

Cena 15,00 zł
(VAT 8%)

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

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

MARZEC / MARCH
ROCZNIK / VOLUME 69

2024 | 3

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

POLSKIE TOWARZYSTWO STATYSTYCZNE
POLISH STATISTICAL ASSOCIATION



WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

MARZEC / MARCH
ROCZNIK / VOLUME 69

2024 | 3 (754)

RADA NAUKOWA / SCIENTIFIC COUNCIL

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

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

KOLEGIUM REDAKCYJNE / EDITORIAL BOARD

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

ZESPÓŁ REDAKCYJNY / EDITORIAL STAFF

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

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

redaktorzy tematyczni / thematic editors: Andrzej Młodak, Małgorzata Tarczyńska-Łuniewska, Wioletta Wrzaszcz, Agnieszka Zgierska

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

ADRES REDAKCJI / EDITORIAL OFFICE ADDRESS

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

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

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

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

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



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

SPIS TREŚCI
CONTENTS

Od redakcji	IV
From the editorial team	
Statystyka w praktyce	
Statistics in practice	
Dorota Rozmus	
Miara stabilności w wyborze liczby grup w taksonomii zagregowanej z zastosowaniem analizy spektralnej i metody propagacji podobieństwa	1
Stability measure in the selection of the number of groups in aggregate taxonomy using spectral analysis and the method of affinity propagation	
Wioletta Grzenda, Agnieszka Marszałek	
Why do young workers quit their first job? Identification of the risk factors using the Cox model and survival trees	18
Dlaczego młodzi pracownicy porzucają pierwszą pracę? Identyfikacja czynników ryzyka z wykorzystaniem modelu Coxa i drzew przetrwania	
Studia interdyscyplinarne. Wyzwania badawcze	
Interdisciplinary studies. Research challenges	
Norbert Duczkowski, Adam Śliwiński, Lubomir Słowik	
Impact of Hofstede's cultural dimensions on insurance demand	38
Wpływ wymiarów kulturowych Hofstede'go na popyt na ubezpieczenia	
Dyskusje. Recenzje. Informacje	
Discussions. Reviews. Information	
Józef Oleński, Jacek Kowalewski	
Recenzja monografii naukowej Andrzeja Młodaka, Michała Pietrzaka, Tomasza Klimanka, Tomasza Józefowskiego i Pawła Lańducha <i>Poufność a użyteczność informacji statystycznych. Dylematy ochrony udostępnianych danych</i>	69
Review of a scientific monograph by Andrzej Młodak, Michał Pietrzak, Tomasz Klimanek, Tomasz Józefowski and Paweł Lańduch, entitled <i>Confidentiality vs. usefulness of statistical information. Dilemmas in protecting shared data</i>	
Joanna Sadowy	
Wydawnictwa GUS. Luty 2024	73
Publications of Statistics Poland. February 2024	
Dla autorów	75
For the authors	
Działy „WS” – tematyka artykułów	86
WS sections – topics of the articles	

OD REDAKCJI

W marcowym numerze „Wiadomości Statystycznych. The Polish Statistician” proponujemy Państwu lekturę trzech artykułów, w których przedstawiono wyniki badań przeprowadzonych z użyciem metod statystycznych.

Dr Dorota Rozmus w pracy *Miara stabilności w wyborze liczby grup w taksonomii zagregowanej z zastosowaniem analizy spektralnej i metody propagacji podobieństwa* porównuje wyniki wyboru optymalnej liczby grup w taksonomii zagregowanej na przykładzie realizacji trzech Celów Zrównoważonego Rozwoju w krajach UE. W badaniu, opierającym się na danych Eurostatu za 2019 r., autorka stosuje miarę PAC i wybrane klasyczne indeksy, tj. Calińskiego-Harabasz, Dunna i Daviesa-Bouldina, a w podejściu zagregowanym – propagację podobieństwa i taksonomię spektralną. Z analizy wynika, że klasyczne indeksy różnią się pod względem wskazywania optymalnej liczby grup. Tego typu rozbieżności można zaobserwować w ramach tej samej metody wybranej jako bazowa do utworzenia macierzy zgodności, ale także gdy porównuje się różne metody służące jako bazowe do utworzenia taksonomii zagregowanej. Autorka stwierdza, że opracowanie specjalnej miary stabilności w odniesieniu do podejścia zagregowanego w taksonomii nie rozwiązuje problemu niejednoznacznego wskazywania liczby grup, na które należy podzielić zbiór obiektów.

Artykuł dr hab. Wioletty Grzendy, prof. SGH, i mgr Agnieszki Marszałek *Why do young workers quit their first job? Identification of the risk factors using the Cox model and survival trees* dotyczy zjawiska porzucania pierwszej pracy przez młodych pracowników. Celem podjętego badania jest określenie i ocena indywidualnych charakterystyk młodych pracowników oraz czynników związanych z wykonywaną pracą, które mają wpływ na czas trwania pierwszego zatrudnienia. Autorki korzystają z danych za lata 2019 i 2020 pochodzących z Badania Aktywności Ekonomicznej Ludności przeprowadzonego przez GUS. Stosują model Coxa ze zmiennymi zależnymi od czasu. Użyte wyniki wskazują, że większe ryzyko porzucenia pierwszej pracy istnieje w przypadku osób z wyższym wykształceniem, a duży wpływ mają na to także rodzaj umowy o pracę, tygodniowy czas pracy i pełnienie funkcji kierowniczych.

Impact of Hofstede's cultural dimensions on insurance demand to temat pracy mgr. Norberta Duczkowskiego, prof. dr. hab. Adama Śliwińskiego i mgr. Lubomira Słowika. Autorzy weryfikują zależności między wymiarami kulturowymi zdefiniowanymi przez holenderskiego psychologa społecznego (tj. dystans władzy, indywidualizm i kolektywizm, męskość i kobiecość, unikanie niepewności, orientacja długoterminowa i krótkoterminowa, pobłażanie i powściągliwość) a popytem na różnego rodzaju ubezpieczenia w wybranych krajach europejskich. W badaniu wykorzystują dane za lata 2004–2020 pobrane z serwisu Insurance Europe i ze strony Geert Hofstede. Na podstawie analizy stwierdzają brak korelacji lub bardzo słabą korelację w przypadku większości analizowanych wymiarów kulturowych i popytu na ubezpieczenia.

Oprócz artykułów naukowych w tym numerze zamieszczamy recenzję książki Andrzeja Młodaka, Michała Pietrzaka, Tomasza Klimanka, Tomasza Józefowskiego i Pawła Lańducha *Poufność a użyteczność informacji statystycznych. Dylematy ochrony udostępnianych danych*. Recenzenci – prof. dr hab. Józef Oleński i dr Jacek Kowalewski – podkreślają, że monografia dostarcza gruntownej wiedzy na temat poufności informacji w badaniach statystycznych.

Zachęcamy także do zapoznania się z zestawieniem najnowszych publikacji GUS opracowanym przez Joannę Sadowy.

Życzymy miłej lektury.

FROM THE EDITORIAL TEAM

The March issue of *Wiadomości Statystyczne. The Polish Statistician* features three articles presenting the results of research conducted using statistical methods.

In her article, entitled *Stability measure in the selection of the number of groups in aggregate taxonomy using spectral analysis and the method of affinity propagation*, Dorota Rozmus, PhD, compares the results of selecting an optimal number of groups in aggregate taxonomy on the example of the implementation of three Sustainable Development Goals in EU countries. The study, which uses Eurostat data relating to the year 2019, is based on the PAC measure and selected classic indices, namely the Caliński-Harabasz, the Dunn and the Davies-Bouldin indices. In the aggregate approach, the author applies the affinity propagation method and spectral clustering. The results of the analysis demonstrate that the classic indices differ from the PAC measure in terms of indicating the optimal number of groups. This type of discrepancy may be observed within the same method chosen to serve as the base method for the construction of a compliance matrix and while comparing various methods acting as base methods for creating an aggregate taxonomy. The author concludes that the development of a special stability measure with regard to the aggregate approach in taxonomy provides no solution to the problem of the ambiguous indication of the number of groups that a set of objects should be divided into.

Why do young workers quit their first job? Identification of the risk factors using the Cox model and survival trees by Wioletta Grzenda, PhD, DSc, professor at the SGH Warsaw School of Economics, and Agnieszka Marszałek, MSc, examines the phenomenon of young workers quitting their first job. The aim of the research is to determine and evaluate individual characteristics of young employees and work-related factors which affect the duration of the first employment. The authors use data for 2019 and 2020 collected during Statistics Poland's Labour Force Survey and apply the Cox model with time-variant variables. The obtained results indicate that people with higher education were more likely to resign from their first job than less-educated workers. The type of employment contract, weekly working hours and holding managerial positions have a major influence on the risk of young people quitting their first job.

In their work, Norbert Duczkowski, MSc, Adam Śliwiński, PhD, DSc, ProfTit, and Lubomir Słowik, MSc, study the *Impact of Hofstede's cultural dimensions on insurance demand*. The authors verify the relationships between the cultural dimensions defined by the Dutch social psychologist (i.e. power distance, individualism vs. collectivism, masculinity vs. femininity, uncertainty avoidance, long-term vs. short-term normative orientation, and indulgence vs. restraint) and the demand for different types of insurance in selected European countries. The study is based on data for 2004–2020 downloaded from the Insurance Europe and the Geert Hofstede websites. The analysis shows no or a very weak correlation between most of the analysed cultural dimensions and insurance demand.

This issue also presents a review of a book by Andrzej Młodak, Michał Pietrzak, Tomasz Klimanek, Tomasz Józefowski and Paweł Lańdych, entitled *Confidentiality vs. usefulness of statistical information. Dilemmas in protecting shared data*. The reviewers, Józef Oleński, PhD, DSc, ProfTit, and Jacek Kowalewski, PhD, find this monograph to be a source of thorough knowledge about confidentiality of information in statistical research.

We would also like to recommend Joanna Sadowy's compilation of the most recent publications of Statistics Poland.

We wish you pleasant reading.

Miara stabilności w wyborze liczby grup w taksonomii zagregowanej z zastosowaniem analizy spektralnej i metody propagacji podobieństwa

Dorota Rozmus^a

Streszczenie. Od lat 90. XX w. częstymi tematami rozważań w dziedzinie taksonomii są podejście zagregowane i stabilność metod grupowania. Dotychczas były one rozpatrywane osobno, ale w ostatnim czasie pojawiła się w literaturze propozycja połączenia tych dwóch pojęć – miara stabilności (ang. *proportion of ambiguously clustered pairs* – PAC), którą można zastosować w podejściu zagregowanym w taksonomii i która ma służyć jako kryterium wyboru optymalnej liczby grup. Celem badania omawianego w artykule jest porównanie wyników wyboru optymalnej liczby grup w taksonomii zagregowanej na przykładzie realizacji trzech Celów Zrównoważonego Rozwoju w krajach UE. Wykorzystano miarę PAC i wybrane klasyczne indeksy: Calińskiego-Harabasz, Dunna i Daviesa-Bouldina. Jako metody bazowe w podejściu zagregowanym zastosowano propagację podobieństwa (ang. *affinity propagation method*) i taksonomię spektralną (ang. *spectral clustering*). Badanie opierało się na danych z bazy Eurostatu za 2019 r. Uzyskane rezultaty świadczą o tym, że zarówno wybór kryterium ustalania liczby grup, jak i metody bazowej w taksonomii zagregowanej wpływają na ostateczne rozstrzygnięcie dotyczące ustalenia liczby grup. Bez względu na to, czy stosowano metodę propagacji podobieństwa czy taksonomię spektralną z klasycznymi indeksami, albo też metody te wykorzystywano jako bazowe w podejściu zagregowanym i wybierano liczbę grup za pomocą miary PAC, rozbieżności we wskazywanej liczbie grup okazywały się bardzo duże.

