliografia

Analytical Methods Committee. (2016). Representative sampling? Views from a regulator and a measurement scientist. *Analytical Methods*, *8*(24), 4783–4784. https://doi.org/10.1039/c6ay90077a.

Babbie, E. (2004). *Badania społeczne w praktyce* (tłum. W. Betkiewicz, M. Bucholc, P. Gadomski, J. Haman, A. Jasiewicz-Betkiewicz, A. Kloskowska-Dudzińska, M. Kowalski, M. Mozga). Wydawnictwo Naukowe PWN.

Bracha, C. (1996). *Teoretyczne podstawy metody reprezentacyjnej*. Wydawnictwo Naukowe PWN.

Bracha, C. (1998). *Metoda reprezentacyjna w badaniu opinii publicznej i marketingu*. Efekt.

Dowlen, O. (2008). *The political potential of sortition. A study of the random selection of citizens for public office*. Imprint Academic.

Ebrahim, S., Smith, G. D. (2013). Commentary: Should we always deliberately be non-representative?. *International Journal of Epidemiology*, *42*(4), 1022–1026. https://doi.org/10.1093/ije/dyt105.

Kozłowski, A., Szreder, M. (2020). *Informacje spoza próby w badaniach statystycznych*. Wydawnictwo Uniwersytetu Gdańskiego.

Lavrakas, P. J. (red.). (2008). Representative sample. W: *Encyclopedia of Survey Research Methods* (t. 2). SAGE Publications. https://doi.org/10.4135/9781412963947.

Pub Affairs Bruxelles. (2021). *Conference on the Future of Europe: European Citizens' Panels set to start*. https://www.pubaffairsbruxelles.eu/eu-institution-news/conference-on-the-future-of-europe-european-citizens-panels-set-to-start/.

Reichenbach, H. (1951). *The Rise of Scientific Philosophy*. University of California Press. https://doi.org/10.2307/jj.8501244.

Richiardi, L., Pizzi, C., Pearce, N. (2013). Commentary: Representativeness is usually not necessary and often should be avoided. *International Journal of Epidemiology*, *42*(4), 1018–1022. https://doi.org/10.1093/ije/dyt103.

Rothman, K. J., Gallacher, J. E. J., Hatch, E. E. (2013). Why representativeness should be avoided. *International Journal of Epidemiology*, *42*(4), 1012–1014. https://doi.org/10.1093/ije/dys223.

Rószkiewicz, M., Perek-Białas, J., Węziak-Białowolska, D., Zięba-Pietrzak, A. (2013). *Projektowanie badań społeczno-ekonomicznych. Rekomendacje i praktyka badawcza*. Wydawnictwo Naukowe PWN.

Rudolph, J. E., Zhong, Y., Duggal, P., Mehta, S. H., Lau, B. (2023). Defining representativeness of study samples in medical and population health research. *BMJ Medicine*, *2*(1), 1–7. https://doi.org/10.1136/bmjmed-2022-000399.

Shell, E. R. (2016). Hurdling Obstacles: Meet Marcia McNutt, scientist, administrator, editor, and now National Academy of Sciences president. *Science*, *353*(6295), 116–119. https://doi.org/10.1126/science.353.6295.116.

Spiegelhalt, D. J., Grieve, A. P., Lindley, D. V., Copas, J. B., Cairns, A. J. G., Chatfield, C., Box, G., Cox, D., Tukey, J. W., Raftery, A. E., Aitkin, M., Barnard, G. A., Donnelly, P., Ehrenberg, A. S. C., Gelfand, A. E., MaHlick, B. K., Gelman, A., Meng, X.-L., Glasbey, C. A., …, Zellner, A. (1995). Discussion of the paper by Draper. *Journal of the Royal Statistical Society, Series B (Methodological), 57*(1), 71–97. https://doi.org/10.1111/j.2517-6161.1995.tb02016.x.

Steczkowski, J. (1995). *Metoda reprezentacyjna w badaniach zjawisk ekonomiczno-społecznych*. Wydawnictwo Naukowe PWN.

Szreder, M. (2010). *Metody i techniki sondażowych badań opinii* (wyd. II zmienione). Polskie Wydawnictwo Ekonomiczne.

Szreder, M. (2012). Wybór próby badawczej – aspekt etyczny i naukowy. *Marketing i Rynek,* (11), 24–27.

Szreder, M. (2019). Istotność statystyczna w czasach big data. *Wiadomości Statystyczne. The Polish Statistician, 64*(11), 42–57. https://doi.org/10.5604/01.3001.0013.7583.

Szymkowiak, M. (2019). *Podejście kalibracyjne w badaniach społeczno-ekonomicznych*. Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu. https://doi.org/10.18559/978-83-66199-89-7.

Tong, C. (2019). Statistical Inference Enables Bad Science; Statistical Thinking Enables Good Science. *The American Statistician*, *73*(S1), 246–261. https://www.tandfonline.com/doi/full/10.1080/00031305.2018.1518264#abstract.

Trian, N. (2021, 1 marca). *Macron's 'direct democracy' to be tested as citizens' panel on climate wraps up*. https://www.france24.com/en/france/20210301-macron-s-direct-democracy-to-be-tested-as-citizens-panel-on-climate-wraps-up.

Wiadomości Statystyczne. The Polish Statistician, 2024, *69*(11), 69–80 https://doi.org/10.59139/ws.2024.11.5
Dyskusje. Recenzje. Informacje / Discussions. Reviews. Information

Międzynarodowa konferencja „Privacy in Statistical Databases 2024"

International conference *Privacy in Statistical Databases 2024*

W dniach 25–27 września 2024 r. w Antibes Juan-les-Pins we Francji odbyła się międzynarodowa konferencja "Privacy in Statistical Databases 2024". To jedenasta edycja organizowanego od dwudziestu lat wydarzenia służącego wymianie doświadczeń w zakresie ogólnoświatowych badań dotyczących ochrony prywatności w statystycznych bazach danych i wytyczaniu nowych kierunków w tej dziedzinie. Organizatorem konferencji – odbywających się zazwyczaj co dwa lata w różnych częściach Europy – jest grupa badawcza CRISES (University of Rovira i Virgili, Hiszpania) działająca pod kierunkiem Josepa Dominga-Ferrera. Rolę gospodarza tegorocznych obrad pełnił francuski EURECOM, czołowa instytucja naukowo-badawcza zajmująca się technologiami informacyjno-komunikacyjnymi. Konferencję otworzyła Melek Önen, przedstawicielka gospodarzy. Następnie Domingo-Ferrer przypomniał historię wydarzenia i krótko omówił cele konferencji, podkreślając znaczenie ochrony prywatności w przypadku danych georeferencyjnych. Podczas konferencji odbyło się dziewięć sesji tematycznych dotyczących różnych aspektów kontroli ujawniania danych statystycznych, w tym generowania danych syntetycznych i stosowania narzędzi sztucznej inteligencji.

Pierwsza sesja tematyczna *Disclosure risk assessment*, której przewodniczyli Marieke de Vries (Statistics Netherlands, Holandia) i Matthias Templ (University of Northwestern Switzerland, Szwajcaria, i Technische Universität Wien, Austria) dotyczyła oceny ryzyka ujawnienia (zagrożenia identyfikacją konkretnej jednostki lub powiązania z nią jej wrażliwego atrybutu).

Elizabeth Green, Felix Ritchie i Paul White (University of the West of England, Wielka Brytania) w wystąpieniu *The Statbarn: A New Model for Output Statistical Disclosure Control* przedstawili nowy model zarządzania wyjściowymi danymi statystycznymi (tzw. koncepcję stajni statystycznej) jako ramę klasyfikacyjną wszystkich pojęć statystycznych z punktu widzenia kontroli ujawniania danych. Prelegenci zaproponowali stworzenie bloków bezpiecznych i niebezpiecznych danych statystycznych. Z uwagi na to, że binarne rozróżnienie okazało się w tym przypadku niewystarczające, zasugerowali zastosowanie reguł grupowych i statystyk homologicznych (w tym korelacji i współczynników powiązań), na podstawie których określa się ryzyko ujawnienia danych. Omówili także postęp prac w tej dziedzinie.

Guillermo Navarro-Arribas (Universitat Autonoma de Barcelona, Hiszpania) i Vincenç Torra (Umeå University, Szwecja) w referacie *Attribute Disclosure Risk in Smart Meter Data: A Position Paper* rozważali metody analizy i ochrony danych pochodzących z inteligentnych liczników zużycia energii i gazu. Badali ryzyko ujawnienia atrybutów w pewnych zbiorach mikrozagregowanych danych i ukazali miary wrażliwości w tym zakresie oparte na regułach *(n,k)*-dominacji i *p%*. Omówili także związane z tym zagrożenia, takie jak: duża liczba komórek wrażliwych, wysokie ryzyko ujawnienia w przypadku danych poddanych mikroagregacji i możliwość ataków zewnętrznych.

Wystąpienie Davida Sáncheza, Najeeba Jebreela, Josepa Dominga-Ferrera i Alberta Blanki-Justicii (University of Rovira i Virgili, Hiszpania) oraz Krisha Muralidhara (University of Oklahoma, Stany Zjednoczone) *An Examination of the Alleged Privacy Threats of Confidence-Ranked Reconstruction of Census Microdata* dotyczyło zagrożeń utraty prywatności w kontekście rekonstrukcji danych spisowych opartej na rangowaniu ufności. Prelegenci przypomnieli, że w 2020 r. z powodu niskiej efektywności i zagrożenia atakami amerykańskie Biuro Spisów zrezygnowało z tradycyjnej wymiany rekordów na rzecz różnicowania prywatności. Nowy typ ataku polega na określaniu ufności, że rekonstruowany rekord znajdzie się w danych oryginalnych. Autorzy empirycznie weryfikowali stosowane porządkowanie ryzyka ujawnienia dla rekordów ocenianego w trakcie wielokrotnych prób rekonstrukcji danych wrażliwych. Inicjalizacja takiego procesu może bazować na wiedzy o cechach rozkładu lub mieć charakter losowy. W czasie prób porównywano m.in. częstość występowania poszczególnych wartości i korelację między rekordami, uwzględniając bloki rekordów podobnych i poziomy przestrzenne. Referenci stwierdzili m.in. słabą efektywność w wykrywaniu możliwości wtórnej identyfikacji lub tworzeniu nieistniejących rekordów.

Simon Xi Ning Kolb, Jui Andreas Tang i Sarah Giessing (Federal Statistical Office of Germany, Niemcy) w wystąpieniu *Synthetic Data: Comparing Utility and Risk in Microdata and Tables* dokonali porównania miar użyteczności i ryzyka ujawnienia dla mikrodanych i tablic. Przedstawili metody oceny użyteczności mikrodanych oparte na współczynnikach skłonności (ang. *propensity score*) w przypadku modeli predykcyjnych, a w przypadku tablic – na odległości Hellingera. Ryzyko ujawnienia dla mikrodanych zostało wyznaczone za pomocą liczby unikatowych par rekordów i homogeniczności grup rekordów (ujawnienie grupowe). W tym celu autorzy zaproponowali współczynnik bazujący na liczbie unikatowych rekordów w zbiorach oryginalnym i syntetycznym. W przypadku danych tabelarycznych ryzyko ujawnienia jest pochodną ryzyka dla mikrodanych, na podstawie których powstały tablice.

Carolina Trindade i Tânia Carvalho (University of Porto, Portugalia), Luís Antunes (University of Porto, TekPrivacy, Portugalia) oraz Nuno Moniz (Lucy Family Institute for Data & Society, University of Notre Dame, Stany Zjednoczone) poświęcili referat *Synthetic Data Outliers: Navigating Identity Disclosure* obserwacjom odstającym w danych syntetycznych, które stanowią główne źródło ryzyka identyfikacji jednostki w możliwym do przeprowadzenia ataku realizowanym poprzez powiązania rekordów. Zaprezentowali miary ryzyka i użyteczności oparte m.in. na pokryciu atrybutów, podobieństwie statystycznym i scenariuszach ataków poprzez powiązania. Autorzy doszli do wniosku, że efektywność ochrony obserwacji odstających zależy w dużej mierze od zastosowanego modelu analitycznego.

Przedmiotem prezentacji Gillian Raab (University of Edinburgh, Wielka Brytania) *Privacy Risk from Synthetic Data: Practical Proposals* były propozycje miar oceny ryzyka ujawnienia atrybutów w danych syntetycznych. Panuje przekonanie, że dane w pełni syntetyczne nie wymagają ochrony, ale część ekspertów dostrzega pewne zagrożenia w tym zakresie. Prelegentka przedstawiła możliwe scenariusze ujawnienia w przypadku danych syntetycznych dotyczących rozwoju regionalnego i wskazała metody oceny ryzyka i użyteczności danych. Należą one do narzędzi rozwijanego pakietu `synthpop` w środowisku R.

Druga sesja tematyczna *Privacy models and concepts*, której przewodniczył Peter-Paul de Wolf (Statistics Netherlands, Holandia), dotyczyła zagadnień związanych z modelami i koncepcjami prywatności.

Giuseppe D'Acquisto (Luiss University, Włochy), Aloni Cohen (University of Chicago, Stany Zjednoczone), Maurizio Naldi (LUMSA University, Włochy) i Kobbi Nissim (Georgetown University, Stany Zjednoczone) poświęcili swoje wystąpienie *From Isolation to Identification* kwestii identyfikacji jednostki przez rekordy izolowane (odróżniające się od innych). Prelegenci zauważyli, że pewien stopień izolacji jest w praktyce nieunikniony, nawet gdy dane są ściśle chronione. Opisali przypadek, gdy izolacja wynikająca z udostępnienia danych może umożliwiać identyfikację jednostki. Sformułowali koncepcję zysku izolacyjnego, zgodnie z którą znalezienie się w próbie jest skorelowane z pewnymi publicznie znanymi atrybutami jednostek. Rozważali odporność na niepełną wiedzę i miarę zysku izolacyjnego. Określili w tym kontekście optymalną strategię i ukazali jej użyteczność (m.in. na przykładzie reguły *k*-anonimowości). Postulowali normalizację miar ujawnienia na podstawie trywialnego punktu odniesienia.

Takumi Sugyiama (Chuo University, Japonia), Hiroto Oosugi (Raksu Co., Ltd., Japonia), Io Yamanaka (Secom Co., Ltd., Japonia) i Kazuhiro Minami (The Institute of Statistical Mathematics, Japonia) w referacie *Utility Analysis of Differentially Private Anonymized Data Based on Random Sampling* dokonali analizy użyteczności

danych z próbek losowych zanonimizowanych przy użyciu różnicowania prywatności. Prelegenci rozważali realizację nierozróżnialności poprzez losowanie próbek, co jest równoważne schematowi Bernoulliego i oceniane za pomocą wiarygodności. Ponadto zaprezentowali metodę wyznaczania parametru różnicowania prywatności przy użyciu algorytmu SafePub w narzędziu anonimizacji ARX i omówili praktyczne problemy związane z realizacją wskazanych czynności (m.in. w świetle eksperymentalnie analizowanych optymalnych parametrów, wskaźników względnej precyzji i innych mierników jakościowych, a także oceny relacji między budżetem prywatności a rozmiarem próby).

Luis Del Vasto Terrientes, Sergio Martínez i David Sánchez (Uniwersytet Rovira i Virgili, Hiszpania) w referacie *Privacy- & Utility-Preserving Data Releases over Fragmented Data Using Differential Privacy via Individual Ranking* zajęli się zagadnieniem udostępniania danych zachowujących użyteczność i prywatność w sytuacji, gdy dane zostały sfragmentowane przez zastosowanie różnicowania prywatności za pomocą indywidualnego rangowania. Prelegenci przedstawili zasady i wskazali zalety tego różnicowania, a także podali przykład jego eksperymentalnego zastosowania w budowaniu modelu systemu bezpiecznego udostępniania danych z zachowaniem ich użyteczności w warunkach fragmentaryzacji.

Trzecia sesja tematyczna *Statistical table protection*, której przewodniczyła Sarah Giessing, dotyczyła ochrony tablic statystycznych. Składały się na nią dwa wystąpienia.

Autorem pierwszego z nich, *Secondary Cell Suppression by Gaussian Elimination: An Algorithm Suitable for Handling Issues with Zeros and Singletons,* był Øyvind Langsrud (Statistics Norway, Norwegia). Referent mówił o wtórnym ukrywaniu komórek poprzez eliminację Gaussa i o sytuacji, gdy ukrycie zer w tablicach częstości nie chroni pozostałych zawartych w nich danych. Następnie zaprezentował algorytm możliwy do zastosowania w przypadku występowania zer i kombinacji z częstością 1, którego główną właściwością jest to, że kolumny poddane ukrywaniu wtórnemu nie mogą zostać całkowicie wyeliminowane. Nie ma też potrzeby tworzenia wirtualnych komórek dla eliminacji zer. Kolumny reprezentujące sumy zer nie mogą być w tym przypadku punktem startowym eliminacji. Autor omówił również realizację rozpatrywanego podejścia dla tablic wielkości, gdy występuje dużo rekordów unikalnych. Zwrócił uwagę, że w tym kontekście można użyć odpowiednich narzędzi programu tau-Argus.

W drugim referacie w tej sesji *Obtaining (ε, δ)-Differential Privacy Guarantees When Using the Poisson Distribution to Synthesize Tabular Data* James Jackson i Brian Francis (Lancaster University, Wielka Brytania), Robin Mitra (UCL – London's Global University, Wielka Brytania) i Ian Dove (Office for National Statistics, Wielka Brytania) przybliżyli zagadnienie uzyskiwania efektywnego probabilistycznego

różnicowania prywatności, gdy do syntetyzacji danych tabelarycznych stosuje się rozkład Poissona. Syntetyzacja jest w tym przypadku przeprowadzana przy użyciu saturowanego modelu zliczeniowego. Użyteczność tego podejścia przedstawiono na przykładzie saturowanego modelu Poissona.

Czwartą sesję tematyczną *Microdata protection*, której przewodniczyła Anna Oganian (National Centre for Health Service, Stany Zjednoczone), poświęconą ochronie mikrodanych otworzył referat Katariiny Perkonoi i Joniego Virty (University of Turku, Finlandia) *Asymptotic Utility of Spectral Anonymization* dotyczący asymptotycznej użyteczności anonimizacji spektralnej. Autorzy omówili m.in. rozkład macierzy wartościami osobliwymi (SVD). Porównali trzy alternatywne podejścia w tym zakresie: pierwsze bazujące na macierzy znaków, drugie – na macierzach ortogonalnych i trzecie – na rozkładzie jednostajnym Haara. Wykazali, że żadna z wymienionych metod nie ma przewagi nad pozostałymi.

James Thomas Brown, Ellen Clayton, Michael Barteny i Bradley Malin (Vanderbilt University, Stany Zjednoczone), Murat Kantarcioglu (University of Texas, Stany Zjednoczone) oraz Yewgeniy Vorobeychik (Washington University in Saint Louis, Stany Zjednoczone) w wystąpieniu *Robin Hood*: *A De-identification Method to Preserve Minority Representation for Disparities Research* nawiązali do metody Robin Hooda, ukrywającej dane w sposób zwiększający oczekiwaną liczbę podejść wymaganych do reidentyfikacji. Referat dotyczył użycia metody deidentyfikacyjnej dla zachowania mniejszościowej reprezentacji dysparytetów. Autorzy, posługując się modelem kontradyktoryjnym, stwierdzili, że deidentyfikacja może maskować i odwracać dysparytety oraz zakłócać reprezentację mniejszościową.

Ostatnią prezentację w tej sesji, *An Optimization Approach to Privacy Preserving Dynamic Data Publishing*, przygotowali Jordi Castro i Adrián Tobar-Nicolau (Universitat Politècnica de Catalunya, Hiszpania) oraz Claudio Gentile (Consiglio Nazionale delle Ricerche, Włochy). W referacie poruszono problematykę podejścia optymalizacyjnego do ochrony prywatności w dynamicznym publikowaniu danych opartym na *m*-niezmienności (mikroagregacja pozwala na podział jednostek na rozłączne skupienia, takie że rozmiar każdego z nich wynosi co najmniej *m*, a jednostki w każdym skupieniu są do siebie możliwie najbardziej podobne, tak aby zminimalizować utratę informacji). W referacie podkreślono specyficzny – z punktu widzenia kontroli ujawniania danych – charakter cyklicznego publikowania informacji.

Piąta sesja tematyczna *Synthetic data generation methods*, prowadzona przez Paula Francisa (Max Planck Institute for Software Systems, Niemcy) i Bradleya Malina, dotyczyła metod generowania danych syntetycznych.

W wystąpieniu *The Production of Bespoke Synthetic Teaching Datasets without Access to the Original Data* Mark Elliott i Claire Little (University of Manchester, Wielka Brytania) oraz Richard Allmendinger (Alliance Manchester Business School, Wielka Brytania) omówili proces polegający na tworzeniu na zamówienie uczących zbiorów danych syntetycznych bez dostępu do danych oryginalnych. Jako podstawę syntezy wykorzystuje się w tym przypadku już zatwierdzone i opublikowane analizy. Prelegenci zaprezentowali m.in. stosowane w tym celu algorytmy ewolucyjne oraz wyjaśnili, na czym polega ich praktyczna realizacja związana z przeprowadzaniem eksperymentów dotyczących kontroli ujawniania danych.

Anna Oganian (National Centre for Health Service, Stany Zjednoczone), Jörg Drechsler (Institute for Employment Research, Niemcy, Ludwig-Maximilians--Universität, Niemcy, i University of Maryland, Stany Zjednoczone) i Mechtab Iqbal (Clemson University, Stany Zjednoczone) w prezentacji *Evaluating the Pseudo Likelihood Approach for Synthesizing Surveys under Informative Sampling* przedstawili ocenę użyteczności podejścia pseudowiarygodności do syntetyzacji wyników badań statystycznych opartych na próbkowaniu informacyjnym, czyli takim, w którym prawdopodobieństwo wyboru do próby zależy od wartości zmiennej docelowej szacowanej przez model. Ukazali różne podejścia do tego zagadnienia, także oparte na metodach bayesowskich, oraz przedstawili własną oryginalną propozycję bazującą na pseudopopulacji i estymacji największej wiarygodności, a także podali przykład jej realizacji w formie eksperymentu, pokazujący użyteczność zaproponowanego narzędzia.

Przedmiotem wystąpienia Hajime Ono (National Institute of Information and Communications Technology, Japonia) i Nobuaki Hoshino (Kanazawa University, Japonia) *Hidden Power of Quasi-multinomial Sampling: Utility Analysis and Bias Correction* była ukryta moc próbkowania quasi-wielomianowego w kontekście użyteczności i korekcji obciążenia. Prelegenci zauważyli, że w praktyce trudno jest zapewnić pożądaną użyteczność danych. Sugerowane przez nich rozwiązanie opiera się na losowaniu próby z danych oryginalnych i sztucznych wedle rozkładu quasi-wielomianowego oraz odpowiedniej korekcji obciążenia. Jako miarę użyteczności autorzy stosują miernik podobieństwa rozkładu brzegowego.