Słowa kluczowe: podejście zagregowane w taksonomii, stabilność, taksonomia spektralna, metoda propagacji podobieństwa

JEL: C38

Stability measure in the selection of the number of groups in aggregate taxonomy using spectral analysis and the method of affinity propagation

Abstract. Since the 1990s, the aggregate approach and the stability of grouping methods have been concepts frequently discussed in the field of taxonomy. So far, they have been considered separately, but recently, there has been a postulate in the literature to combine these two concepts in the form of the PAC (proportion of ambiguously clustered pairs) stability measure, which can be used in the aggregate approach to taxonomy and which is intended to serve as

^a Uniwersytet Ekonomiczny w Katowicach, Wydział Finansów, Polska / University of Economics in Katowice, Faculty of Finance, Poland. ORCID: <https://orcid.org/0000-0002-0565-5319>.
E-mail: dorota.rozmus@ue.katowice.pl.

a criterion for selecting an optimal number of groups. The aim of the research presented in this article is to compare the results of the selection of the optimal number of groups in aggregate taxonomy on the example of the attainment of three Sustainable Development Goals (SDGs) by EU countries. The PAC measure and selected classic indices, namely the Caliński-Harabasz, the Dunn and the Davies-Bouldin indices were applied for this purpose. The affinity propagation method and spectral clustering served as base methods in the aggregate approach. The study used Eurostat data for 2019. The obtained results demonstrate that both the choice of the criterion for determining the number of groups and the choice of the base method in aggregate taxonomy have an influence on the final decision on how to determine the number of groups. Regardless of whether the affinity propagation method or spectral taxonomy with classic indices was used, or whether these methods were used as base ones in the aggregate approach and the number of groups was selected using the PAC measure, the differences between the indicated numbers of groups were very large.

Keywords: aggregate approach to taxonomy, stability, spectral taxonomy, affinity propagation method

1. Wprowadzenie

Duża dokładność wyników grupowania stanowi kluczowy cel analizy taksonomicznej. Precyzja w tym zakresie przesądza o słuszności decyzji podejmowanych na podstawie uzyskanych rezultatów. Przez dokładność grupowania należy rozumieć zdolność rozpatrywanej metody analizy skupień do rozpoznawania rzeczywistej struktury grup – pod warunkiem że znana jest rzeczywista struktura grup (w literaturze opisano wiele indeksów badających zgodność między rzeczywistą strukturą grup a tą, którą uzyskuje się w wyniku zastosowania dowolnej metody taksonomicznej). Gdy struktura grup nie jest jednak znana *a priori*, jakość grupowania można ocenić tylko przez pryzmat homogeniczności (wewnętrznej jednorodności skupień) i heterogeniczności (odrębności skupień). Wciąż proponowane są metody i rozwiązania umożliwiające uzyskanie dokładniejszych wyników niż te, które można otrzymać dzięki tradycyjnym algorytmom grupowania (np. k -średnich, k -medoidów i metod hierarchicznych). Przykładami takich rozwiązań są metody taksonomii zagregowanej (Dudoit i Fridlyand, 2003; Fred i Jain, 2002; Hornik, 2005; Kuncheva i Vetrov, 2006; Leisch, 1999; Monti i in., 2003).

Pożądaną cechą dowolnego algorytmu taksonomicznego jest również stabilność danej metody względem niewielkich zakłóceń występujących w zbiorze danych (np. wybór podzbioru danych, małe wahania wartości cech) lub parametrów algorytmu (np. losowa inicjalizacja algorytmu). Nieformalnie kryterium stabilności zakłada, że jeśli metoda taksonomiczna jest wielokrotnie stosowana w przypadku niezależnych podprób (przy niezmiennych parametrach algorytmu) i daje podobne wyniki grupowania, to można je uznać za stabilne i odzwierciedlające rzeczywistą strukturę grup (Shamir i Tishby, 2008). Volkovich i in. (2010) stwierdzają nawet, że liczba grup, która maksymalizuje stabilność grupowania, może służyć jako informacja

o rzeczywistej liczbie grup w zbiorze danych. Z tego powodu to kryterium stosuje się najczęściej do określania liczby grup, na które zostanie podzielony zbiór obiektów.

Zagraniczni badacze proponują wiele sposobów pomiaru stabilności (np. Ben-Hur i Guyon, 2003; Brock i in., 2008; Fang i Wang, 2012; Henning, 2007; Lord i in., 2017; Marino i Presti, 2019; Suzuki i Shimodaira, 2006). W Polsce prace z zakresu zaproponowanych metod zagregowanych i metod badania stabilności opublikowała Rozmus (2011, 2013), a za granicą – Leone i in. (2007), Meng i in. (2016) oraz Yu i in. (2015). Propozycję miary stabilności w taksonomii zagregowanej przedstawili natomiast Şenbabaoglu i in. (2014). Na podstawie schematu agregacji – bazującego na macierzy współwystąpień (ang. *consensus matrix*) – oraz kryterium doboru liczby grup (empiryczna funkcja dystrybuanty [ang. *cumulative distribution function* – CDF] i proporcjonalna zmiana powierzchni pod CDF [Δk]) zaproponowanych przez Montiego i in. (2003) wprowadzili miarę stabilności PAC (ang. *proportion of ambiguously clustered pairs*), określającą udział niejednoznacznych grupowań par obiektów przy wielokrotnie dokonywanym podziale zbioru obiektów na grupy podobnych do siebie obserwacji.

Celem badania omawianego w artykule jest porównanie wyników wyboru wartości parametru k w taksonomii zagregowanej na przykładzie realizacji trzech Celów Zrównoważonego Rozwoju w krajach UE. Wykorzystano miarę PAC i wybrane klasyczne indeksy służące do określania liczby grup, a mianowicie indeksy Calińskiego-Harabasz, Dunna i Daviesa-Bouldina. Jako metody bazowe w podejściu zagregowanym zastosowano propagację podobieństwa (ang. *affinity propagation method*) i taksonomię spektralną (ang. *spectral clustering*).

2. Grupowanie na podstawie macierzy współwystąpień

Grupowanie na podstawie macierzy współwystąpień jest jednym z przykładów podejścia zagregowanego w taksonomii. Polega ono na tym, że wykorzystując wybraną metodę taksonomiczną (np. k -średnich) jako metodę bazową, dokonuje się wielokrotnego grupowania danych – za każdym razem na innym podzbiorze obiektów wybranym z pierwotnego zbioru danych. Warto nadmienić, że nie jest to nowa koncepcja, pierwotnie została bowiem opracowana przez Fred i Jaina (2002, 2005) pod nazwą *co-association matrix*.

Algorytm grupowania zagregowanego z wykorzystaniem macierzy współwystąpień wymaga, aby w pierwszym kroku określić B podprób bootstrapowych \mathbf{X}^b ($b = 1, 2, \dots, B$), które zostały wylosowane z pierwotnego zbioru danych $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, gdzie N oznacza liczbę obserwacji w zbiorze danych. Dla każdego grupowania z osobna wyznaczana jest macierz zgodności \mathbf{C}^b (ang. *connectivity matrix*). Macierze tego rodzaju posłużą potem do utworzenia macierzy współwystąpień \mathbf{Q}^k .

Macierz zgodności \mathbf{C}^b jest $(N \times N)$ -wymiarową macierzą konstruowaną po zastosowaniu pojedynczego algorytmu taksonomicznego do podzbioru \mathbf{X}^b i na podstawie grupowania uzyskanego za pomocą tego algorytmu. Jej elementy definiuje się następująco:

$$C^b(i, j) = \begin{cases} 1, & \text{gdy obiekty } i \text{ oraz } j \text{ należą do tej samej grupy,} \\ 0 & \text{w przeciwnym przypadku.} \end{cases} \quad (1)$$

Z kolei macierz współwystąpień \mathbf{Q}^k – przy podziale na k grup – jest $(N \times N)$ -wymiarową macierzą, której elementem na przecięciu i -tego wiersza i j -ej kolumny jest odsetek grupowań (wśród B podzbiorów), w których i -ty oraz j -y element należą do tej samej grupy. Elementy macierzy współwystąpień, oznaczane jako $Q(i, j)$, są nazywane indeksami współwystąpień (ang. *consensus index*). Stanowią znormalizowane sumy elementów macierzy zgodności po wszystkich podziorach \mathbf{X}^b ($b = 1, 2, \dots, B$):

$$Q(i, j) = \frac{\sum_b C^b(i, j)}{\sum_b I^b(i, j)}, \quad (2)$$

gdzie \mathbf{I}^b jest $(N \times N)$ -wymiarową macierzą wskaźnikową, której element z i -tego wiersza oraz j -ej kolumny $I^b(i, j)$ przybiera wartość 1, gdy obiekty i -ty oraz j -y znalazły się w podpróbie \mathbf{X}^b , lub 0 – w przeciwnym przypadku.

Konieczność uwzględniania macierzy wskaźnikowej \mathbf{I}^b wynika z tego, że w schemacie tego podejścia zagregowanego losowane są próby \mathbf{X}^b , które nie zawierają wszystkich obserwacji z pierwotnego zbioru danych.

Każdy element macierzy \mathbf{Q}^k jest liczbą z przedziału $[0, 1]$. Doskonała zgodność pojawi się w momencie, gdy wszystkie elementy tej macierzy będą równe 0 lub 1. Ta własność macierzy współwystąpień zainspirowała Şenbabaoğlu i in. (2014) do zaproponowania miary stabilności w podejściu zagregowanym, tj. miary PAC, czyli proporcji niejednoznacznie grupowanych par obiektów.

3. Miara stabilności PAC

Znając macierz współwystąpień \mathbf{Q}^k , można sporządzić histogram dla $\frac{N(N-1)}{2}$ wartości indeksów współwystąpień $Q(i, j)$, dla $i < j$, zawartych w tej macierzy. Doskonałą zgodność wszystkich par obiektów wśród B podziałów (wykr. 1) reprezentują dwa słupki odpowiadające wartościom 0 i 1. Słupki między 0 i 1, w środkowej części histogramu (wykr. 2), oznaczają wszelkie niezgodności w grupowaniu par obiektów.

Kolejnym krokiem w konstruowaniu miary PAC jest wyznaczenie odpowiedniej CDF w przedziale $[0, 1]$. Dla zadanego histogramu jest ona definiowana jako:

$$CDF(x) = \frac{\sum_{i < j} I\{Q(i, j) \leq x\}}{\frac{N(N-1)}{2}}, \quad (3)$$

gdzie $I\{\cdot\}$ jest funkcją wskaźnikową, a N to liczba obserwacji w zbiorze danych.

Po przeanalizowaniu wykresu funkcji dystrybuanty nasuwają się następujące uwagi dotyczące miary PAC:

- lewa dolna część wykresu reprezentuje punkty z pierwotnego zbioru danych, które przy wielokrotnym grupowaniu nigdy nie znalazły się w tej samej grupie lub rzadko się w niej znajdowały;
- prawa górna część reprezentuje obserwacje, które przy wielokrotnym grupowaniu zawsze albo często umieszczane są w tej samej grupie;
- środkowa część wykresu reprezentuje obserwacje, których współwystępowanie w tej samej grupie różni się w zależności od wybranego podzbioru danych.

Na podstawie wyżej wymienionych własności można zatem stwierdzić, że wykres funkcji dystrybuanty odzwierciedla optymalną wartość parametru k , czyli liczby grup. Dla poprawnie dobranej wartości k wykres dystrybuanty powinien być bowiem płaski w środkowej części. Na tej podstawie miara PAC (Şenbabaoğlu i in., 2014) została zdefiniowana jako frakcja par obiektów, dla których indeks $Q(i, j)$ znajduje się w przedziale (x_1, x_2) , który to przedział zawiera się między 0 a 1:

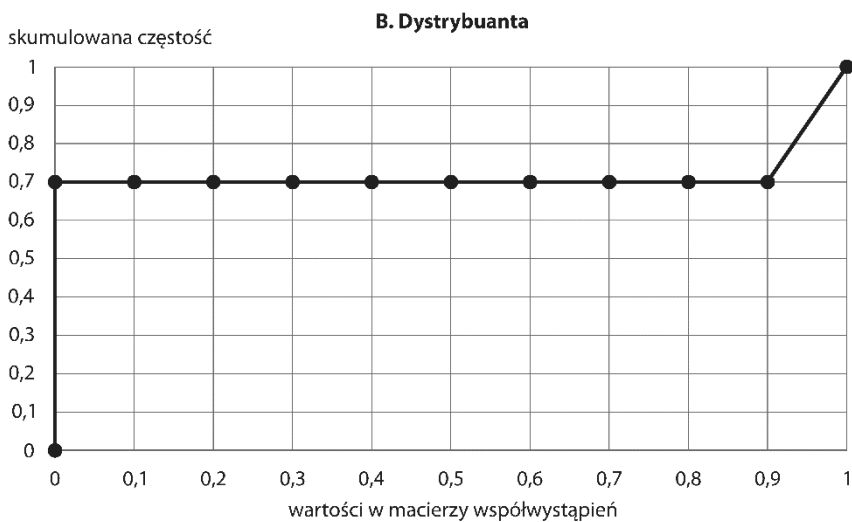
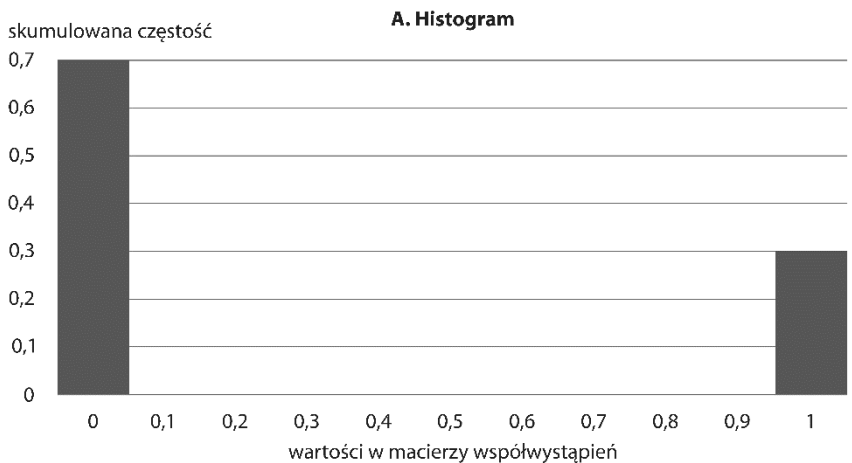
$$PAC_k(x_1, x_2) = CDF_k(x_2) - CDF_k(x_1). \quad (4)$$

We wzorze (4) x_1 i x_2 to progi zależne od danych, które zazwyczaj są wybierane odpowiednio w pobliżu 0 i 1. W swoim badaniu Şenbabaoğlu i in. (2014) przyjęli wartości równe odpowiednio $x_1 = 0,1$ i $x_2 = 0,9$. Takie same wartości zostały przyjęte dla x_1 i x_2 w omawianym badaniu.

Niskie wartości miary PAC sugerują, że środkowa część wykresu dystrybuanty jest płaska. Optymalną wartość parametru k dobieramy zatem jako:

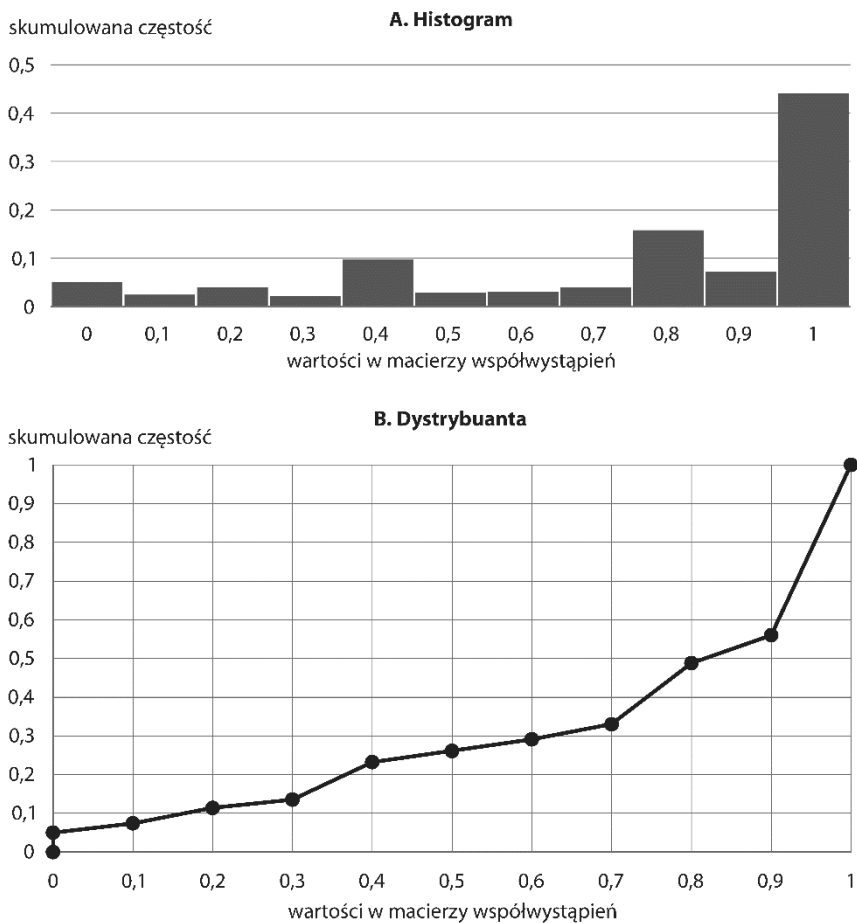
$$k_{opt} = \arg \min_k PAC_k. \quad (5)$$

Wykr. 1. Histogram wartości zawartych w macierzy współwystąpień reprezentujący doskonałą zgodność grupowania wszystkich par obiektów wśród B podziałów i odpowiadająca mu dystrybuanta



Źródło: opracowanie własne.

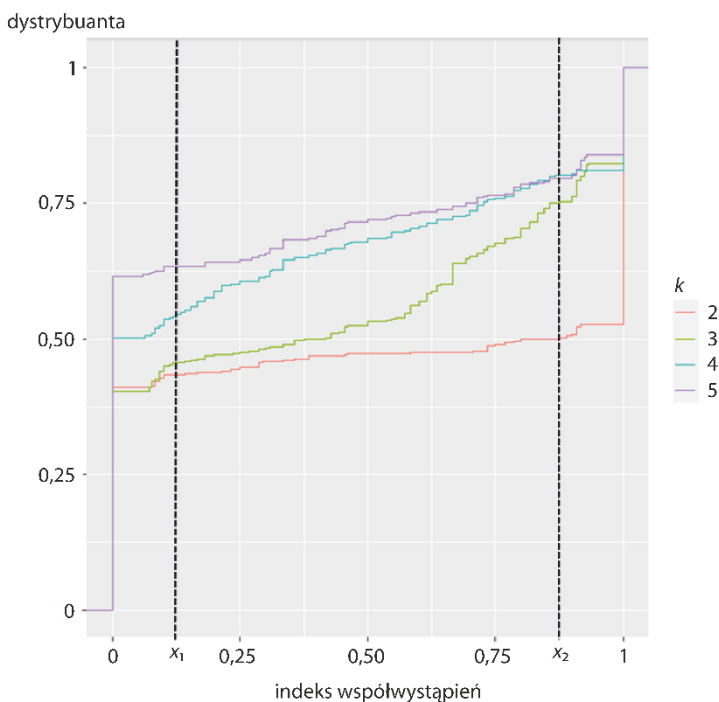
Wykr. 2. Histogram wartości zawartych w macierzy współwystąpień reprezentujący wszelkie niezgodności w grupowaniu wszystkich par obiektów wśród B podziałów i odpowiadająca mu dystrybuanta



Źródło: opracowanie własne.

Na wyk. 3 zilustrowano koncepcję struktury miary PAC. Indeks współwystąpień, dany wzorem (2), oraz wartości x_1 i x_2 są dobierane przez badacza (standardowe ustawienia to $x_1 = 0,1$ i $x_2 = 0,9$).

Wykr. 3. Koncepcja miary stabilności PAC



Uwaga. x_1 i x_2 – granice przedziału, w którym rozpatrywana jest wartość indeksu współwystąpień;
 k – potencjalnie rozpatrywana liczba grup, dla których będzie wyliczana wartość miary PAC, $k_{opt}=2$.
 Źródło: opracowanie własne.

Na podstawie wykresów CDF dla różnych wartości parametru ($k = 2, 3, 4$ i 5) można zauważyć, że najbardziej płaski kształt w swojej środkowej części przyjmuje dystrybuanta dla $k = 2$. Stąd wniosek, że zbiór rozpatrywanych obiektów powinien zostać podzielony na dwie grupy.

4. Taksonomiczna metoda propagacji podobieństwa

Frey i Dueck (2007) opisują taksonomiczną metodę propagacji podobieństwa (ang. *affinity propagation method*) jako algorytm, który identyfikuje reprezentantów (ang. *exemplars*) wśród obiektów w zbiorze danych i tworzy wokół nich grupy obiektów podobnych. Istota tej metody polega na jednoczesnym traktowaniu wszystkich obiektów w zbiorze jako potencjalnych reprezentantów. Wymiana informacji po-

między obiektami trwa aż do momentu, gdy uda się znaleźć odpowiedni zbiór reprezentantów i odpowiadający mu podział obiektów. Celem tej metody jest maksymalizacja sumy podobieństw między obiektami i ich reprezentantami.

Metoda propagacji podobieństwa składa się z siedmiu kroków (Frey i Dueck, 2007):

1. Określenie macierzy podobieństw między obiektami przy zastosowaniu ujemnego kwadratu odległości euklidesowej:

$$s(i, j) = -\|x_i - x_j\|^2. \quad (6)$$

2. Ustalenie preferencji wejściowej (ang. *preference*):

$$s(j, j) = p. \quad (7)$$

która dla każdej obserwacji wskazuje tendencję do tego, by stała się ona reprezentantem (Bodenhofer i in., 2011).

Preferencja wejściowa jest (oprócz macierzy podobieństwa – zob. część 4) najważniejszym parametrem wejściowym metody. Można ją wybrać indywidualnie dla każdej obserwacji, a także przyjąć taką samą wartość dla wszystkich obserwacji w zbiorze danych. Co ważne, preferencje wejściowe w dużej mierze determinują liczbę skupień. Frey i Dueck (2007, s. 972) wprowadzili następującą praktyczną zasadę: „Wspólną wartością może być mediana podobieństw w macierzy podobieństw (skutkująca umiarkowaną liczbą skupień) lub ich minimum (skutkujące małą liczbą skupień)”.

3. Wymiana między obserwacjami dwóch rodzajów informacji:
 - odpowiedzialność (ang. *responsibility*) $r(i, j)$ – przenoszona od obserwacji x_i do potencjalnego reprezentanta x_j ; odzwierciedla, w jakim stopniu obserwacja otrzymująca informację powinna być reprezentantem dla obserwacji wysyłającej informację, biorąc pod uwagę informację płynącą z innych punktów będących potencjalnymi reprezentantami dla obserwacji x_i ;
 - osiągalność (ang. *availability*) $a(i, j)$ – przekazywana od potencjalnego reprezentanta x_j do obserwacji x_i ; odzwierciedla, w jakim stopniu obserwacja wysyłająca informację powinna być reprezentantem dla obserwacji otrzymującej informację, biorąc pod uwagę informację płynącą z innych obserwacji, dla których x_j jest potencjalnym reprezentantem.
4. Zainicjowanie algorytmu przy założeniu, że początkowe wartości osiągalności przyjmują wartość $a(i, j) = 0$.
5. Obliczanie wartości odpowiedzialności i osiągalności z formuły:

$$r(i, j) \leftarrow s(i, j) - \max_{j' \neq j} \{a(i, j') + s(i, j')\}, \quad (8)$$

gdzie

$$a(i, j) = \begin{cases} \min \left\{ 0, r(j, j) + \sum_{i' \in \{i, j\}} \max \{0, r(i', j)\} \right\}, & \text{gdy } i \neq j \\ \sum_{i' \neq j} \max \{0, r(i', j)\}, & \text{gdy } i = j. \end{cases} \quad (9)$$

6. Zakończenie procedury przesyłania informacji:

- po określonej liczbie iteracji;
- jeśli zmiana w przesyłanej informacji będzie mniejsza niż z góry ustalona wartość;
- gdy przez określoną liczbę iteracji nie zachodzą zmiany w przesyłanej informacji.