Michel Béra (CNAM-Laboratoire Cédric/MSDMA, Francja), Vasiliki Daskalaki, Spiros Kolovos, Antonia Spinakis i Photis Stavropoulos (Quantos S.A. Statistics & Information Systems, Grecja) oraz Konstantinos Spinakis (Ecole Polytechnique Fédérale de Lausanne, Szwajcaria) w wystąpieniu *Evaluation of Synthetic Data Quality Using the Quantile at Risk* przedstawili sposób oceny jakości danych syntetycznych z wykorzystaniem kwantyli na poziomie ryzyka i omówili zastosowanie tego rozwiązania. Wykazali, że metryka oparta na wspomnianych kwantylach uzupełnia tradycyjne metryki użyteczności, zapewniając kompleksową ocenę jakości danych syntetycznych.

Fabio Ricciato (Eurostat, Luksemburg) w prezentacji *A Cautionary Reflection on (Pseudo)Synthetic Data from Deep Learning on Personal Data* podzielił się refleksjami dotyczącymi danych pseudosyntetycznych uzyskiwanych w wyniku głębokiego uczenia przeprowadzanego na danych dotyczących osób. Autor porównał cechy zbiorów syntetycznych i pseudosyntetycznych oraz modeli z nadmierną parametryzacją (klasycznego i głębokiego uczenia), a także ich dopasowanie do danych i relację zachodzącą między użytecznością a ochroną prywatności. Przypomniał, że odmienność rozkładów nie oznacza jeszcze ochrony prywatności. Sieć generatywna trenowana na dużą skalę na danych rzeczywistych może bowiem nauczyć się pewnego rodzaju mechanizmu oszukującego ochronę prywatności.

Jonathan Latner (Institute for Employment Research, Niemcy), Marcel Neunhoffer (Ludwig-Maximilians-Universität, Niemcy) i Jörg Drechsler podczas prelekcji *Generating Synthetic Data Is Complicated: Know Your Data and Know Your Generator* zastanawiali się nad poziomem złożoności generowania danych syntetycznych oraz kluczowym znaczeniem znajomości danych i generatora w tym procesie. Autorzy omówili trudności związane z poruszanymi zagadnieniami oraz przybliżyli znaczenie wiedzy o mechanizmie generowania i miarach użyteczności stosowanych w tym przypadku.

Jorge Cisneros i Timothy Wojan (Oak Ridge Insititute for Science & Education, Stany Zjednoczone), Audrey Kindlon i Heather Madray (National Center for Science & Engineering Statistics, Stany Zjednoczone), Matthew Williams i Jennifer Ozawa (Research Triangle Institute International, Stany Zjednoczone) oraz Christie Task i Damon Streat (Knexus Research, Stany Zjednoczone) przedstawili referat *Developing Synthetic Microdata through Machine Learning for the Annual Business Survey* dotyczący tworzenia mikrodanych syntetycznych z wykorzystaniem uczenia maszynowego na przykładzie rocznego badania działalności gospodarczej. Do realizacji tego celu prelegenci użyli generatora CenSyn w pakiecie `synthpop` środowiska R i R `tidysynthesis`, stworzonego przez Urban Institute i US Internal Revenue Service[1].

Szóstą sesję tematyczną *Synthetic data generation software*, poświęconą informatycznym narzędziom generowania danych syntetycznych, której przewodniczyła Gillian Raab, otworzyło wystąpienie Paula Francisa (Max Planck Institute for Software Systems, Niemcy) *A Comparison of SynDiffix Multi-table versus Single-table Synthetic Data*. Prelegent porównał podejście multi-table z programów SynDiffix z klasycznym single-table. W tym kontekście omówił wykorzystywane w amerykańskiej statystyce narzędzia do oceny użyteczności wygenerowanych danych (typu SDNIST) i proponowane przez siebie rozwiązanie dotyczące tej oceny. Autor badał

[1] Agencja rządowa zajmująca się ściąganiem podatków, podlegająca Departamentowi Skarbu Stanów Zjednoczonych.

też m.in. efektywność analizy głównych składowych w tym ujęciu. Podkreślił, że omawiane narzędzie jest proste w użytkowaniu, i zwrócił uwagę na wysoką jakość otrzymywanych w ten sposób danych.

Emma Fössing (Georg-August-Universität, Niemcy) i Jörg Drechsler w prezentacji *An Evaluation of Synthetic Data Generators Implemented in the Python Library Synthcity* poruszyli temat oceny jakości generatorów danych syntetycznych zaimplementowanych do biblioteki Synthcity w Pythonie. Prelegenci, wykorzystując dane z zakresu szkolnictwa, dokonali przeglądu 12 dostępnych generatorów danych syntetycznych i miar użyteczności.

Oscar Thees (University of Northwestern Switzerland, Szwajcaria), Jiří Novák (University of Northwestern Switzerland, Swiss Data Anonymization Competence Center – SwissAnon i University of Zurich, Szwajcaria) oraz Matthias Templ (University of Northwestern Switzerland i Technische Universität Wien, Austria) w wystąpieniu *Evaluation of Synthetic Data Generators on Complex Tabular Data* przedstawili ocenę generatorów danych syntetycznych dla kompleksowych danych tabelarycznych. Uwzględnili modelowanie warunkowe i łączne oraz predykcję wskaźników skłonności do analizy efektywności. Posłużyli się danymi z Europejskiego Badania Dochodów i Warunków Życia (EU-SILC) i skorzystali z narzędzi pakietów `simpop` i `synthpop` środowiska R.

W siódmej sesji, prowadzonej przez Jörga Drechslera, zaprezentowano studia przypadku z różnych krajów. Mohamed Aghaddar, Liu Nuo Su, Lucas Barnhoorn i Peter-Paul de Wolf (Statistics Netherlands, Holandia) oraz Manel Slokom (Statistics Netherlands i Centrum Wiskunde & Informatica, Holandia) w wystąpieniu *A Case Study Exploring Data Synthesis Strategies on Tabular vs. Aggregated Data Sources for Official Statistics* omówili przykład oceny strategii syntetyzacji danych ze źródeł tabelarycznych i zagregowanych. W referacie rozważano różne kierunki takich działań: od poziomu makro do poziomu mikro i mikro-mikro oraz trzy odmienne scenariusze (brak wiedzy, wiedza wewnętrzna i wiedza zewnętrzna). Zaprezentowano też odpowiednie modele i omówiono efekty ich praktycznego zastosowania na danych rzeczywistych. Prelegenci stwierdzili, że podejście od makro do mikro daje lepsze rezultaty.

Manel Slokom, Shruti Agrawal, Nynke C. Krol i Peter-Paul de Wolf (Statistics Netherlands, Holandia) w prezentacji *Relational or Single: A Comparative Analysis of Data Synthesis Approaches for Privacy and Utility on a Use Case from Statistical Office* przedstawili wyniki analizy porównawczej różnych podejść do syntezy danych pod kątem prywatności i użyteczności w przypadku danych pochodzących z urzędu statystycznego. Do oceny użyteczności wykorzystano modele regresji liniowej trenowane na danych syntetycznych i rzeczywistych, a testowane na danych

rzeczywistych. Do konstrukcji danych syntetycznych użyto modelowania hierarchicznego (m.in. drzewa analityczno-regresyjnego CART).

Krish Muralidhar i Steven Ruggles (University of Minnesota i Integrated Public Use Microdata Series, Stany Zjednoczone) w wystąpieniu *Escalation of Commitment: A Case Study of the United States Census Bureau Efforts to Implement Differential Privacy for the 2020 Decennial Census* zastanawiali się nad kwestią eskalacji zaangażowania, analizując studium przypadku wysiłków Biura Spisów Stanów Zjednoczonych na rzecz wdrożenia różnicowania prywatności na potrzeby spisu ludności w 2020 r. Prelegenci zaprezentowali kalendarium działań Biura Spisów dotyczących rekonstrukcji ochrony prywatności. Dokonali też oceny ryzyka ujawnienia wynikającego z funkcjonowania tablic upublicznionych w 2010 r., której wówczas zabrakło. Wykazali, że procedury stosowane do oceny ryzyka są nierzetelne i przeszacowują je. Zauważyli, że obecnie nie ma jednak innej możliwości niż intensyfikacja nowej procedury w tym zakresie, chociaż nie zapewnia to ani należytej ochrony prywatności, ani dokładności.

Ostatni referat w tej sesji, *Applications of Statistical Disclosure Control Methods to Protect the Confidentiality in Agricultural Census Microdata,* wygłosili Andrzej Młodak (Urząd Statystyczny w Poznaniu i Uniwersytet Kaliski im. Prezydenta Stanisława Wojciechowskiego) i Tomasz Józefowski (Uniwersytet Ekonomiczny w Poznaniu i Urząd Statystyczny w Poznaniu). Wystąpienie dotyczyło kompleksowego algorytmu zaburzania mikrodanych w Portalu Geostatystycznym GUS z wykorzystaniem mikroagregacji opartej na odległości Gowera dla zmiennych kategorialnych i nakładania szumu skorelowanego dla zmiennych ciągłych. Do oceny jakości i efektywności zaburzania wykorzystano znane i nowe miary dostępne w pakiecie `sdcMicro` środowiska R.

Ósmej sesji *Spatial and georeferenced data*, poświęconej danym przestrzennym i georeferencyjnym, przewodniczył Krish Muralidhar.

Pierwsze wystąpienie, *Masking Georeferenced Health Data – An Analysis Taking the Example of Partially Synthetic Data on Sleep Disorder,* przygotowali Simon Cremer i Lydia Jehmlich (Cologne University of Technology, Arts and Sciences, Niemcy) oraz Rainer Lenz (Cologne University of Technology, Arts and Sciences i Technical University of Dortmund, Niemcy). Dotyczyło ono maskowania georeferencyjnych danych o zdrowiu, a dokładniej – analizy bazującej na przykładzie częściowo syntetycznych danych z zakresu zaburzeń snu. Rozpatrywane dane dotyczyły ok. 10% ludności Kolonii (ponad 109 000 osób). Dane ukrywano tzw. metodą pączka i za pomocą przekodowania. Pomiar ryzyka ujawnienia bazował na przestrzennej regule *k*-anonimowości i ilorazie najbliższych sąsiedztw.

Drugi referat w tej sesji, *Privacy and Disclosure Risks in Spatial Dynamic Microsimulations*, przygotowany przez Hannę Brenzel i Martina Palma (Federal Statistical Office of Germany, Niemcy) oraz Jana Weymeirscha i Ralfa Münnicha (Trier University, Niemcy), został poświęcony ocenie ryzyka zagrożenia prywatności w przestrzennych mikrosymulacjach dynamicznych dokonanej w ramach projektu Mikrosim. Jego celem było uzyskanie możliwie najbardziej szczegółowych i rzetelnych, a przy tym bezpiecznych danych przeznaczonych dla użytkowników ze środowiska akademickiego. Szczególną uwagę zwrócono na powiązanie tego programu z realizacją potrzeb statystyki publicznej. Rzadkie rekordy i powiązania w małych podpopulacjach mogą zostać w tym przypadku efektywnie zreplikowane, a zastosowanie odpowiedniego wariantu metody Monte Carlo dla różnych symulacji umożliwia zachowanie oryginalnych cech rozkładów. Ochrona danych bazowała głównie na zaokrąglaniu stochastycznym, a użyteczność tej metody uwidoczniono m.in. na przykładzie analizy zgonów w Trewirze.

Dziewiątą sesję *Machine learning and privacy*, poświęconą uczeniu maszynowemu i prywatności, poprowadził Jordi Castro. Otworzyła ją prelekcja *Combinations of AI Models and XAI Metrics Vulnerable to Record Reconstruction Risk*, którą wygłosili Ryotaro Tomy i Hiroaki Kikuchi (Meiji University, Japonia), przedstawiająca kombinację modeli sztucznej inteligencji (AI) i metryk XAI wrażliwych na ryzyko rekonstrukcji rekordów. AI obejmuje m.in. sieci neuronowe i lasy losowe, a Explanable AI (XAI) to jej wariant zapewniający przejrzystość modeli i wyjaśnienie relacji między cechami w zbiorach wejściowym i wyjściowym. Wkład każdej cechy szacowano za pomocą wartości Shapleya. Wykorzystano model LIME (interpretowalne lokalnie objaśnienie modelowo-agnostyczne). Okazało się, że ryzyko dla wartości Shapleya jest wyższe niż dla LIME.

Saloni Kwatra i Vincenç Torra (Umeå University, Szwecja) w wystąpieniu *DISCOLEAF: Personalized DIScretization of COntinuous Attributes for LEArning with Federated Decision Trees* zaprezentowali metodę spersonalizowanej dyskretyzacji ciągłych atrybutów dla federacyjnego uczenia maszynowego z drzewami decyzyjnymi DISCOLEAF. Prelegenci omówili szczegóły tego podejścia i efekty jego zastosowania, zwłaszcza w kontekście oceny wpływu parametryzacji na jakość danych wynikowych.

Prezentacja autorstwa Oualida Zariego, Ayşe Ünsal i Melek Önen (EURECOM) oraz Javiera Parra-Arnau (Karlsruhe Institute of Technology, Niemcy, i Universitat Politècnica de Catalunya, Hiszpania) *Node Injection Link Stealing Attack* dotyczyła ryzyka kradzieży poprzez atakowanie łączy w grafowych sieciach neuronowych (ang. *graph neural networks* – GNN). Autorzy przedstawili możliwy schemat i przebieg takiego ataku, na który – w kontekście możliwego wycieku informacji wrażliwych – bardzo podatne są generatywne sieci neuronowe. Możliwym sposobem obrony jest w tym przypadku różnicowanie prywatności.

Marko Miletic i Murat Sariyar (Bern University of Applied Sciences, Szwajcaria) poświęcili referat *Assessing the Potentials of LLMs and GANs as State-of-the-art Tabular Synthetic Data Generation Methods* zagadnieniu oceny potencjału dużego modelu językowego (LLM) i generatywnych sieci kontradyktoryjnych (GAN) jako narzędzi generowania syntetycznych danych tabelarycznych. Ocenę użyteczności oparli na odległości Wassersteina, zaproponowali również nową miarę ryzyka ujawnienia. Doszli do wniosku, że LLM ma przewagę nad CTGAN (czyli GAN dla danych tabelarycznych) pod względem użyteczności wygenerowanych danych. Wysunęli też propozycję nowej miary ryzyka o nazwie *identyfikowalność dzielonych sąsiedztw* (ang. *shared neighbor identifiability*).

Osamu Saisho, Takayuki Miura, Kazuki Iwahana, Masanobu Kii i Rina Okada (NTT Communications, Japonia) w wystąpieniu *Active Learning for Human Annotation of Privacy-Preserved Synthetic Data* pochylili się nad aktywnym uczeniem w zakresie adnotacji ludzkich (dane weryfikowane lub klasyfikowane według osób) w ramach danych syntetycznych zachowujących prywatność. Prelegenci zastanawiali się nad znaczeniem tego rodzaju informacji w danych syntetycznych i ukazali schemat aktywnego uczenia dla danych syntetycznych bazujący na danych otwartych. Zaproponowali metodę, która wykorzystuje wyniki grupowania wraz z niewielką ilością danych publicznie dostępnych w celu dalszej poprawy wydajności adnotacji.

Sesję zamknął referat Marianne Abi-Kanaan (Université de Franche-Comté, Francja, i Université Antonine, Liban), Jeana-François Couchota (Université de Franche-Comté, Francja) oraz Talara Atechiana i Rony Darazi (Université Antonine, Liban) *Improving Utility in a DP-Fied ML Algorithm with a Metaheuristics-Based Privacy Budget Allocation Strategy*, poświęcony zwiększaniu użyteczności w algorytmie największej wiarygodności wzmocnionym różnicowaniem prywatności z metaheurystyczną strategią alokacji budżetu prywatności. Kluczowy okazuje się w tym przypadku model drzewa decyzyjnego z gradientem różnicowania prywatności (GBDT). Jako propozycję rozwiązania problemu poprawy użyteczności finalnych danych prelegenci zaprezentowali procedurę GA-DiffBoost. Swój pomysł zilustrowali wynikami eksperymentu, w którym miara straty była wyznaczana bez optymalizacji współczynnika zmienności.

Wystąpienia prelegentów dostarczyły uczestnikom konferencji gruntownej wiedzy o nowoczesnych rozwiązaniach z zakresu kontroli ujawniania danych, a zwłaszcza generowania danych syntetycznych z wykorzystaniem narzędzi sztucznej inteligencji i narzędzi służących do oceny jakości i użyteczności otrzymywanych w ten sposób informacji przeznaczonych do udostępnienia. Z uwagi na potrzeby użytkowników, a zwłaszcza oczekiwania większego zakresu i bardziej szczegółowych danych, oraz coraz większą wiedzę na temat możliwych naruszeń prywatności wdrożenie

tego rodzaju narzędzi w polskiej statystyce publicznej i u innych gestorów danych stanie się wkrótce koniecznością. Więcej informacji na temat konferencji można znaleźć pod adresem https://crises-deim.urv.cat/psd2024/.

Andrzej Młodak
Uniwersytet Kaliski im. Prezydenta Stanisława Wojciechowskiego, Międzywydziałowy Zakład Matematyki i Statystyki; Urząd Statystyczny w Poznaniu, Ośrodek Statystyki Małych Obszarów, Polska / University of Kalisz, Inter-faculty Department of Mathematics and Statistics; Statistical Office in Poznań, Centre of Small Area Estimation, Poland
ORCID: https://orcid.org/0000-0002-6853-9163

Tomasz Józefowski
Uniwersytet Ekonomiczny w Poznaniu, Instytut Informatyki i Ekonomii Ilościowej; Urząd Statystyczny w Poznaniu, Ośrodek Statystyki Małych Obszarów, Polska / Poznań University of Economics and Business, Institute of Informatics and Quantitative Economics; Statistical Office in Poznań, Centre of Small Area Estimation, Poland
ORCID: https://orcid.org/0000-0001-9485-1946

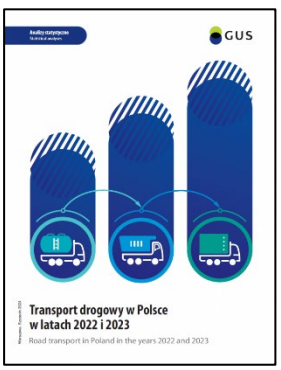
Wiadomości Statystyczne. The Polish Statistician, 2024, *69*(11), 81–82
Dyskusje. Recenzje. Informacje / Discussions. Reviews. Information

WYDAWNICTWA GUS. PAŹDZIERNIK 2024
PUBLICATIONS OF STATISTICS POLAND. OCTOBER 2024

W ofercie wydawniczej Głównego Urzędu Statystycznego z ubiegłego miesiąca warto zwrócić uwagę na następującą publikację:
Among Statistics Poland's publications from the previous month, we would like to recommend:

Transport drogowy w Polsce w latach 2022 i 2023
Road transport in Poland in the years 2022 and 2023

Siódma edycja publikacji zawierającej podstawowe dane na temat transportu drogowego w Polsce, wydawanej co dwa lata przez Główny Urząd Statystyczny i Urząd Statystyczny w Szczecinie.



Język: polski, angielski
Language: Polish, English
Seria: Analizy statystyczne
Series: Statistical analyses
Dostępne wersje: drukowana i elektroniczna
Available in: printed and electronic form

W opracowaniu przedstawiono wyniki badań statystycznych i dane uzyskane przez statystykę publiczną ze źródeł administracyjnych dotyczące m.in. wielkości przewozów ładunków i pasażerów, infrastruktury drogowej, pojazdów samochodowych i oddziaływania ruchu drogowego na środowisko. Zobrazowano sytuację przedsiębiorstw transportu drogowego, m.in. podano informacje o sieci drogowej i środkach transportu drogowego, transporcie towarowym i pasażerskim, ruchu drogowym i wypadkach drogowych. Zamieszczono także dane Eurostatu o transporcie drogowym w Unii Europejskiej.

W październiku br. ukazały się ponadto:
- *Aktywność ekonomiczna ludności Polski – 2 kwartał 2024 r.*;
- „Biuletyn statystyczny" nr 9/2024;
- *Ceny robót budowlano-montażowych i obiektów budowlanych (sierpień 2024 r.)*;

- *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach 2000–2024 (październik 2024)*;
- *Ludność. Stan i struktura ludności oraz ruch naturalny w przekroju terytorialnym w 2024 r. Stan w dniu 30 czerwca*;
- *Oświata i wychowanie w roku szkolnym 2023/2024*;
- *Pracujący w gospodarce narodowej w 2023 r.*;
- *Produkcja budowlano-montażowa w 2023 r.*;
- *Produkcja ważniejszych wyrobów przemysłowych we wrześniu 2024 r.*;
- *Rocznik Statystyczny Handlu Zagranicznego 2024*;
- *Ruch graniczny oraz wydatki cudzoziemców w Polsce i Polaków za granicą w 2023 r.*;
- *Rynek wewnętrzny w 2023 r.*;
- *Sytuacja społeczno-gospodarcza kraju – 1–3 kwartał 2024 r.*;
- *Szkolnictwo wyższe i jego finanse w 2023 r.*;
- „Wiadomości Statystyczne. The Polish Statistician" nr 10/2024;
- *Zeszyt metodologiczny. Badania produkcji przemysłowej*;
- *Zeszyt metodologiczny. Statystyka krótkookresowa według europejskiej koncepcji jednostki rodzaju działalności*;
- *Żegluga śródlądowa w Polsce w latach 2022–2023.*

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

Wszystkie publikacje GUS w wersji elektronicznej są dostępne na stronie stat.gov.pl/publikacje/publikacje-a-z. Wersje drukowane (jeśli zostały wydane) można zamawiać pod adresem: zws-sprzedaz@stat.gov.pl.
All the publications of Statistics Poland available in electronic form can be accessed at stat.gov.pl/en/publications. Printed versions (if available) may be ordered at: zws-sprzedaz@stat.gov.pl.

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 czasopism naukowych Ministerstwa Nauki i Szkolnictwa Wyższego. Zgodnie z komunikatem Ministra Nauki z dnia 5 stycznia 2024 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych „WS" otrzymały 70 punktów.

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

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

1. Zgłaszanie artykułów

Prace przeznaczone do opublikowania w „WS" należy przesyłać za pośrednictwem platformy Editorial System: www.editorialsystem.com/ws.

Zgłaszany artykuł powinien być zanonimizowany, tj. pozbawiony informacji o autorze/autorach (również we właściwościach pliku), podziękowań i informacji o źródłach finansowania, a także innych informacji wskazujących na afiliację lub umożliwiających zidentyfikowanie autora. Jeżeli w pracy występują tablice, wykresy lub mapy, powinny być umieszczone w treści artykułu. Materiały graficzne, razem z danymi do nich, należy ponadto załączyć jako osobny plik / osobne pliki, najlepiej w formacie Excel. **Prosimy o niestosowanie stylów i ograniczenie formatowania do wymogów redakcyjnych.** Więcej informacji w pkt 4 *Wymogi redakcyjne*.