7. Przydział obiektów do skupień $\hat{\mathbf{c}} = (\hat{c}_1, \dots, \hat{c}_k)$ według reguły:

$$\hat{c}_k = \arg \max_k [a(i, j) + r(i, j)], \quad (10)$$

gdzie \hat{c}_k jest reprezentantem skupienia, do którego przydzielona jest obserwacja x_i .

W czasie implementacji rozpatrywanej metody w bibliotece `apcluster` w programie R domyślnie korzysta się z reguły mediany, jeśli użytkownik nie poda niestandardowej reguły doboru wartości preferencji wejściowych. W obliczeniach przeprowadzonych w omawianym badaniu przyjęto wartość domyślną dla p .

5. Taksonomia spektralna

W taksonomii spektralnej stosuje się wartości własne pochodzące ze spektralnej dekompozycji macierzy podobieństwa opisującej badane obiekty. Następnie największe wartości własne i odpowiadające im wektory własne są wykorzystywane do ostatecznego podziału obserwacji. W literaturze zaproponowano kilka metod spektralnych, z których każda w nieco inny sposób stosuje wektory własne (Kannan i in., 2004; Ng i in., 2001; Shi i Malik, 2000).

W niniejszym badaniu zastosowana została metoda, którą zaproponowali Ng i in. (2001).

Dany jest zbiór obserwacji $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, który należy podzielić na k grup. Odbywa się to w sześciu etapach:

1. Konstrukcja macierzy podobieństwa (ang. *affinity matrix*) $\mathbf{A} \in R^{N \times N}$, której elementy zostały zdefiniowane jako:

$$A_{ij} = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right), \quad (11)$$

gdy $i \neq j$ oraz $A_{ii} = 0$ a σ to parametr skalujący dobierany przez badacza.

2. Definiowanie \mathbf{D} jako macierzy diagonalnej, której element (i, i) jest sumą i -tego wiersza macierzy \mathbf{A} , i konstrukcja macierzy:

$$\mathbf{L} = \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{-1/2}. \quad (12)$$

3. Znalezienie k pierwszych wektorów własnych $(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k)$ macierzy \mathbf{L} i zestawienie ich w kolumnach. Konstrukcja macierzy:

$$\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k] \in R^{N \times k}. \quad (13)$$

4. Konstrukcja macierzy \mathbf{Y} poprzez normalizację każdego wiersza macierzy \mathbf{Z} tak, aby miały jednakową długość, czyli:

$$y_{ij} = \frac{z_{ij}}{(\sum_j z_{ij}^2)^{1/2}}. \quad (14)$$

5. Podział każdego wiersza macierzy \mathbf{Y} , czyli punktu w przestrzeni \mathbb{R}^k , na k grup z zastosowaniem metody k -średnich (lub innej).
6. Przydzielenie każdej pierwotnej obserwacji x_i do k -tej grupy tylko wtedy, gdy i -ty wiersz macierzy \mathbf{Y} został przydzielony do k -tej grupy.

6. Wyniki badania

Do przeprowadzenia analizy empirycznej wykorzystano dane z bazy Eurostatu za 2019 r. dotyczące realizacji trzech wybranych Celów Zrównoważonego Rozwoju w krajach UE:

- celu 8: Wzrost gospodarczy i godna praca;
- celu 9: Innowacyjność, przemysł, infrastruktura;
- celu 12: Odpowiedzialna konsumpcja i produkcja¹.

¹ Pobrano dane z bazy <https://ec.europa.eu/eurostat/data/database>, zgromadzone w zakładce *Sustainable development indicators*, odnoszące się do celów 8, 9 i 12.

Jako metody grupowania wykorzystane w budowaniu macierzy zgodności przyjęto analizę spektralną i propagację podobieństwa. Każdej metody użyto 100 razy, za każdym razem na innym podzbiornym danych obejmującym 70% obserwacji z pierwotnego zbioru danych. Wszystkie obliczenia przeprowadzono w programie R z wykorzystaniem biblioteki `dicer` (Chiu i Talhouk, 2018).

Porównano wartość parametru k wskazywaną przez miarę stabilności PAC z wartościami wskazywanymi przez wybrane klasyczne indeksy, a mianowicie indeksy Calińskiego-Harabasa, Dunna i Daviesa-Bouldina. Liczba grup wahała się od 2 do 7. Wyniki analiz dla poszczególnych Celów Zrównoważonego Rozwoju zestawiono w tabl. 1–3.

Tabl. 1. Wartości miary stabilności PAC i klasycznych indeksów dla $k = 2, \dots, 7$ dla danych dotyczących celu 8: Wzrost gospodarczy i godna praca

k	Taksonomia spektralna				Metoda propagacji podobieństwa			
	PAC	Caliński-Harabasz	Dunn	Davies-Bouldin	PAC	Caliński-Harabasz	Dunn	Davies-Bouldin
2	0,463	47,936	0,113	0,463	0,071	47,936	0,113	0,548
3	0,503	28,070	0,041	0,516	0,434	69,258	0,058	0,521
4	0,632	39,903	0,045	0,451	0,429	64,865	0,055	0,563
5	0,561	30,528	0,031	0,500	0,291	128,101	0,119	0,474
6	0,463	23,781	0,042	0,458	0,265	84,247	0,055	0,512
7	0,429	67,738	0,074	0,492	0,188	227,133	0,177	0,386

Uwaga. Pogrubieniem wyróżniono optymalne wartości parametru k .

Źródło: obliczenia własne na podstawie danych Eurostatu za 2019 r.

Na podstawie wyników przedstawionych w tabl. 1 można stwierdzić, że w przypadku agregacji taksonomii spektralnej wartość parametru k wskazana przez miarę stabilności PAC jest zgodna jedynie z wartością wskazaną przez indeks Calińskiego-Harabasa. W przypadku agregacji metody propagacji podobieństwa między miarą stabilności PAC a pozostałymi indeksami nie ma żadnej zgodności co do wskazywanej wartości parametru k . Przyglądając się wynikom uzyskanym dzięki zastosowaniu metody propagacji podobieństwa, można zauważyć, że każda miara wskazała inną optymalną wartość parametru k .

W badaniu liczby skupień krajów UE pod względem celu 8 zaskakuje ogromny rozrzut wartości parametru k , czyli optymalnej liczby skupień, sugerowanych przez różne kryteria. Jedyną „niezapropionowaną” wartością parametru k w przypadku omawianego celu jest 6. Może to świadczyć np. o głębokim zróżnicowaniu krajów członkowskich pod względem realizacji celu dotyczącego wzrostu gospodarczego i zapewnienia godnej pracy. Trzeba zaznaczyć, że najczęściej sugerowaną liczbą skupień jest 7, więc może warto byłoby rozpatrzyć jeszcze wyższe wartości parametru k .

Tabl. 2. Wartości miary stabilności PAC i klasycznych indeksów dla $k = 2, \dots, 7$ dla danych dotyczących celu 9: Innowacyjność, przemysł, infrastruktura

k	Taksonomia spektralna				Metoda propagacji podobieństwa			
	PAC	Caliński-Harabasz	Dunn	Davies-Bouldin	PAC	Caliński-Harabasz	Dunn	Davies-Bouldin
2	0,717	33,593	0,350	0,846	0,071	47,936	0,113	0,548
3	0,839	25,374	0,262	0,998	0,434	69,258	0,058	0,521
4	0,606	27,059	0,286	0,933	0,429	64,865	0,055	0,563
5	0,466	24,669	0,303	0,850	0,291	128,101	0,119	0,474
6	0,497	21,340	0,233	0,956	0,265	84,247	0,055	0,512
7	0,474	18,792	0,263	1,062	0,188	227,133	0,177	0,386

Uwaga. Jak przy tabl. 1.

Źródło: obliczenia własne na podstawie danych Eurostatu za 2019 r.

Przedstawione w tabl. 2 wyniki dotyczące celu 9: Innowacyjność, przemysł, infrastruktura wskazują, że miara stabilności PAC nie jest zgodna z żadnym klasycznym indeksem w kwestii wyboru optymalnej wartości parametru k – i to zarówno w przypadku gdy pojedyncze grupowania tworzone z zastosowaniem analizy spektralnej, jak i metody propagacji podobieństwa. Jedyne zgodności, jakie można dostrzec, zachodzą między indeksem Calińskiego-Harabasza i indeksem Dunna, jednak wartości te były różne dla grupowania zagregowanego z wykorzystaniem taksonomii spektralnej ($k = 2$) i metody propagacji podobieństwa ($k = 7$).

Analizując wszystkie wyniki przedstawione w tabl. 2, należy zwrócić uwagę, że najczęściej wskazywane wartości parametru k w przypadku celu 9 to 2 i 7, czyli krańcowe wartości założonego przedziału zmienności wartości parametru k . Świadczy to albo o małym zróżnicowaniu krajów UE pod względem realizacji celu 9 (dla $k = 2$), albo wręcz przeciwnie – o dużych różnicach między nimi, jeśli chodzi o realizację omawianego celu (dla $k = 7$).

Tabl. 3. Wartości miary stabilności PAC i klasycznych indeksów dla $k = 2, \dots, 7$ dla danych dotyczących celu 12: Odpowiedzialna konsumpcja i produkcja

k	Taksonomia spektralna				Metoda propagacji podobieństwa			
	PAC	Caliński-Harabasz	Dunn	Davies-Bouldin	PAC	Caliński-Harabasz	Dunn	Davies-Bouldin
2	0,817	11,641	0,278	0,937	0,677	17,896	0,116	1,080
3	0,765	21,280	0,166	0,870	0,519	21,280	0,166	0,870
4	0,730	18,751	0,166	0,947	0,463	19,163	0,165	0,955
5	0,685	17,433	0,165	0,886	0,426	20,940	0,206	0,738
6	0,566	14,519	0,165	0,949	0,370	22,363	0,257	0,760
7	0,484	12,691	0,165	0,900	0,320	9,654	0,115	1,088

Uwaga. Jak przy tabl. 1.

Źródło: obliczenia własne na podstawie danych Eurostatu za 2019 r.

Na podstawie danych przedstawionych w tabl. 3 można zaobserwować, że w przypadku przyjęcia taksonomii spektralnej jako metody bazowej miara stabilności PAC nie wykazuje zgodności co do optymalnej wartości parametru k z żadnym z zastosowanych indeksów. Dla taksonomii zagregowanej, wykorzystującej metodę propagacji podobieństwa, miara stabilności PAC, podobnie jak indeks Daviesa-Bouldina, jako optymalną liczbę grup wykazuje $k = 7$. Warto również odnotować zgodność indeksów Calińskiego-Harabasz i Dunna, które sugerują z kolei wartość $k = 6$.

Analizując wszystkie wyniki przedstawione w tabl. 3, należy zwrócić uwagę, że dwie najczęściej wskazywane liczby skupień to 6 i 7, co może świadczyć o wyraźnym zróżnicowaniu krajów UE pod względem realizacji celu 12.

Na podstawie badania przeprowadzonego dla taksonomii zagregowanej można stwierdzić, że wybór zarówno kryterium ustalania optymalnej liczby grup (miara stabilności PAC lub klasyczny indeks), jak i metody bazowej w tym podejściu (propagacja podobieństwa lub taksonomia spektralna) wpływa na to, jaka wartość parametru k zostanie wskazana.

Podsumowując, można sformułować następujące wnioski:

- cel 8 – proponowana wartość parametru k charakteryzuje się największą zmiennością; obejmuje wartości 2–5 i 7; tak duża rozbieżność może świadczyć o głębokim zróżnicowaniu krajów członkowskich w zakresie realizacji tego celu;
- cel 9 – proponowane są tylko skrajne wartości parametru k (2 lub 7), co może świadczyć albo o niewielkim zróżnicowaniu krajów członkowskich pod względem realizacji tego celu (dla $k = 2$), albo wręcz przeciwnie – o dużym zróżnicowaniu (dla $k = 7$);
- cel 12 – najczęściej wskazywanymi wartościami parametru k są 7 i 8, co może świadczyć o wyraźnych różnicach w osiąganiu tego celu przez kraje członkowskie.