Razem z artykułem należy przesłać skan/zdjęcie oświadczenia o oryginalności pracy i niezłożeniu jej w innym wydawnictwie. **Załączenie oświadczenia jest warunkiem poddania pracy ocenie wstępnej i skierowania do recenzji.**

Zgłoszenie artykułu do opublikowania w „WS" oznacza zgodę na jego udostępnienie na licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0).

Autorzy mają prawo do samodzielnego umieszczania w wybranych przez siebie repozytoriach artykułu w wersji zarówno zgłoszonej do „WS", jak i zaakceptowanej do opublikowania oraz opublikowanej, z zastrzeżeniem wymogu niezwłocznego podania w repozytorium informacji o numerze „WS", w którym praca się ukazała, wraz z linkiem do niej (DOI).

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, zgodności jego tematyki z profilem czasopisma, struktury i zawartości pracy pod kątem wymogów redakcyjnych oraz oryginalności (wykrywanie programem antyplagiatowym treści zapożyczonych, a także wygenerowanych za pomocą narzędzi sztucznej inteligencji). Na jej podstawie formułowane są uwagi i zalecenia dla autora. Poprawiona/uzupełniona przez autora praca jest kierowana do recenzji. W przypadku negatywnej weryfikacji artykuł zostaje odrzucony, a autor otrzymuje decyzję wraz z uzasadnieniem.
2. **Ocena recenzentów**, dokonywana przez specjalistów w danej dziedzinie. Artykuł oceniają dwaj recenzenci spoza jednostki naukowej, przy której afiliowany jest autor, i spoza Zespołu Redakcyjnego „WS"; w przypadku pracy w języku angielskim co najmniej jeden recenzent jest afiliowany przy jednostce zagranicznej. W razie sprzecznych opinii dwóch recenzentów powoływany jest trzeci recenzent. Recenzenci kierują się kryteriami oryginalności i jakości opracowania zarówno w odniesieniu do treści, jak i formy artykułu.

   Autor pracy, która otrzymała dwie pozytywne oceny, wprowadza poprawki zalecane przez recenzentów i przesyła do redakcji skorygowaną wersję tekstu. Jeśli pojawi się różnica zdań dotycząca zasadności proponowanych zmian, autor jest zobligowany do uzasadnienia swojego stanowiska.
3. **Ocena redakcji**, decydująca o przyjęciu pracy do publikacji. Polega m.in. na weryfikacji dokonania przez autora zmian w artykule stosownie do uwag recenzentów. Redakcja 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 do publikacji są zamieszczane na stronie internetowej czasopisma w zakładce Early View, gdzie znajdują się do czasu opublikowania w konkretnym wydaniu „WS".
4. **Opracowanie redakcyjne, autoryzacja i korekta.** Artykuł zakwalifikowany do druku jest poddawany opracowaniu redakcyjnemu, a następnie – po autoryzacji – przekazywany do składu, łamania i opracowania graficznego. Następnie wykonywane są co najmniej dwie korekty wydawnicze. Autor wykonuje korektę autorską na etapie drugiej korekty wydawniczej.

   Redakcja zastrzega sobie prawo do zmiany tytułu i śródtytułów, modyfikowania tablic, wykresów i innych elementów graficznych oraz przeredagowywania treści bez naruszenia zasadniczej myśli autora.

   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

Wszyscy uczestnicy procesu publikacyjnego są zobowiązani do przestrzegania zasad etyki publikacyjnej. Zasady przyjęte w „Wiadomościach Statystycznych. The Polish Statistician" („WS") opierają się na wytycznych Komitetu ds. Etyki Publikacyjnej (Committee on Publication Ethics – COPE), które są dostępne na stronie internetowej www.publicationethics.org.

W celu zapewnienia transparentności w publikowaniu wyników badań naukowych wymagane jest, aby każdy uczestnik procesu publikacyjnego zgłaszał potencjalne konflikty interesów. Przez konflikt interesów rozumiane jest

> wszystko, co zakłóca lub może być w sposób uzasadniony postrzegane jako zakłócające pełne i obiektywne prezentowanie i recenzowanie artykułów przesłanych do czasopisma, podejmowanie decyzji redakcyjnych w ich sprawie lub ich publikowanie. Konflikty interesów mogą mieć charakter finansowy lub niefinansowy, zawodowy lub osobisty i mogą powstać w stosunkach z instytucją lub inną osobą [na podstawie: https://journals.plos.org/plosone/s/competing-interests].

Redakcja nie toleruje przejawów nierzetelności naukowej, takich jak:

- plagiat – przywłaszczenie cudzego utworu lub jego fragmentu bez podania informacji o źródle;
- autoplagiat – ponowne publikowanie własnego utworu lub jego części;
- fabrykowanie danych – oparcie pracy naukowej na nieprawdziwych wynikach badań;
- autorstwo widmowe (*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 tworzeniu artykułu, aby lista autorów wyglądała bardziej imponująco;
- autorstwo grzecznościowe (*gift authorship*) – dodawanie jako współautorów osób, których wkład jest oparty jedynie na słabym powiązaniu z badaniem, w ramach przysługi, uznania lub uprzejmości.

Odpowiedzialność poszczególnych uczestników procesu publikacyjnego w zakresie etyki publikacyjnej jest przedstawiona poniżej.

3.1. Odpowiedzialność autorów

3.1.1. Oryginalność pracy

Artykuły naukowe zgłaszane do publikacji w „WS" muszą stanowić własność intelektualną autorów i być pracami oryginalnymi, nie mogą naruszać praw autorskich innych osób, być wcześniej publikowane ani złożone w innym wydawnictwie (także w innej wersji językowej), a w przypadku wykorzystania narzędzi sztucznej inteligencji autorzy muszą mieć większościowy wkład twórczy w powstanie artykułu, co deklarują w oświadczeniu. W przypadku złożenia artykułu w innym wydawnictwie przed ukazaniem się go w „WS" autorzy są zobowiązani do niezwłocznego powiadomienia o tym redakcji „WS".

Jeżeli materiały, na podstawie których powstał artykuł, były prezentowane publicznie, np. podczas konferencji, to autorzy powinni poinformować o tym redakcję, zgłaszając tekst do publikacji w „WS".

Jeśli autorzy zgłoszonego artykułu umieścili go w repozytorium przed opublikowaniem w „WS", to niezwłocznie po ukazaniu się numeru „WS" z tym artykułem powinni podać przy artykule zamieszczonym w repozytorium link do publikacji w „WS".

3.1.2. Autorstwo

Autorzy ponoszą odpowiedzialność za treści prezentowane w artykułach.

W artykule muszą być wskazane wszystkie osoby, które wniosły znaczący wkład w jego powstanie. Niedopuszczalne jest autorstwo widmowe, gościnne i grzecznościowe.

Autor zgłaszający artykuł określa procentowy udział autorów i ich wkład odpowiednio dla:

- koncepcji i projektu badania;
- gromadzenia lub zestawiania danych;
- analizy i interpretacji danych;
- napisania artykułu;
- krytycznego zrecenzowania artykułu;
- zatwierdzenia ostatecznej wersji artykułu.

Wszelkie zmiany na liście autorów (dodawanie lub usuwanie nazwisk i zmiana kolejności autorów) po zgłoszeniu artykułu do publikacji w „WS" wymagają przesłania do redakcji formularza zmiany na liście autorów podpisanego przez wszystkich autorów. Redakcja nie rozstrzyga ewentualnych sporów między autorami, a w przypadku braku możliwości uzgodnienia między nimi wspólnego stanowiska wycofuje artykuł z publikacji.

W przypadku śmierci jednego z autorów przed opublikowaniem artykułu współautorzy poręczają za niego w zakresie jego wkładu i potencjalnych konfliktów interesów.

Wkład innych osób w powstanie artykułu, który nie spełnia kryteriów autorstwa, taki jak wspieranie badania, ogólny mentoring, pełnienie funkcji koordynatora badania i inne powiązane działania, można wskazać w części artykułu pt. „Podziękowania".

Każdy autor powinien posługiwać się identyfikatorem Open Researcher Contributor ID.

3.1.3. Rzetelność badań

Artykuły naukowe powinny zawierać precyzyjny opis badanych zjawisk i stosowanych metod oraz autorskie wnioski.

3.1.4. Cytowanie

Wszystkie zawarte w artykule informacje, dane i stwierdzenia niebędące autorskimi i wykraczające poza wiedzę powszechną muszą być opatrzone przypisem bibliograficznym, niezależnie od tego, czy są ujęte w ramy cytatu, czy nie są dosłownie przytaczane.

Autorzy artykułu ponoszą odpowiedzialność za właściwe oznaczanie cytowanych prac innych autorów.

3.1.5. Dane i odtwarzalność badań

Autorzy powinni dokładnie opisać dane użyte w badaniu empirycznym, aby umożliwić powtórzenie badania. Są także zobowiązani do udostępnienia surowych danych badawczych na

prośbę redakcji. Jeżeli spełnienie tej prośby nie jest możliwe z istotnych powodów, powinni uzasadnić swoją odmowę.

3.1.6. Użycie narzędzi sztucznej inteligencji

Podczas zbierania i analizy danych, pisania artykułu i opracowywania elementów graficznych autorzy mogą wspomagać się narzędziami sztucznej inteligencji, ale to oni powinni mieć większościowy wkład twórczy w powstanie artykułu i są w pełni odpowiedzialni za treści wygenerowane automatycznie, a tym samym za wszelkie związane z tym naruszenia etyki publikacyjnej. Są także zobowiązani do poinformowania redakcji o użyciu narzędzi sztucznej inteligencji. Takie narzędzia nie mogą być wskazane jako współautorzy.

Artykuł, w przypadku którego autorzy nie mają większościowego wkładu twórczego i który w przeważającej części powstał przy użyciu narzędzi sztucznej inteligencji, nie może być uznany za oryginalną pracę naukową i przyjęty do publikacji.

Niniejsze wytyczne nie obejmują narzędzi, które są używane do poprawy pisowni, gramatyki i ogólnej edycji.

Ostateczną decyzję o tym, czy użycie narzędzi sztucznej inteligencji jest właściwe lub dopuszczalne w przypadku danego artykułu, podejmuje redaktor naczelny.

3.1.7. Konflikt interesów

Autorzy są zobowiązani do zgłoszenia redakcji wszystkich potencjalnych konfliktów interesów odnoszących się do badania przedstawionego w artykule.

Autorzy podają w artykule źródła finansowania badania.

Niezgłoszenie istniejącego konfliktu interesów może skutkować odrzuceniem artykułu.

Ujawnienie konfliktu interesów autorów, który miał nadmierny wpływ na artykuł lub jego recenzje, po publikacji będzie skutkowało retrakcją artykułu.

3.1.8. Współpraca

Autorzy biorą udział w procesie recenzowania *double-blind peer review*, dokonywanej przez co najmniej dwóch niezależnych ekspertów z danej dziedziny. Po otrzymaniu minimum dwóch pozytywnych recenzji autorzy wprowadzają zalecane przez recenzentów poprawki i przesyłają do redakcji zaktualizowaną wersję artykułu wraz z poświadczeniem uwzględnienia poprawek.

W przypadku różnicy zdań co do zasadności proponowanych zmian i nieuwzględnienia którejś z zalecanych poprawek autorzy uzasadniają swoje stanowisko.