7. Podsumowanie

W artykule przedstawiono proporcję niejednoznacznie grupowanych par obiektów PAC jako miarę stabilności zaproponowaną przez Şenbabaoglu i in. (2014). To kolejna koncepcja miary stabilności, która mimo że odnosi się do podejścia zagregowanego w taksonomii, to opiera się na przekonaniu, że wartość parametru k wskazywana przez miarę stabilności powinna oznaczać rzeczywistą strukturę grup.

Wyniki badania przeprowadzonego na zbiorach danych za 2019 r., opisujących trzy wybrane Cele Zrównoważonego Rozwoju w krajach UE, pozwalają przede wszystkim stwierdzić, że wskazywana przez klasyczne indeksy wartość parametru k niemalże w każdym przypadku całkowicie różni się z wartościami wskazywanymi przez miarę stabilności PAC. Pierwszym wyjątkiem – dla celu 8 – jest taksonomia zagregowana bazująca na taksonomii spektralnej, gdzie zarówno miara stabilności PAC, jak i indeks Calińskiego-Harabasz jako optymalną wskazały taką samą liczbę

grup, a mianowicie 7. Drugim wyjątkiem – zaobserwowanym wtedy, gdy w taksonomii zagregowanej jako bazową wybierze się metodę propagacji podobieństwa – jest cel 12, w przypadku którego miara stabilności PAC i indeks Daviesa-Bouldina jako optymalną wskazały tę samą wartość parametru k .

Z przeprowadzonych badań wynika również, że klasyczne indeksy różnią się między sobą w zakresie wskazywania optymalnej liczby grup. Tego typu rozbieżności można zaobserwować w ramach tej samej metody wybranej jako bazowa do utworzenia macierzy zgodności, ale także porównując różne metody służące jako bazowe do utworzenia taksonomii zagregowanej.

Warto podkreślić, że wyniki omówione w artykule są zbieżne z wynikami badań przedstawionymi we wcześniejszych pracach autorki (Rozmus, 2021, 2022), w których również wykazano, że stosowanie różnych miar stabilności i klasycznych indeksów w odniesieniu do taksonomii zagregowanej, jak również różnych metod bazowych do budowy taksonomii zagregowanej nie daje zbieżnych wyników w kontekście wyboru wartości parametru k . Uzyskane rezultaty pokazują dodatkowo, że nawet opracowanie specjalnej miary stabilności w odniesieniu do podejścia zagregowanego w taksonomii nie rozwiązuje problemu niejednoznacznego wskazywania liczby grup, na które należy podzielić zbiór obiektów.

Bibliografia

- Ben-Hur, A., Guyon, I. (2003). Detecting stable clusters using principal component analysis. W: M. J. Brownstein, A. B. Kohodursky (red.), *Functional Genomics: Methods and Protocols* (s. 159–182). Humana press. <https://doi.org/10.1385/1-59259-364-X:159>.
- Bodenhofer, U., Kothmeier, A., Hochreiter, S. (2011). APCluster: an R package for affinity propagation clustering. *Bioinformatics*, 27(17), 2463–2464. <https://doi.org/10.1093/bioinformatics/btr406>.
- Brock, G., Pihur, V., Datta, S., Datta, S. (2008). cValid: An R Package for Cluster Validation. *Journal of Statistical Software*, 25(4), 1–22. <https://doi.org/10.18637/jss.v025.i04>.
- Chiu, D. S., Talhouk, A. (2018). diceR: an R package for class discovery using an ensemble driven approach. *BMC Bioinformatics*, 19(11), 1–4. <https://doi.org/10.1186/s12859-017-1996-y>.
- Dudoit, S., Fridlyand, J. (2003). Bagging to improve the accuracy of a clustering procedure. *Bioinformatics*, 19(9), 1090–1099. <https://doi.org/10.1093/bioinformatics/btg038>.
- Fang, Y., Wang, J. (2012). Selection of the number of clusters via the bootstrap method. *Computational Statistics and Data Analysis*, 56(3), 468–477. <https://doi.org/10.1016/j.csda.2011.09.003>.
- Fred, A. L. N., Jain, A. K. (2002). Data clustering using evidence accumulation. W: *2002 International Conference on Pattern Recognition* (s. 276–280). IEEE. <https://doi.org/10.1109/ICPR.2002.1047450>.
- Fred, A. L. N., Jain, A. K. (2005). Combining multiple clusterings using evidence accumulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6), 835–850. <https://doi.org/10.1109/TPAMI.2005.113>.

- Frey, B. J., Dueck, D. (2007). Clustering by Passing Messages Between Data Points. *Science*, 315(5814), 972–976. <https://doi.org/10.1126/science.1136800>.
- Henning, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics and Data Analysis*, 52(1), 258–271. <https://doi.org/10.1016/j.csda.2006.11.025>.
- Hornik, K. (2005). A CLUE for CLUster Ensembles. *Journal of Statistical Software*, 14(12), 1–25. <https://doi.org/10.18637/jss.v014.i12>.
- Kannan, R., Vempala, S., Vetta, A. (2004). On clustering: Good, Bad and Spectral. *Journal of the ACM*, 51(3), 497–515. <https://doi.org/10.1145/990308.990313>.
- Kuncheva, L. I., Vetrov, D. P. (2006). Evaluation of Stability of k -Means Cluster Ensembles with Respect to Random Initialization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11), 1798–1808. <https://doi.org/10.1109/TPAMI.2006.226>.
- Leisch, F. (1999). *Bagged Clustering* (SFB Working Papers No. 51). <https://doi.org/10.57938/9b129f95-b53b-44ce-a129-5b7a1168d832>.
- Leone, M., Sumedha, Weigt, M. (2007). Clustering by soft-constraint affinity propagation: applications to gene-expression data. *Bioinformatics*, 23(20), 2708–2715. <https://doi.org/10.1093/bioinformatics/btm414>.
- Lord, E., Willems, M., Lapointe, F. J., Makarenkov, V. (2017). Using the stability of objects to determine the number of clusters in datasets. *Information Sciences*, 393, 29–46. <https://doi.org/10.1016/j.ins.2017.02.010>.
- Marino, V., Presti, L. L. (2019). Stay in touch! New insights into end-user attitudes towards engagement platforms. *Journal of Consumer Marketing*, 36(6), 772–783. <https://doi.org/10.1108/JCM-05-2018-2692>.
- Meng, J., Hao, H., Luan, Y. (2016). Classifier ensemble selection based on affinity propagation clustering. *Journal of Biomedical Informatics*, 60, 234–242. <https://doi.org/10.1016/j.jbi.2016.02.010>.
- Monti, S., Tamayo, P., Mesirov, J., Golub, T. (2003). Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Machine Learning*, 52(1–2), 91–118. <https://doi.org/10.1023/A:1023949509487>.
- Ng, A. Y., Jordan, M. I., Weiss, Y. (2001). On Spectral Clustering: Analysis and an algorithm. W: T. G. Dietterich, S. Becker, Z. Ghahramani (red.), *Advances in Neural Information Processing Systems 14*. The MIT Press.
- Rozmus, D. (2011). Porównanie stabilności zagregowanych algorytmów taksonomicznych opartych na macierzy współwystąpień. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu. Research Papers of Wrocław University of Economics*, (176), 212–220.
- Rozmus, D. (2013). Porównanie dokładności taksonomicznej metody propagacji podobieństwa oraz zagregowanych algorytmów taksonomicznych opartych na idei metody bagging. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu. Research Papers of Wrocław University of Economics*, (279), 106–114.
- Rozmus, D. (2021). The Number of Groups in an Aggregated Approach in Taxonomy with the Use of Stability Measures and Classical Indices – A Comparative Analysis. *Acta Universitatis Lodzianis. Folia Oeconomica*, 6(357), 55–67. <https://doi.org/10.18778/0208-6018.357.04>.
- Rozmus, D. (2022). Cluster Ensemble Stability in Clustering of EU Members in Terms of Sustainable Development Goals. W: K. Jajuga, G. Dehnel, M. Walesiak (red.), *Modern Classification and Data Analysis. Methodology and Applications to Micro- and Macroeconomic Problems* (s. 289–301). Springer. https://doi.org/10.1007/978-3-031-10190-8_20.

- Şenbabaoğlu, Y., Michailidis, G., Li, J. Z. (2014). Critical limitations of consensus clustering in class discovery. *Scientific Reports*, 4, 1–13. <https://doi.org/10.1038/srep06207>.
- Shamir, O., Tishby, N. (2008). Cluster stability for finite samples. W: J. C. Platt, D. Koller, Y. Singer, S. T. Roweis (red.), *Advances in Neural Information Processing Systems 20 (NIPS 2007)* (s. 1297–1304). Curran Associates.
- Shi, J., Malik, J. (2000). Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 888–905. <https://doi.org/10.1109/34.868688>.
- Suzuki, R., Shimodaira, H. (2006). Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, 22(12), 1540–1542. <https://doi.org/10.1093/bioinformatics/btl117>.
- Volkovich, Z., Barzily, Z., Toledano-Kitai, D., Avros, R. (2010). The Hotteling's metric as a cluster stability measure. *Computer Modelling and New Technologies*, 14(4), 65–72. http://www.cmnt.lv/upload-files/ns_3914_4_cmnt2010.pdf.
- Yu, Z., Li, L., Liu, J., Zhang, J., Han, G. (2015). Adaptive noise immune cluster ensemble using affinity propagation. *IEEE Transactions on Knowledge and Data Engineering*, 27(12), 3176–3189. <https://doi.org/10.1109/TKDE.2015.2453162>.

Why do young workers quit their first job? Identification of the risk factors using the Cox model and survival trees¹

Wioletta Grzenda,^a Agnieszka Marszałek^b

Abstract. According to Statistics Poland's data, the situation of young people in the Polish labour market has improved significantly in recent years. Therefore, on the one hand, it is easier for young people entering the labour market to find a job, and on the other, it is increasingly difficult for employers to keep such people in their organisation. The aim of this study is to identify and assess individual characteristics of young workers and work-related factors that affect the length of the time they spend in their first job. The study is based on data for 2019 and 2020 from Statistics Poland's Labour Force Survey. It is of key importance in the research on the professional activity of young people to take into account in modelling the high volatility of their characteristics over time. Therefore, we used the Cox model with time-variant variables to identify factors of risk of quitting a young employee's first job. One of the findings of the study was that people with higher education were more likely to quit their jobs than people with lower-level education. As regards work-related factors, in addition to the type of employment contract, the weekly working time and holding or not a managerial position were the important ones affecting the decision to continue or quit. Furthermore, groups of employees homogeneous in terms of the duration of their first job were identified using survival trees. We found that employees with fixed-term contracts were less likely to quit their jobs than those with permanent contracts, but working part-time.

Keywords: labour market, first job, young employees, Cox model, survival trees

JEL: J62, J64, C14, C41

Dlaczego młodzi pracownicy porzucają pierwszą pracę? Identyfikacja czynników ryzyka z wykorzystaniem modelu Coxa i drzew przetrwania

Streszczenie. Według danych Głównego Urzędu Statystycznego w ostatnich latach w Polsce sytuacja młodych osób na rynku pracy znacznie się poprawiła. W konsekwencji osobom wchodzącym na rynek pracy łatwiej jest znaleźć zatrudnienie, a pracodawcom coraz trudniej zatrzymać takie osoby w organizacji. Celem badania omawianego w artykule jest identyfikacja

¹ Artykuł został opracowany na podstawie referatu wygłoszonego na 40th International Conference MSA 2022 joined with MASEP, która odbyła się w dniach 7–9 listopada 2022 r. w Łodzi / The article is based on a paper delivered at the 40th International Conference MSA 2022 joined with MASEP, held on 7–9th November 2022 in Łódź.

^a Szkoła Główna Handlowa w Warszawie, Kolegium Analiz Ekonomicznych, Instytut Statystyki i Demografii, Polska / SGH Warsaw School of Economics, Collegium of Economic Analysis, Institute of Statistics and Demography, Poland. ORCID: <https://orcid.org/0000-0002-2226-4563>. Autor korespondencyjny / Corresponding author, e-mail: wgrzend@sgh.waw.pl.

^b Badaczka niezależna / Independent researcher. ORCID: <https://orcid.org/0000-0003-4906-6484>.
E-mail: marszalek.agnieszka@outlook.com.

i ocena indywidualnych charakterystyk młodych pracowników oraz czynników związanych z wykonywaną pracą, które mają wpływ na czas trwania pierwszego zatrudnienia. W badaniu wykorzystano dane za lata 2019 i 2020 pochodzące z Badania Aktywności Ekonomicznej Ludności przeprowadzonego w Polsce przez Główny Urząd Statystyczny. Kluczowym elementem badań aktywności zawodowej młodych osób na rynku pracy powinno być uwzględnienie w modelowaniu dużej zmienności w czasie ich indywidualnych charakterystyk, dlatego do ustalenia czynników ryzyka porzucenia pierwszej pracy wykorzystano model Coxa ze zmiennymi zależnymi od czasu. Stwierdzono m.in., że większe ryzyko zaprzestania wykonywania pracy istniało w przypadku osób z wyższym wykształceniem niż wśród osób z niższym wykształceniem. Oprócz rodzaju umowy o pracę duży wpływ na to ryzyko miały też tygodniowy czas pracy i pełnienie funkcji kierowniczych. Ponadto wśród respondentów wyodrębniono, z wykorzystaniem drzew przetrwania, grupy homogeniczne ze względu na czas trwania pierwszego zatrudnienia. Okazało się, że istnieje mniejsze prawdopodobieństwo rozwiązania stosunku pracy przez pracowników zatrudnionych na czas określony niż przez pracowników zatrudnionych na czas nieokreślony, ale pracujących w niepełnym wymiarze godzin.

Słowa kluczowe: rynek pracy, pierwsza praca, młodzi pracownicy, model Coxa, drzewa przeżycia

1. Introduction

The situation of young people in the labour market in Poland has improved significantly in recent years. According to Eurostat data, the unemployment rate among people aged 15 to 29 in 2020 was 7.1% – one of the lowest in Europe (Eurostat, 2022). The year 2020 marked the outbreak of the COVID-19 pandemic in Poland and several other countries, which significantly impacted the global economy, including the job market. In 2020, the unemployment rate among Poles aged 15 to 29 increased by 0.5% compared to 2019; however, it was 0.6% lower than in 2018 and significantly lower than in previous years. This made it increasingly difficult for employers to retain young workers in their companies or institutions. Too-high a rate of voluntary employee turnover is undesirable, as it entails high costs for the organisation (Steenackers & Guerry, 2016) and, consequently, leads to an even higher turnover (Boon & Biron, 2016). Among other things, hiring a new employee requires appropriate training, even if they have all the required professional competencies. It is therefore desirable not only to define a suitable applicant profile in terms of the position to be filled, but also to assess the risk of the applicant quitting his or her job. According to Domurath et al. (2023), the desire to change jobs grows stronger over time, and this effect is particularly strong for employees working in low-growth ventures. Moreover, people under the age of 30 often change their jobs as they are still figuring out what they want to do in life (Arnett, 2006).

Currently, people born in the 1980s and 1990s, i.e., Generation Y and older representatives of Generation Z, are considered to be young workers in the labour market. According to Gadomska-Lila (2015), these people have significantly different work attitudes than the older generations. Hence, the factors that drive

them to change employers may be different from those motivating people from older generations. According to Horáčková and Kopáček (2018), the different attitude of young people towards work aggravates the workforce management problem and makes these workers less retainable. Considering the findings of Arnett (1998, 2006), in this study we will focus on the first job of people aged 18 to 29. The data for the study came from the Polish Labour Force Surveys (LFS) held in 2019–2020. The study includes, among other respondents, those who were employed for the first time or had just terminated their employment at their first employer at the time of the survey. Both the individual characteristics of young workers and factors directly related to their jobs were included in the analysis.

The aim of this study is to identify and assess individual characteristics of young workers and work-related factors that affect the length of the time they spend in their first job.

Therefore, our first research question is: which characteristics of young employees influence the risk that they would quit their first job? Here we have to remember, though, that determining the impact of individual employee characteristics on this risk might not be enough. The decision to remain in particular employment may also be influenced by the conditions of this employment and the characteristics of the enterprise. Therefore, the second question we posed is: which work-related factors influence the risk that young workers would quit their first job? In addition to identifying these factors, it is important to examine their importance and check under what conditions young people want to continue working in the same enterprise. Hence the third research question: what is the profile of a young employee who is not very likely to leave his or her employment? Finally, the fourth research question: which work-related factors can keep young workers in an enterprise for longer, and which push them away?

2. The first job and switching jobs

According to Ignaczak's research (2014) based on the Canadian General Social Survey, successive cohorts of workers tend to spend less and less time in their first job. This trend was observed for both women and men, but more significantly for men. In our study, we focus on generations Y and Z. Generation Y is very familiar with modern means of communication and the virtual world. These are the people who live at a fast pace and focus on themselves and their own development (Robak, 2017). In the literature, these individuals have been described as having little loyalty towards their employer, looking for a job that they will enjoy and, at the same time, one that will guarantee them financial independence (Wawrzonek, 2014). Consequently, people from this generation are prone to change employers frequently in

search of a satisfactory job, and their employers face the problem of qualified staff retention and high turnover of employees (Hassan et al., 2023). Generation Z has a high level of self-confidence and currently is the most technologically-competent generation of all. Individuals from Generation Z are often highly educated and have competencies that are now in strong demand in the labour market – they are not afraid of risk and make decisions quickly. In addition, they are geared towards fast-track professional careers and at the same time want their working and private lives to be a coherent whole (Żarczyńska-Dobiesz & Chomątowska, 2014).

A study by McGinnis Johnson and Ng (2016) found that there is a strong correlation between the education level of people from Generation Y and their employer-switching behaviours. While examining the duration of the first job according to the educational background on the basis of data from the Canadian General Social Survey, Ignaczak et al. (2022) indicated significant differences between genders. They observed that for women, in contrast to men, a higher level of education reduced the hazard of changing jobs. For men, entering the labour market at a young age and working part-time resulted in a higher likelihood of quitting their first job. Immigrant and married men, on the other hand, were less likely to quit their jobs. As regards women, the probability of leaving their first job increased in the case of part-time female employees and mothers. Similar results were obtained by Steenackers and Guerry (2016) when analysing the labour market behaviour of young people in Belgium. They found that young women were more likely than young men to change employers in the early stages of their careers. This reflects their search for a rewarding and stable job that would allow them to take care of their children after starting a family (Grzenda, 2019; Steenackers & Guerry, 2016).

Young people at the beginning of their professional careers are more likely to change employers than their older colleagues (Steenackers & Guerry, 2016). There is a particularly high probability of employee departure in the case of entry-level jobs. It is rare for an employee to start and end their career with the same employer (Pandey, 2019). Changing employers is exceptionally common in the private sector and among Generation Y workers (Steenackers & Guerry, 2016). From the company's point of view, excessive employee turnover is highly undesirable and employers use various strategies to deal with this problem. Understanding, at the individual level, factors influencing the behaviour of young people in the labour market can significantly improve employers' effectiveness of retaining these people in the company and thus makes it more viable to fully use the potential of young employees (Das & Baruah, 2013).

Young people do not always change jobs voluntarily, but anyway, the experience they gain in their first job enables them to perform much better in the labour market

afterwards. According to a study by Cockx and Picchio (2012), even short-lived jobs for graduates, lasting for a quarter of a year or shorter, increase the likelihood of finding long-term employment later on. Similar results were obtained for young people taking up jobs while still in education (Baert et al., 2022).

3. Research method

Our study used data from Statistics Poland's LFS to model the time spent in the first job. It is a quarterly panel survey with rotational sampling conducted by Statistics Poland and covering persons aged 15 and older (Główny Urząd Statystyczny [GUS], 2018). The LFS is based on a representative sample of Polish households. In the first quarter of 2019, the sample consisted of 69,526 people (GUS, 2019). In accordance with the study's main precepts, a randomly selected respondent is surveyed for two consecutive quarters, then does not participate in the study for the next two quarters, and is surveyed again in the following two. The LFS contains retrospective questions that make it possible to analyse the length of the employment. Our analysis covers respondents aged 18–29 who, at the time of the survey, were employed for the first time or had just quitted their first job. Only people who started working at the age of 18 or above were considered in our research. The assumed upper age limit of respondents was determined on the basis of Arnett's research (1998, 2006), which sets the full social maturity of a person around the age of 30. The number of observations performed for the purpose of the analysis was 384. Of the surveyed people, at the time of the last survey, only 13 (3.39%) were still in their first job.

In the case of young people, it is very common for their characteristics such as marital status, education, or place of residence to change over a short period. Factoring in the changes in the values of explanatory variables requires the use of appropriate data analysis methods. Hence, in this study, the time-dependent Cox model (Blanche et al., 2022; Fisher & Lin, 1999; Kalbfleisch & Prentice, 2002) and survival trees (Bou-Hamad et al., 2011; De Rose & Pallara, 1997; Shimokawa et al., 2015) were used to analyse the time spent by young workers in their first job. The risk determinants relating to people aged 18 to 29 quitting their first job were identified by means of the generalisation of the Cox model containing both time-invariant and time-variant variables (Hosmer et al., 2008). In the second part of the analysis, groups of employees homogeneous in terms of the duration of their first job were identified using survival trees. The proposed approach and the use of up-to-date data provide a new insight into the behaviour of young people on the labour market in Poland. The obtained results can be used, for instance, to define the profile of a young employee who is highly probable to quit his or her job.

The dependent variable, which in the survival analysis is time to failure, was defined as the number of months from the beginning of the first job to its termination for those who quit work during the observation period, or until the last survey for those who were still working for the same company or institution at the time of the last survey. The minimum time spent in a job was one month, and the maximum 123 months. In addition, a censoring variable was created and assigned the value of unity if the event occurred during the observation period. Thus, the value was one if the respondent terminated employment with the first employer, and zero if he or she did not.

Due to the very small representation of one of the two following categories, the information about the type of institution where the respondent worked (private or public) and whether the person was self-employed or not (less than 1%) were not included in the analysis. Detailed information on the overtime, salary, having or not an additional job and participating or not in training was also skipped, as less than 5% of respondents provided this information. The age variable was created by categorising the continuous variable describing the age of the respondent. The first category reflects the age range into which young people are classified, while the second group corresponds to those in the ‘emerging adulthood’ stage (Arnett, 2006). The set of potential independent variables selected for modelling is presented in Table 1.

Table 1. Individual characteristics and work-related factors

Variable	Categories	Number	Percent
Individual characteristics			
Gender	man	226	58.85
	woman	158	41.15
Macroregion of Poland (NUTS 1)	South (Małopolskie, Śląskie)	53	13.80
	Northwest (Wielkopolskie, Zachodniopomorskie, Lubuskie)	63	16.41
	Southwest (Dolnośląskie, Opolskie)	31	8.07
	North (Kujawsko-Pomorskie, Warmińsko-Mazurskie, Pomorskie)	63	16.41
	Central (Łódzkie, Świętokrzyskie)	45	11.72
	East (Lubelskie, Podkarpackie, Podlaskie)	80	20.83
	Mazovia Voivodship (Warszawski Stołeczny, Mazowiecki Regionalny)	49	12.76
Marital status	single	328	85.42
	married	56	14.58
Age	18–25 years	214	55.73
	26–29 years	170	44.27
Finished process of education	no	64	16.67
	yes	320	83.33

Table 1. Individual characteristics and work-related factors (cont.)