Autorzy zatwierdzają artykuł po opracowaniu redakcyjnym (autoryzują go) i biorą udział w korekcie autorskiej.

W razie zgłaszania przez czytelników zastrzeżeń do opublikowanych artykułów ich autorzy są zobligowani do udzielenia odpowiedzi za pośrednictwem redakcji.

3.2. Odpowiedzialność redakcji

3.2.1. Obiektywizm i uczciwość

Redakcja podejmuje decyzję o publikacji danego artykułu, kierując się kryteriami merytorycznej oceny wartości artykułu, jego oryginalności, rzetelności i jasności przekazu, a także ścisłego

związku z celem i zakresem tematycznym „WS". Ocenia artykuły niezależnie od płci, rasy, pochodzenia etnicznego, narodowości, religii, wyznania, światopoglądu, niepełnosprawności, wieku lub orientacji seksualnej ich autorów.

3.2.2. Przeciwdziałanie nierzetelności naukowej

Redakcja nie toleruje przejawów nierzetelności naukowej, takich jak: plagiat, autoplagiat, fabrykowanie danych oraz autorstwo widmowe, gościnne i grzecznościowe.

Jeżeli na którymkolwiek etapie procesu publikacyjnego powstaje uzasadnione podejrzenie, że autorzy dopuścili się nierzetelności naukowej, redakcja skrupulatnie bada sprawę zgodnie z zasadami COPE określonymi na stronie https://publicationethics.org/guidance/Flowcharts. W przypadku udowodnienia nierzetelności autorów zgłoszony przez nich artykuł zostaje odrzucony (w przypadku opublikowanego artykułu – wycofany), a autorzy otrzymują informację o podjętej decyzji wraz z uzasadnieniem. Redakcja informuje o nierzetelności autorów odpowiednie podmioty (instytucje zatrudniające autorów, towarzystwa naukowe itp.).

W celu uzyskania obiektywnej oceny oryginalności nadsyłanych artykułów przed skierowaniem ich do recenzji redakcja wykorzystuje system antyplagiatowy. W przypadku wykrycia znacznego podobieństwa artykułu do innych prac lub wysokiego prawdopodobieństwa użycia narzędzi sztucznej inteligencji redaktor naczelny, po zasięgnięciu opinii pozostałych członków redakcji i Rady Konsultacyjnej, podejmuje decyzję o przyjęciu lub odrzuceniu artykułu. W przypadku odrzucenia autor otrzymuje decyzję wraz z uzasadnieniem.

3.2.3. Konflikt interesów

Redaktorzy są zobowiązani do zgłoszenia wszelkich potencjalnych konfliktów interesów odnoszących się do autorów, badań przedstawianych w artykułach i instytucji je finansujących. Nie mogą być zaangażowani w decyzje redakcyjne dotyczące artykułów ich autorstwa zgłoszonych do publikacji w „WS". W przypadku gdy ich własne interesy mogą utrudniać im bezstronną ocenę danego artykułu i dotyczącą go decyzję o publikacji, powinni wycofać się z jego oceny lub dyskusji na jego temat.

W celu zapobiegania konfliktom interesów między recenzentami a autorami oraz zapewnienia uczciwego i bezstronnego procesu recenzowania redakcja wybiera recenzentów spośród specjalistów spoza jednostki, do której afiliowani są autorzy, i spoza Zespołu Redakcyjnego „WS".

Jeżeli po opublikowaniu artykułu zostanie ujawniony konflikt interesów autorów, to redakcja zbada, czy miał on nadmierny wpływ na artykuł lub jego recenzje. W przypadku gdy taki wpływ zostanie stwierdzony, artykuł podlega retrakcji.

3.2.4. Poufność

Informacje dotyczące artykułu są poufne. Redaktorowi ani żadnemu innemu pracownikowi redakcji nie wolno ich ujawnić nikomu poza autorami, recenzentami, doradcami i – jeśli to uzasadnione – wydawcą.

W przypadku podjęcia decyzji o niepublikowaniu artykułu nie może on zostać w żaden sposób wykorzystany przez wydawcę lub uczestników procesu publikacyjnego bez pisemnej zgody autorów.

3.2.5. Dyskusja na temat opublikowanych artykułów

Każdy może zgłosić redakcji błędy lub naruszenia dostrzeżone w opublikowanych artykułach. Postępowanie redakcji w takich przypadkach zostało określone w punktach 6–8.

Redakcja publikuje również nadesłane polemiki z opublikowanymi artykułami.

3.2.6. Korekta opublikowanego artykułu

W przypadku odkrycia przez autorów lub czytelników istotnych błędów w opublikowanym artykule (np. błędów, które wpływają na interpretację danych lub przedstawionych informacji), redakcja dokonuje korekty artykułu lub zamieszcza sprostowanie. Drobne usterki redakcyjne lub techniczne, które nie wpływają na znaczenie lub interpretację artykułu, zazwyczaj nie są korygowane po publikacji. Nie są też dodawane do artykułu treści wykraczające poza jego pierwotny zakres, takie jak dodatkowe odniesienia lub aktualizacje oparte na informacjach niedostępnych w momencie publikacji artykułu. Po zgłoszeniu przez autorów lub czytelników błędu redakcja ocenia, czy jest on na tyle istotny, że wymaga korekty. W przypadku korekty opublikowanego artykułu jego nowa wersja jest zamieszczana na stronie internetowej „WS" wraz ze stosowną adnotacją.

3.2.7. Wycofanie (retrakcja) opublikowanego artykułu

Jeżeli po opublikowaniu w artykule zostaje wykryty poważny błąd lub naruszenie (np. oszustwo, plagiat, naruszenie praw autorskich, powielona publikacja, nieujawniony konflikt interesów, wykorzystanie informacji poufnych niezgodnie z prawem), które unieważniają przedstawione w artykule ustalenia, to artykuł podlega retrakcji. Redakcja postępuje wtedy w następujący sposób:

- w najbliższym numerze „WS" publikowana jest notatka o wycofaniu artykułu podpisana przez autorów lub redaktora naczelnego, z podaniem daty i powodu wycofania artykułu oraz linkiem do oryginalnego artykułu;
- oryginalny artykuł pozostaje niezmieniony, z wyjątkiem umieszczenia znaku wodnego na każdej stronie pliku PDF o treści „artykuł wycofany".

3.2.8. Zastrzeżenia redakcji dotyczące opublikowanego artykułu

Jeżeli istnieją uzasadnione obawy co do rzetelności badania przedstawionego w opublikowanym artykule lub podejrzenia jakichkolwiek nieprawidłowości (dowody na niepoprawność badania przeprowadzonego przez autorów nie są rozstrzygające, ale charakter wątpliwości uzasadnia powiadomienie czytelników; istnieje uzasadniona obawa, że ustalenia są niewiarygodne lub że mogło dojść do nieprawidłowości), redakcja może opublikować notatkę z zastrzeżeniami, że do wyników przedstawionego w nim badania należy podchodzić z ostrożnością. Takie zastrzeżenia są publikowane jedynie w przypadku, gdy dochodzenie dotyczące artykułu nie przyniosło rezultatów. Redakcja może opublikować swoje zastrzeżenia również wtedy, gdy dochodzenie w sprawie wątpliwego artykułu jest w toku.

3.3. Odpowiedzialność recenzentów

3.3.1. Rzetelność i terminowość

Recenzenci przyjmują artykuł do zrecenzowania, jeśli posiadają odpowiednią wiedzę w określonej dziedzinie, aby rzetelnie ocenić pracę, a także gdy mogą wywiązać się z terminu ustalonego przez redakcję, aby nie opóźniać publikacji.

3.3.2. Obiektywizm

Recenzenci uczestniczą w procesie opartym na modelu *double-blind peer review*, zgodnie z którym nie znają tożsamości autorów ani ich tożsamość nie jest znana autorom.
Recenzenci oceniają artykuł zgodnie z kryteriami zawartymi w karcie recenzji „WS". Powinni uzasadnić swoją ocenę, przedstawiając stosowną argumentację. Są zobligowani do zachowania obiektywności i powstrzymania się od osobistej krytyki.

3.3.3. Wsparcie redakcji

Recenzenci wspierają redakcję w ocenie artykułów zgłoszonych do publikacji. Ich zadaniem jest wyrażenie opinii, czy artykuł:
- może być opublikowany w obecnej formie;
- może być opublikowany po uwzględnieniu zalecanych poprawek;
- wymaga znacznej modyfikacji i ponownej oceny recenzenta (w ponownej ocenie zapada ostateczna decyzja o dopuszczeniu do publikacji lub odrzuceniu);
- nie powinien zostać opublikowany.

3.3.4. Wsparcie autora

Recenzenci powinni wskazać ważne dla wyników badań opublikowane prace, które w ich ocenie powinny zostać przywołane w ocenianym artykule.

3.3.5. Użycie narzędzi sztucznej inteligencji

Niedopuszczalne jest korzystanie z narzędzi sztucznej inteligencji podczas sporządzania recenzji, z wyjątkiem narzędzi, które są używane do poprawy pisowni, gramatyki i ogólnej edycji.

3.3.6. Przeciwdziałanie nierzetelności naukowej

W razie stwierdzenia wysokiego poziomu zbieżności treści recenzowanej pracy z innymi opublikowanymi materiałami lub podejrzenia innych przejawów nierzetelności naukowej recenzenci informują o tym redakcję.

3.3.7. Konflikt interesów

Recenzenci są zobowiązani do zgłoszenia redakcji – zgodnie z ich stanem wiedzy – wszelkich potencjalnych konfliktów interesów odnoszących się do autorów, przedstawionych w artykule badań i instytucji je finansujących. Jeżeli uznają, że istnieje taki konflikt interesów, to powinni odstąpić od recenzowania artykułu.

3.3.8. Poufność

Recenzenci powinni traktować artykuły przesłane im do zrecenzowania jako poufne. Nie mogą ich udostępniać ani omawiać z osobami spoza redakcji, chyba że redakcja wyrazi na to zgodę. Po ukończeniu recenzji przechowywanie przesłanych przez redakcję materiałów (w jakiejkolwiek formie) oraz posługiwanie się nimi przez recenzentów jest niedozwolone.

3.4. Odpowiedzialność wydawcy

3.4.1. Ochrona własności intelektualnej

Materiały opublikowane w „WS" są chronione prawem autorskim. Od 2022 r. autorzy udzielają wydawcy – Głównemu Urzędowi Statystycznemu – licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0), która jest dostępna na stronie https://creativecommons.org/licenses/by-sa/4.0/legalcode.pl. Szczegółowa informacja o prawach autorskich (copyright) jest podawana przy każdym artykule, zarówno w wersji elektronicznej, jak i drukowanej.

3.4.2. Otwarty dostęp

Wydawca udostępnia pełną treść artykułów w internecie w trybie otwartego dostępu, tj. bezpłatnie i bez technicznych ograniczeń. Użytkownicy mogą czytać, pobierać, kopiować, drukować i wykorzystywać do innych celów artykuły zamieszczone na stronie internetowej czasopisma, zgodnie z zapisami:

- ustawy o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego w przypadku artykułów zgłoszonych do 31.12.2021 r.;
- licencji Creative Commons w przypadku artykułów zgłoszonych po 31.12.2021 r.
  Inne sposoby wykorzystania treści artykułów „WS" wymagają zgody wydawcy.

3.4.3. Sprostowania i przeprosiny

Wydawca deklaruje gotowość do opublikowania sprostowań i przeprosin.