Variable	Categories	Number	Percent
Individual characteristics (cont.)			
Education	tertiary	150	39.06
	post-secondary and vocational secondary	94	24.48
	general secondary	54	14.06
	basic vocational	35	9.11
	lower secondary, primary, incomplete primary	51	13.28
Place of residence	city with 100,000 or more residents	130	33.85
	city with less than 100,000 residents	84	21.88
	rural areas	170	44.27
Work-related factors			
Weekly working time	part-time (less than 40 hours per week)	33	8.59
	full-time (40 hours per week)	316	82.29
	overtime (more than 40 hours per week)	35	9.12
Type of employment contract	permanent	230	59.90
	fixed-term	154	40.10
Shift work	no	260	67.71
	yes	124	32.29
Managerial position	no	345	89.84
	yes	39	10.16
Earnings	up to PLN 2,500	77	20.05
	over PLN 2,500	55	14.32
	no data available	252	65.63
Employer's NACE (classification of activities)	wholesale and retail trade; repair of motor vehicles including motorcycles; transportation and storage	97	25.26
	manufacturing; electricity, gas, steam, hot water and air conditioning manufacturing and supply; water supply; sewerage, waste management and remediation activities; construction	128	33.33
	human health and social work activities	22	5.73
	education; public administration and defence; compulsory social security	27	7.03
	others	110	28.65
Workplace in the same municipality as the place of residence	no	141	36.72
	yes	243	63.28
Employer size	up to 10 people	97	25.26
	10 people and more	287	74.74
Full-time job	no	26	6.77
	yes	358	93.23
Work consistent with education level	yes	165	42.97
	partially	64	16.67
	no	155	40.36
Individual professional development plan	no	38	9.90
	yes	288	75.00
	others	58	15.10

Source: authors' work based on Statistics Poland's Labour Force Survey.

In this study, to identify factors affecting the probability of quitting the first job, we used the generalisation-of-the-proportional-hazards model containing both time-invariant and time-variant variables (Hosmer et al., 2008). Let $[x_1, \dots, x_k]^T$ denote the vector of independent variables, and $[\beta_1, \dots, \beta_k]$ denote the vector of regression coefficients, due to which the hazard rate for the model of proportional hazards takes the form:

$$h(t|x_1, \dots, x_k) = h_0(t) \exp\left(\sum_{i=1}^k \beta_i x_i\right), \tag{1}$$

where $h_0(t)$ denotes the baseline hazard. The baseline hazard can be regarded as the hazard function for an individual whose all covariates have the values of zero. It is a non-negative function of a continuous random variable T , for $t \geq 0$ (Cox, 1975; Cox & Oakes, 1984). Due to the nonparametric specification of the hazard function, the partial likelihood method is used to estimate the covariate effect in this model.

The Cox model is based on the assumption of the proportionality of hazards. In this study, the assumption was verified using the time-dependent covariate method. This method involves including time-dependent variables in a model with the set of potential explanatory variables by creating interaction between the independent variables and a non-zero function of the time variable $g_i(t)$. The most popular functions are $g_i(t) = t$ and $g_i(t) = \ln(t)$. We choose the second function similarly as Ng'andu (1997). The model then takes the following form:

$$h(t|x_1, \dots, x_k) = h_0(t) \exp\left\{\sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \gamma_i x_i g_i(t)\right\}, \tag{2}$$

where γ_i denotes the parameter of interaction for the i -th ($i = 1, \dots, k$) explanatory variable. Hazards are assumed to be proportional at times t , if $\gamma_i = 0$, for $i = 1, \dots, k$. This assumption is verified using the likelihood ratio test.

When the assumption of proportional hazards is not met, it is possible to construct a generalised Cox model with nonproportional hazards, called a time-dependent Cox model (Blanche et al., 2022; Fisher & Lin, 1999; Kalbfleisch & Prentice, 2002). We used the generalisation-of-the-proportional-hazards model containing both time-invariant and time-variant variables (Hosmer et al., 2008). The hazard rate for this model is defined as:

$$h(t|x_1, \dots, x_{k_1}, x_{1t}, \dots, x_{k_2t}) = h_0(t) \exp\left\{\sum_{h=1}^{k_1} \beta_h x_h + \sum_{l=1}^{k_2} \gamma_l x_{lt}\right\}, \tag{3}$$

where x_h ($h = 1, \dots, k_1$) denotes time-invariant variables, and x_{lt} ($l = 1, \dots, k_2$) stands for time-variant variables.

In the second stage of the research, the survival tree was used to find groups of employees homogeneous in terms of the time spent in their first job. Survival trees are an alternative to classical survival analysis methods (De Rose & Pallara, 1997). These techniques, also known as recursive partitioning methods, involve the construction of prediction rules in the form of a binary tree. The advantage of this method over classical methods is its structure, which requires fewer assumptions and allows relatively simple analysis of the non-standard and non-linear structure of a dataset. Using this model provides easily-interpretable decision rules which reflect the behavioural patterns of individuals during the surveyed period. For each group formed on the basis of a behavioural pattern, it is possible to calculate a survival function and a hazard ratio (De Rose & Pallara, 1997).

Basic recursive partitioning involves splitting data according to the value of one explanatory variable in a single step. In the case of the continuous or ordinal variable X , the potential splitting rule takes the form $X \leq c$, where c is a constant. For the categorical variable X , the splitting rule takes the form $X \in \{c_1, \dots, c_k\}$, where c_1, \dots, c_k are the possible values of variable X . The division starts with the root node, which contains all the observations. In this step, all possible binary splitting rules are searched for and verified using a predetermined partitioning criterion. For survival trees, the log-rank test (Klein & Moeschberger, 2006) is commonly used as a partitioning criterion. This criterion assures the best separation of the median survival times over two children of nodes. The best split is the one for which the largest significant test statistic value was obtained (Bou-Hamad et al., 2011).

4. Factors contributing to quitting the first job

In the first stage of the research, we checked which of the demographic and socio-economic characteristics of the respondents and the work-related factors change over time. It turned out that only two characteristics are constant over time: 'gender' and 'macroregion of Poland'. Thus, we used the hazards model containing both time-invariant and time-variant variables (3). For the remaining explanatory variables, the examined period was divided into sub-episodes in which the values of the considered variables were constant for each respondent. Before starting to construct the Cox model, the assumption of hazard proportionality was verified. This assumption was not met for the 'age' variable. Therefore, a time-dependent indicator variable was created for this factor. This variable was constructed using a non-parametric hazard function, with the 30th month of employment chosen as the split point. This point was chosen taking into account the shape of the hazard function. In this way, two variables were created: the first one corresponding to the time interval from the first month to the 30th month of employment, and the second

corresponding to the time interval from the 31st month of employment to the end of the study. Finally, only those factors that had a statistically significant impact on the risk of quitting the first job were included in the model. The results obtained from the generalised Cox model are presented in Tables 2 and 3.

Table 2. Hazard of terminating employment: the results of the Cox model (Type 3 Tests)

Covariate	DF	Wald Chi-Square	p-value
Marital status	1	4.3791	0.0364
Age: up to the 30th month of work	1	36.7849	<0.0001
after 30 months of work	1	56.0383	<0.0001
Education level	4	42.0982	<0.0001
Weekly working time	2	13.0962	0.0014
Type of employment contract	1	89.5275	<0.0001
Managerial position	1	9.2081	0.0024
Employer's NACE	4	21.3222	0.0003
Individual professional development plan	2	4.3302	0.1147

Source: authors' work based on Statistics Poland's Labour Force Survey.

Table 3. Hazard of terminating employment: the results of the Cox model

Covariate	Parameter estimate	Standard error	p-value	Hazard ratio
Age (ref. 18–25 years)				
26–29 years: up to the 30th month of work	-1.163	0.192	<0.0001	0.312
after 30 months of work	-1.383	0.185	<0.0001	0.251
Marital status (ref. single)				
Married	-0.360	0.172	0.036	0.698
Education level (ref. lower secondary, primary, incomplete primary)				
Higher	0.613	0.202	0.002	1.846
Post-secondary and vocational secondary ...	-0.133	0.179	0.458	0.876
General secondary	-0.102	0.224	0.651	0.903
Basic vocational	-0.607	0.219	0.006	0.545
Weekly working time (ref. full-time)				
Part-time	-0.324	0.203	0.111	0.723
Overtime	-0.701	0.205	0.001	0.496
Type of employment contract (ref. permanent)				
Fixed-term	1.266	0.134	<0.0001	3.545
Managerial position (ref. no)				
Yes	-0.560	0.184	0.002	0.572
Individual professional development plan (ref. no)				
Yes	-0.108	0.197	0.584	0.898
Others	-0.378	0.184	0.040	0.685

Table 3. Hazard of terminating employment: the results of the Cox model (cont.)

Covariate	Parameter estimate	Standard error	p-value	Hazard ratio
Employer’s NACE (ref. wholesale and retail trade; repair of motor vehicles including motorcycles; transportation and storage)				
Manufacturing; electricity, gas, steam, hot water and air conditioning, manufacturing and supply; water supply; sewerage, waste management and remediation activities; construction	-0.610	0.142	<.0001	0.544
Human health and social work activities	-0.600	0.280	0.832	0.942
Education; public administration and defence; compulsory social security	0.035	0.287	0.901	1.036
Others	-0.184	0.156	0.239	0.832

Source: authors’ work based on Statistics Poland’s Labour Force Survey.

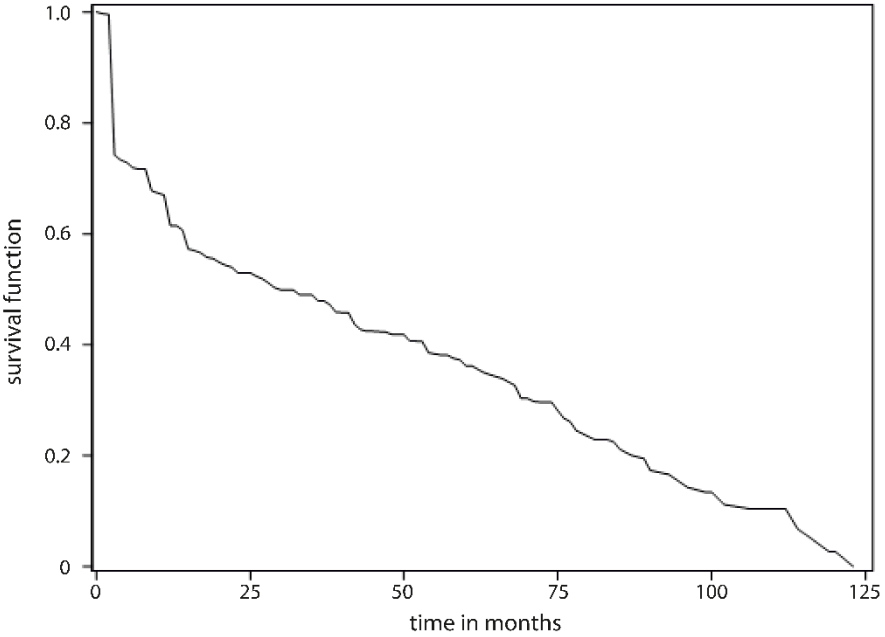
The estimation of the Cox model with time-dependent variables revealed that the risk of leaving the first job decreased with age. People aged 26 to 29 were 68.8% less likely to terminate their first employment contracts in the first 30 months, and 74.9% less likely to do so when the employment relationship lasted longer than 30 months than people aged 18 to 25. Married people were 30.2% less likely to quit their jobs than the unmarried. Furthermore, statistically significant parameter estimates were obtained for some levels of the respondent’s education variable. People with higher education were 84.6% more likely, and those with basic vocational education were 45.5% less likely to terminate their job than the least educated people.

The type of employment contract is an especially important work-related factor. Respondents with a fixed-term employment contract were approximately 3.5 times more likely to quit their job than people with a permanent contract. Another important work-related factor was weekly working time. People who reported working overtime were 50.4% less likely to leave their employer than those working full-time (40 hours per week). In addition, people in managerial positions were 42.8% less likely to quit than those in other positions. Having an individual professional development plan proved statistically insignificant. Young people working in industries such as manufacturing, electricity, gas, steam and air conditioning supply, water supply, sewerage and waste management, and remediation activities and construction were 45.5% less likely to terminate their employment contracts than respondents working in industries such as wholesale and retail trade, repair of motor vehicles including motorcycles, and transportation and storage. Differences in relation to the reference category for the other categories of the employer’s NACE variable were found to be statistically insignificant. These interpretations remain valid under the ceteris paribus assumption.

In the second part of the study, we determined the survival curve for all respondents using the Kaplan-Meier method (Figure 1). Next, using two survival trees, groups of respondents homogeneous in terms of the duration of their first job were identified. For each of them, a survival curve was estimated. These curves made it possible to assess the probability of continuing employment at each point over time. Moreover, the applied approach allowed the detection of the relations between the examined determinants of the duration of the first employment. In the study, individual characteristics of respondents and work-related factors were considered separately. The results are shown in Figures 2 and 3.