3.5. Odwołania i skargi

3.5.1. Odwołania

Autorzy mogą się odwołać od decyzji o niepublikowaniu artykułu. W tym celu powinni skontaktować się z redaktorem naczelnym lub sekretarzem redakcji i przedstawić stosowną argumentację. Odwołania autorów są rozpatrywane przez redaktora naczelnego.

3.5.2. Skargi

Każdy uczestnik procesu publikacyjnego oraz czytelnicy mają prawo do złożenia skargi. Skargę należy przesłać do adres redakcji udostępniony w zakładce Kontakt.

4. Wymogi redakcyjne

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

4.1. Struktura i zawartość artykułu

Wymagane elementy artykułu recenzowanego:

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

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

- *Bibliografia*, zawierająca pełny wykaz prac i materiałów przywołanych w artykule, przygotowana zgodnie z wymogami czasopisma (zob. Przywoływanie źródeł w artykułach napisanych w języku polskim oraz Bibliografia załącznikowa w artykułach napisanych w języku polskim).

  Wszystkie części powinny być opatrzone numerami.

8. Jeżeli podczas gromadzenia i analizy danych, pisania artykułu lub opracowywania elementów graficznych do niego autor korzystał z narzędzi sztucznej inteligencji, to powinien podać w tekście, jakich narzędzi i do czego użył.

   W przypadku artykułu nierecenzowanego nie są wymagane streszczenie, słowa kluczowe ani kody JEL. Bibliografia załącznikowa jest opcjonalna.

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, streszczenie, 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, śródtytuły wyższego rzędu;
   - 12 pkt – tekst główny, śródtytuły niższego rzędu;
   - 10 pkt – pozostałe elementy.
8. Marginesy – 2,5 cm z każdej strony.
9. Interlinia – 1,5 wiersza; tablice i przypisy – 1 wiersz; przed tytułami rozdziałów i podrozdziałów oraz po nich – pusty wiersz.
10. Wcięcie akapitowe – 0,4 cm; bibliografia – bez wcięcia, wysunięcie 0,4 cm.
11. Przy wyliczeniach należy posłużyć się listą punktowaną z punktorami w postaci kropek (wysunięcie 0,4 cm, wcięcie 0 cm); wiersze (oprócz ostatniego) zakończone średnikiem.
12. Strony ponumerowane automatycznie.
13. Tablice i elementy graficzne (wykresy, mapy, schematy) muszą być przywołane w tekście.
14. Wykresy, mapy i schematy należy zamieścić w tekście głównym. Wykresy powinny być edytowalne (optymalnie wykonane w programie Excel; w przypadku wykonania w programie graficznym powinny mieć postać wektorową). Wykresy i inne materiały graficzne należy przekazać osobno, najlepiej w pliku programu Excel lub innym edytowalnym w pakiecie Microsoft Office.

15. Tablice muszą być edytowalne. Nie należy stosować rastrów, cieniowania, pogrubiania czy też podwójnych linii itp.
16. Wskazówki dotyczące opracowywania map znajdują się w publikacji *Mapy statystyczne. Opracowanie i prezentacja danych*, dostępnej na stronie internetowej GUS.
17. Pod tablicami i każdym elementem graficznym należy podać źródło opracowania, a także objaśnić użyte w nich skróty i symbole.
18. Literowe symbole liczb i innych wielkości niezłożonych należy zapisywać małą lub dużą literą i pismem pochyłym (np. $a, A, y(x), a_i$); wektorów – pismem pochyłym i pogrubionym (np. $\boldsymbol{a}, \boldsymbol{A}, \boldsymbol{w}, \boldsymbol{y}(x), \boldsymbol{w}_i$); macierzy – pismem prostym i pogrubionym (np. $\mathbf{A}, \mathbf{a}, \mathbf{M}, \mathbf{Y}(x), \mathbf{M}_i$).
19. Objaśnienia znaków umownych i zapisów w tablicach: kreska (–) – zjawisko nie wystąpiło; zero (0) – zjawisko istniało w wielkości mniejszej od 0,5; (0,0) – zjawisko istniało w wielkości mniejszej od 0,05; kropka (.) – brak informacji, konieczność zachowania tajemnicy statystycznej, wypełnienie pozycji jest niemożliwe lub niecelowe; „w tym" – oznacza, że nie podaje się wszystkich składników sumy.
20. Stosowane są następujące skróty: tablica – tabl., wykres – wykr.
21. Wszystkie zawarte w artykule informacje, dane i stwierdzenia wykraczające poza wiedzę powszechną – np. wyniki badań innych autorów, zarówno o charakterze empirycznym, jak i koncepcyjnym – muszą być opatrzone przypisem bibliograficznym. Przez wiedzę powszechną należy rozumieć informacje ogólnie znane i niebudzące wątpliwości ani kontrowersji w danej grupie społecznej, np. utworzenie GUS w 1918 r. lub powstanie UE w 1993 r. na podstawie traktatu z Maastricht. Natomiast dane statystyczne udostępniane lub publikowane np. przez GUS lub Eurostat nie należą do takich informacji. Charakteru wiedzy powszechnej nie mają również stwierdzenia odnoszące się do idei, zjawisk i procesów społecznych, politycznych czy gospodarczych. Nawet pozornie zdroworozsądkowe idee zmieniają bowiem swój sens w zależności od kultury, języka lub dyscypliny naukowej, a także bywają w rozmaity sposób konceptualizowane, jak np. pojęcie poznania w naukach społecznych.

    Podanie źródła jest konieczne niezależnie od tego, czy informacje lub stwierdzenia są ujęte w ramy cytatu, czy przedstawione bez dosłownego przytoczenia, np. w formie parafrazy. Jeżeli stwierdzenie może budzić jakiekolwiek wątpliwości odbiorców, autor powinien wskazać stosowne źródło podawanej informacji.
22. Przypisy rzeczowe, słownikowe lub informacyjne należy umieszczać na dole strony. Przypisy bibliograficzne, zgodnie ze standardem APA (American Psychological Association), należy podawać w tekście głównym.
23. Bibliografię należy przygotować zgodnie ze standardem APA.

4.3. Przywoływanie źródeł w artykułach napisanych w języku polskim

4.3.1. Ogólne zasady APA

Wyszczególnienie	Przykład przywołania	
	w odsyłaczu	w treści zdania
Autor indywidualny		
Jeden autor	(Iksiński, 2001)	Iksiński (2001)
Dwóch autorów	(Iksiński i Nowak, 1999)	Iksiński i Nowak (1999)
Trzech autorów lub więcej	(Jankiewicz i in., 2003)	Jankiewicz i in. (2003)
Autor instytucjonalny		
Nazwa funkcjonuje jako powszechnie znany skrótowiec: pierwsze przywołanie w tekście	(International Labour Organization [ILO], 2020)	International Labour Organization (ILO, 2020)
kolejne przywołanie	(ILO, 2020)	ILO (2020)
Pełna nazwa	(Stanford University, 1995)	Stanford University (1995)
Niepełne dane bibliograficzne		
Brak ustalonego autorstwa	(*Skrócony tytuł...*, 2015)	*Pełny tytuł* (2015)
Brak roku wydania	(Iksiński, b.r.)	Iksiński (b.r.)
Inne przypadki		
Przywoływanie kilku prac (porządek prac – chronologiczny, porządek autorów – alfabetyczny)	(Iksiński, 1997, 1999, 2004a, 2004b; Nowak, 2002)	Iksiński (1997, 1999, 2004a, 2004b) i Nowak (2002)
Przywoływanie publikacji za innym autorem (uwaga: w bibliografii należy wymienić tylko pracę czytaną)	(Nowakowski, 1990, za: Zieniecka, 2007)	Nowakowski (1990, za: Zieniecka, 2007)
Praca tłumaczona, przedruk lub wydanie wznowione	(Adamski, 1857/2020)	Adamski (1857/2020)

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

4.3.2. Szczegółowe wewnętrzne zasady „WS"

4.3.2.1. Adresy portali internetowych, w tym baz danych Głównego Urzędu Statystycznego

Adresy portali internetowych, które są przywoływane w artykule jedynie w celach informacyjnych, należy umieszczać w przypisach dolnych.

W przypadku korzystania z danych pobranych z baz Głównego Urzędu Statystycznego prosimy o podanie w miejscu, w którym baza jest przywoływana po raz pierwszy, pełnej nazwy bazy i jej skrótu (jeśli istnieje), nazwy jej właściciela oraz adresu internetowego w przypisie dolnym. W kolejnych przywołaniach, np. w źródle pod wykresem, należy posługiwać się już tylko pełną lub skróconą nazwą bazy.

Przykłady baz danych GUS	
pierwsze przywołanie	kolejne przywołania
Bank Danych Lokalnych (BDL) Głównego Urzędu Statystycznego + link podany w przypisie dolnym: https://bdl.stat.gov.pl	BDL
Baza Demografia Głównego Urzędu Statystycznego + link podany w przypisie dolnym: https://demografia.stat.gov.pl	Baza Demografia
Dziedzinowe Bazy Wiedzy (DBW) Głównego Urzędu Statystycznego + link podany w przypisie dolnym: https://dbw.stat.gov.pl	DBW

4.3.2.2. Akty prawne

Jeśli autor powołuje się w pracy na akty prawne, powinien za pierwszym razem podać ich pełny oficjalny tytuł; przy kolejnych przywołaniach najczęściej wystarczy nazwa skrócona. W przypadku aktów prawnych zapisanych w innym alfabecie niż łaciński tytuł trzeba poddać transkrypcji. (Informacje dotyczące miejsca publikacji aktu prawnego, takie jak numer dziennika urzędowego, należy podać tylko w opisie zamieszczonym w bibliografii załącznikowej).

Przykłady aktów prawnych	
pierwsze przywołanie	kolejne przywołania
Ustawa z dnia 29 czerwca 1995 r. o statystyce publicznej (dalej: ustawa o statystyce publicznej)	ustawa o statystyce publicznej
Rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 1260/2013 z dnia 20 listopada 2013 r. w sprawie statystyk europejskich w dziedzinie demografii (dalej: rozporządzenie nr 1260/2013)	rozporządzenie nr 1260/2013
Statistics Act	Statistics Act

4.4. Bibliografia załącznikowa w artykułach napisanych w języku polskim

4.4.1. Zasady ogólne

Bibliografia powinna być zamieszczona na końcu opracowania. Opisy bibliograficzne powinny być sporządzone w alfabecie łacińskim.

Źródła należy uszeregować alfabetycznie według nazwiska pierwszego autora, a w przypadku dwóch lub więcej prac tego samego autora – chronologicznie według roku publikacji. Prace bez znanego roku publikacji (oznaczone „b.r.") występują przed pracami ze znanym rokiem publikacji. Jeśli kilka prac tego samego autora zostało opublikowanych w tym samym roku, należy podać je w kolejności alfabetycznej według tytułu i odpowiednio oznaczyć literami a, b, c itd.

Opis bibliograficzny materiałów dostępnych w internecie powinien zawierać link prowadzący do źródłowej strony internetowej lub link DOI. Nie należy podawać linków prowadzących do baz czasopism czy repozytoriów.

4.4.2. Przykłady opisów bibliograficznych

Typ źródła	Przykład opisu bibliograficznego
Artykuł w czasopiśmie	
W wersji: drukowanej	Nazwisko, X. (rok). Tytuł artykułu. *Tytuł czasopisma*, *rocznik* (zeszyt), strona początku–strona końca.
elektronicznej, z DOI	Nazwisko, X., Nazwisko 2, Y. (rok). Tytuł artykułu. *Tytuł czasopisma*, *rocznik*(zeszyt), strona początku–strona końca. https://doi.org/xxx.
elektronicznej, bez DOI	Nazwisko, X., Nazwisko 2, Y., Nazwisko 3, Z. (rok). Tytuł artykułu. *Tytuł czasopisma*, *rocznik*(zeszyt), strona początku–strona końca. https://xxx.
Opublikowany w trybie online first	Nazwisko, X. (rok). Tytuł artykułu. *Tytuł czasopisma*. Opublikowany w trybie online first. https://xxx.
Artykuł w gazecie codziennej	
W wersji: drukowanej	Nazwisko, X. (rok, dzień i miesiąc). Tytuł artykułu. *Tytuł gazety*, strona lub strona początku–strona końca.
elektronicznej	Nazwisko, X. (rok, dzień i miesiąc). Tytuł artykułu. *Tytuł gazety*. https://xxx. Nazwisko, X. (b.r.). Tytuł artykułu. *Tytuł gazety*. https://xxx. Tytuł artykułu. (rok, miesiąc i dzień). *Tytuł gazety*. https://xxx.
Książka	
W wersji: drukowanej	Nazwisko, X. (rok). *Tytuł książki*. Wydawnictwo.
elektronicznej, z DOI	Nazwisko, X. (rok). *Tytuł książki*. Wydawnictwo. https://doi.org/xxx.
elektronicznej, bez DOI	Nazwisko, X. (rok). *Tytuł książki*. Wydawnictwo. https://xxx.
W przekładzie	Nazwisko, X. (rok). *Tytuł książki* (tłum. Y. Nazwisko). Wydawnictwo.
Wydanie wielotomowe: tom zatytułowany	Nazwisko, X. (rok). *Tytuł książki: numer tomu. Tytuł tomu*. Wydawnictwo.
tom niezatytułowany	Nazwisko, X. (rok). *Tytuł książki* (numer tomu). Wydawnictwo.
Kolejne wydanie	Nazwisko, X. (rok). *Tytuł książki* (numer wydania). Wydawnictwo.
Pod redakcją (niezależnie od języka, w którym książka została wydana)	Nazwisko, X. (red.). (rok). *Tytuł książki*. Wydawnictwo.
Przedruk lub wznowienie	Nazwisko, X. (rok). *Tytuł książki*. Wydawnictwo. (Wydanie pierwotne rok).
W przekładzie	Nazwisko, X. (rok). *Tytuł książki*. (tłum. Y. Nazwisko). Wydawnictwo. (Wydanie pierwotne rok).

Typ źródła	Przykład opisu bibliograficznego
Rozdział i hasło słownikowe/encyklopedyczne	
Rozdział w pracy zbiorowej	Nazwisko, X. (rok). Tytuł rozdziału. W: Y. Nazwisko, Z. Nazwisko 2 (red.), *Tytuł książki* (s. strona początku–strona końca). Wydawnictwo. https://doi.org/xxx lub https://xxx.
Hasło ze słownika lub z encyklopedii w wersji: drukowanej	Nazwisko autora hasła, X. (rok). Hasło. W: Y. Nazwisko (red.), *Tytuł*. Wydawnictwo. Hasło. (rok). W: Y. Nazwisko (red.), *Tytuł*. Wydawnictwo.
elektronicznej	Hasło. (rok, dzień i miesiąc lub „b.r."). W: *Tytuł* (np. *Wikipedia* lub *Słownik języka polskiego PWN*). https://xxx.
Raporty i szara literatura	
Autor: indywidualny	Nazwisko, X. (rok). *Tytuł raportu*. Wydawnictwo. https://doi.org/xxx lub https://xxx.
instytucjonalny	Nazwa instytucji. (rok). *Tytuł raportu*. Wydawnictwo (tylko jeśli wydawcą jest inna instytucja niż instytucja autorska). https://doi.org/xxx lub https://xxx.
Working papers	Nazwisko, X. (rok). *Tytuł pracy* (nazwa serii i numer). https://doi.org/xxx lub https://xxx.
Materiały z konferencji	
Opublikowane jako: druk zwarty	zob. przykład opisu książki lub rozdziału
druk ciągły	zob. przykład opisu artykułu w czasopiśmie
Niepublikowane (jedynie wygłoszone)	Nazwisko, X. (rok, dzień i miesiąc). *Tytuł pracy* [typ wystąpienia, np. referat lub prezentacja]. Nazwa i miejsce (miasto, kraj) konferencji.
Rozprawa doktorska	
Niepublikowana	Nazwisko, X. (rok). *Tytuł pracy* [niepublikowana rozprawa doktorska]. Nazwa instytucji nadającej tytuł doktorski.
Opublikowana, dostępna w internecie	Nazwisko, X. (rok). *Tytuł pracy* [rozprawa doktorska, nazwa instytucji nadającej tytuł doktorski]. https://xxx.
Maszynopis	
Niepublikowany / przygotowywany przez autora / zgłoszony do publikacji, ale jeszcze niezaakceptowany	Nazwisko, X. (rok). *Tytuł* [maszynopis niepublikowany / w przygotowaniu / zgłoszony do publikacji].
Artykuł zaakceptowany do publikacji w czasopiśmie	Nazwisko, X. (w druku). Tytuł artykułu. *Tytuł czasopisma.*
Opublikowany nieformalnie (np. na stronie internetowej autora)	Nazwisko, X., Nazwisko 2, Y. (rok). *Tytuł*. https://xxx.

Typ źródła	Przykład opisu bibliograficznego
Akt prawny[a]	
Polski i UE	Pełny tytuł aktu prawnego wraz z numerem/pozycją w dzienniku urzędowym.
Inny	Pełny tytuł aktu prawnego w języku oryginalnym (w przypadku zapisu w innym alfabecie niż łaciński należy podać tylko transkrypcję) wraz z numerem/pozycją w dzienniku urzędowym. https://xxx.
Tekst na stronie internetowej (dostępny tylko online)	
Znana data publikacji, zawartość strony się nie zmienia (jest archiwizowana)	Nazwisko, X. (rok, dzień i miesiąc). *Tytuł*. https://xxx.
Nieznana data publikacji, zawartość strony się zmienia (nie jest archiwizowana)	Nazwa instytucji. (b.r.). *Tytuł*. Pobrane dzień, miesiąc i rok pobrania z https://xxx.
Zbiór danych	
Dane opublikowane: znana data publikacji, zawartość zbioru się nie zmienia (jest archiwizowana)	Nazwisko, X. (rok). *Nazwa zbioru danych* [zbiór danych]. Wydawca. https://xxx.
nieznana data publikacji, zawartość zbioru się zmienia (nie jest archiwizowana)	Nazwa instytucji. (b.r.). *Nazwa zbioru danych* [zbiór danych]. Wydawca (tylko jeśli wydawcą jest inna instytucja niż instytucja autorska / właściciel danych). Pobrane dzień, miesiąc i rok pobrania z https://xxx.
Materiały audiowizualne	
Nagranie wideo	Nazwisko, X. (rok, dzień i miesiąc). *Tytuł* [wideo]. Nazwa kanału, na którym nagranie zostało udostępnione (np. YouTube). https://xxx.
Webinar	Nazwisko, X. (rok, dzień i miesiąc). *Tytuł* [webinar]. Nazwa instytucji. https://xxx.
Posty w serwisach społecznościowych	
Post na portalu X lub Instagramie	Nazwisko, X. lub nazwa instytucji [@nazwa użytkownika] (rok, dzień i miesiąc). *Treść – do 20 wyrazów* [post]. Nazwa serwisu społecznościowego (X lub Instagram). https://xxx.
Post na Facebooku	Nazwisko, X. lub nazwa instytucji (rok, dzień i miesiąc). *Treść – do 20 wyrazów* [post]. Facebook. https://xxx. Nazwa instytucji [nazwa użytkownika] (rok, dzień i miesiąc). *Treść – do 20 wyrazów* [post]. Facebook. https://xxx.

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

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

[a] Wewnętrzne zasady „WS".

STAŁE DZIAŁY „WS" – ZAKRES TEMATYCZNY
PERMANENT SECTIONS OF *WS* – THEMATIC SCOPE

Tematy artykułów	Topics of the articles
Studia metodologiczne / Methodological studies	
• Oryginalne lub udoskonalone rozwiązania metodologiczne, które mogą znaleźć zastosowanie w analizach statystycznych i służyć podnoszeniu ich jakości • Projektowanie badań statystycznych	• Original or developed methodological solutions which can be applied to statistical analyses and serve to improve their quality • Planning statistical surveys
Statystyka w praktyce / Statistics in practice	
• Nowatorskie zastosowania narzędzi i modeli statystycznych oraz analiza i ocena statystyczna zjawisk społeczno-gospodarczych i innych, prowadzona w szczególności na danych pochodzących z zasobów statystyki publicznej • Wykorzystanie narzędzi informatycznych do uzyskiwania i przetwarzania informacji statystycznych, naliczania i kontroli ujawniania danych oraz prezentacji i rozpowszechniania danych wynikowych	• Innovative applications of statistical tools and models as well as statistical analysis and assessment of social, economic and other phenomena, performed mainly on data produced by official statistics • Application of IT tools to obtain and process statistical information, to calculate data and control the statistical disclosure, and to present and disseminate output data
Studia interdyscyplinarne. Wyzwania badawcze / Interdisciplinary studies. Research challenges	
• Wyzwania badawcze wynikające z rosnących potrzeb użytkowników danych statystycznych i wymagające stosowania rozwiązań z różnych dziedzin nauki • Problematyka wykraczająca poza konwencjonalne tematy związane ze statystyką • Wyniki badań prowadzonych w obrębie różnych dyscyplin z wykorzystaniem metod statystycznych	• Research challenges resulting from growing needs of statistical data users and requiring the application of solutions from various fields of science • Problems beyond the conventional thematic scope related to statistics • Results of research carried out in the framework of several fields of science using statistical methods
Edukacja statystyczna / Statistical education	
• Metody i efekty nauczania statystyki na wszystkich poziomach edukacji • Popularyzacja myślenia statystycznego i rzetelnego posługiwania się informacjami statystycznymi	• Methods and effects of statistical education at all levels of education • Popularisation of statistical thinking and of diligent use of statistical information
Spisy powszechne – problemy i wyzwania / Issues and challenges in census taking	
• Rozwiązania metodologiczne i organizacyjne możliwe do zastosowania podczas przygotowywania i prowadzenia spisów • Praktyczne aspekty związane z gromadzeniem i udostępnianiem danych ze spisów, w tym dotyczące obciążenia odpowiedzi i ochrony tajemnicy statystycznej	• Methodological and organisational solutions which may be implemented in the process of preparing and conducting censuses • Practical aspects of collecting and disseminating census data, including those related to response burden and the protection of statistical confidentiality
Z dziejów statystyki / From the history of statistics	
• Historia prowadzenia obserwacji statystycznych, w tym rozwój metodologii i narzędzi oraz instytucji statystycznych w Polsce i za granicą • Życie i osiągnięcia wybitnych statystyków	• History of statistical observations, including the development of statistical methodologies, tools and institutions in Poland and abroad • Life and achievements of prominent statisticians
In memoriam	
• Nekrologi i artykuły wspomnieniowe o osobach zasłużonych dla statystyki	• Obituaries and articles remembering important people in the world of statistics
Dyskusje. Recenzje. Informacje / Discussions. Reviews. Information	
• Dyskusje i polemiki • Sprawozdania z konferencji naukowych • Recenzje książek oraz zestawienia nowości wydawniczych GUS	• Discussions and polemics • Reports from scientific conferences • Book reviews and compilations of Statistics Poland's new publications