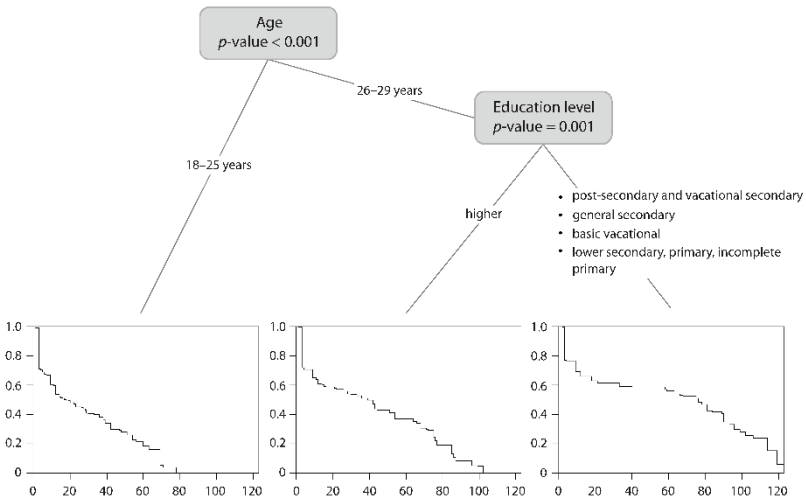
Considering the individual characteristics of the respondents only (Figure 2), it turned out that the key factors influencing the duration of the first job were 'age' and 'education level'. A conclusion could be formed on the basis of the estimated survival curves that the probability of staying in the first employment throughout the studied period decreased more quickly in the case of respondents aged 18 to 25 than in the case of those aged 26 to 29. Moreover, this probability decreased more quickly in relation to people aged 26 to 29 with higher education than in the case of respondents aged 26 to 29 with education level lower than tertiary. Regarding people aged 18 to 25, the probability of staying in the same job decreased very quickly during the first year of employment, and was about 0.6 afterwards. After two years, the probability of remaining in the same organisation was approximately 0.4. For people aged 26 to 29 with the education level lower than tertiary, the probability of not terminating their first employment also decreased rapidly during the first year of employment. However, in their case, the probability of staying in their first job between the second and fifth year of employment remained at a similar level, i.e. at around 0.6. For people aged 26 to 29 with tertiary education, the probability of remaining in the same organisation decreased at a similar rate throughout the whole studied period.

Figure 1. Survival curve plotted by using the Kaplan-Meier method



Source: authors' work based on Statistics Poland's Labour Force Survey.

Figure 2. Survival tree for individual characteristics of young employees in Poland

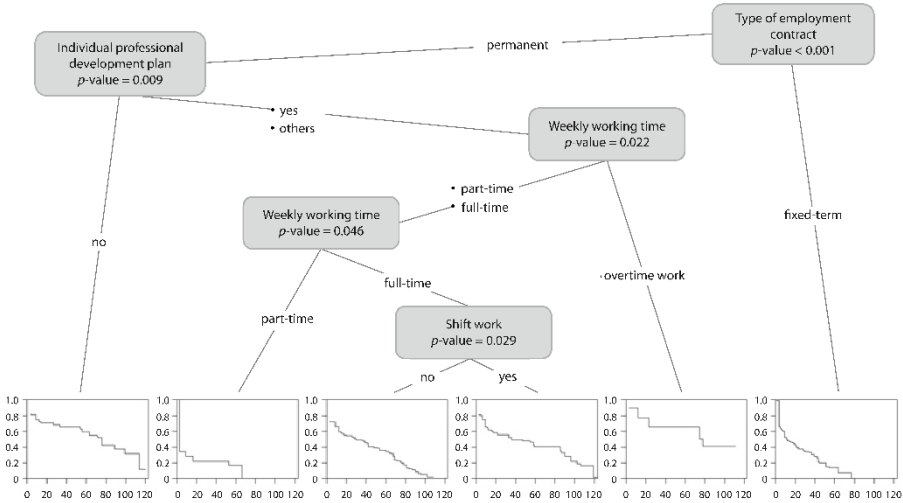


Note. x axis – time in months, y axis – survival function.
Source: authors' work based on Statistics Poland's Labour Force Survey.

Using the survival tree analysis and the work-related factors, six subgroups of respondents statistically significantly different in terms of survival functions were identified (Figure 3). It was found that the key factor influencing the termination of the first job was ‘type of employment contract’. The survival function for respondents with a fixed-term contract decreased very quickly in the first year of employment; the probability of employment termination after one year reached 0.5. For people with a permanent contract, the second significant work-related factor was an individual professional development plan. Interestingly, for those who did not have such a plan, the probability of quitting decreased very slowly compared to other groups of respondents.

Outside the group of respondents who worked overtime weekly, the probability of quitting did not drop below 0.4, which may stem from two reasons. Firstly, the sample size of this group was not substantial. Additionally, workers who did overtime tended to change employment relatively often. For other respondents, weekly working time proved to be the key factor. For respondents who worked full-time, it was important whether they worked in shifts. Working in such a system resulted in a lower probability of quitting within the studied period than in the case of working according to a fixed schedule. The study used SAS software and R language.

Figure 3. Survival tree for work-related factors



Note. As in Figure 2.
Source: authors' work based on Statistics Poland's Labour Force Survey.

5. Discussion

Demographic changes result in a decrease in labour supply (Lewandowski & Magda, 2018). This paper focuses on young workers and their first job. Preliminary data analysis confirmed our assumptions about the high volatility of most of the individual characteristics of young people as well as work-related factors over time. In addition, some variables appeared to be time-dependent.

In the Introduction, we formulated four research questions:

1. Which characteristics of young employees influence the risk that they would leave their first job?
2. Which work-related factors influence the risk that they would leave their first job?
3. What is the profile of a young employee who is not very likely to leave his or her employer?
4. Which work-related factors can keep young workers in the organisation for longer, and which push them away?

For the first and second research questions, we used a time-dependent Cox model to identify risk factors for the termination of the first job. For the third and fourth, we identified groups of employees homogeneous in terms of the time they spent in their first job. These groups were identified using survival trees. Additionally, the survival curves were drawn for each group, which enabled the assessment of changes over time in the probability of staying with the same employer.

On the basis of the results obtained by means of the time-dependent Cox model, it can be concluded that for the surveyed respondents, the probability of terminating their first job decreased with age. This outcome is consistent with previous research on young people in Belgium, which showed a positive correlation between the age of workers and the duration of the employment relationship (Steenackers & Guerry, 2016). In addition, we found that married people were less likely to quit their jobs than other people. Our results also comply with previous studies for men in Canada (Ignaczak et al., 2022) and women in Poland (Grzenda, 2019). Therefore, it can be assumed that young people, after getting married, seek job stability to ensure financial security and adequate conditions for expanding their family. Moreover, in the case of young pregnant women, such results may be the consequence of the labour market regulations in Poland.

Ignaczak et al. (2022) indicated that the effect of education on the probability of quitting the first job differed between genders. Our study does not confirm such differences. We found that people with higher education had a greater tendency to leave their first job than the least educated people. The result obtained for Poland might be affected by the fact that people with higher education at the initial stage of their professional career change jobs more often in search of a job that is consistent

with their education, interests, and qualifications or closer to their financial expectations. According to AbouAssi et al. (2021), millennials are perceived as employees who are driven by a higher pay and who feel dissatisfied with their jobs relatively quickly, as a consequence of which they are prone to leave their jobs.

The second group of determinants of the probability of quitting that we considered were work-related factors. As we expected, in this group of factors, the type of employment contract had the greatest impact on the risk of terminating a job. Young people with fixed-term contracts were over three times more likely to quit their jobs than those with permanent contracts. There are several potential reasons for this. On the one hand, it might be due to the type of contract; some of them expire without the possibility of renewal (e.g. internship contracts, apprenticeships) or are not renewed due to other reasons. Among the latter, there is the young employee's decision made due to his or her dissatisfaction with the working conditions offered by the employer, the effect of the employee's insufficient skills required for his or her position, or the lack of suitable commitment on their part. The discontinuation of the contract may also result from reasons within the employer's responsibility (e.g. the company's economic situation, caused, for instance, by the recent COVID-19 pandemic).

As reported by Portugal and Varejão (2022), fixed-term contracts are often used by employers as a form of verification of the skills and the commitment of applicants for given positions, which in the case of the positive assessment of the employee are then changed into permanent contracts. An interesting result was obtained for the 'weekly working time' variable. The study showed that those who worked overtime were less likely to terminate their job than those working full-time without overtime. This may be due to young people's determination to retain their position, their drive for quick promotion, or having a job they are passionate about. A similar conclusion was reached by Musinszki et al. (2020) for Generation Y and Z in Hungary and Romania. In their study, it was found that young people work overtime because of fear of losing their job or because overtime is paid, which in effect is treated as additional earnings. Based on the results of our research, people in managerial positions were less likely to terminate their first job.

To answer the third and fourth questions, individual characteristics of respondents and work-related factors were analysed separately using survival trees. The results are consistent with the outcome of the time-dependent Cox model. In the case of individual respondent characteristics, the survival curves indicate a faster decline in the probability that people aged 18 to 25 would remain in their first job than in the case of other respondents throughout the study period. Moreover, people aged 26 to 29 with tertiary education were more likely to quit their first jobs over the analysed period than less-educated people in the same age group.

As regards work-related factors, it was found that throughout the study period, those with a permanent contract, an individual professional development plan and not working full-time, were less likely to stay with their first employer than those with a fixed-term contract. This result is consistent with the findings of Steenackers and Guerry (2016), who showed a link between part-time work and a higher likelihood of quitting. In contrast, people with a permanent contract and without an individual professional development plan were the most likely to remain with their first employer. This surprising result concerning an individual professional development plan may be related to the size of the organisation in which someone is employed, i.e. such plans are most often created by human resources departments of larger organisations, while missing in small companies. Furthermore, based on the shape of the majority of the estimated survival curves, it can be inferred that the minimum period of employment (365 days in Poland) needed to obtain unemployment benefits has a strong impact on the probability of terminating the first employment contract. This thesis is supported by a study by Immervoll and Knotz (2018).

Importantly, for some respondents, their first professional experience coincided with the COVID-19 pandemic. Its adverse effects predominantly affected individuals under 24 years of age (GUS, 2021), which was also reflected in the results of our research. Moreover, the pandemic impacted the economic activity of women and men in a different way. Comparing 2020 to 2019, greater discrepancies in the economic activity were observed in individual quarters among women than among men. This could be attributed to constraints on access to childcare facilities and remote learning (GUS, 2021). Our study did not show differences in the duration of employment based on gender. This may be due to the age of the respondents, as well as their other individual characteristics.

6. Conclusions

Our paper identifies the individual characteristics of young workers and work-related factors that affect the time spent in their first job. Most of the studies conducted so far have shown clear differences in the duration of employment according to gender. Our research indicates changes taking place in this area. We found that gender is not a factor that currently affects the time spent by young people in Poland in their first job. We found that the probability of quitting the first job decreases with the age of the respondent, but also depends on the level of education. Moreover, offering young workers an indefinite-term first contract makes them stay in the job for longer. However, even with such a contract, in the case of part-time work, the probability of quitting is higher than in the case of fixed-term contracts.

Results of our research may be helpful in defining the profile of an applicant who is more likely to continue in a job, and may also help companies provide working conditions that would encourage employees to stay with the same employer.

Our study had some limitations, i.e. it was not possible to use weights while constructing survival trees. There is such a possibility, however, in the case of a semiparametric model, which was one of the research tools used in this work. The results presented in this article may constitute the basis for future research on the employer-switching behaviour of young people. Assessing the probability of quitting the first job, but also the frequency of job changes, could help researchers and employers gain a deeper understanding of the labour market behaviour of young people in Poland.

References

- AbouAssi, K., McGinnis Johnson, J., & Holt, S. B. (2021). Job mobility among millennials: Do they stay or do they go?. *Review of Public Personnel Administration*, 41(2), 219–249. <https://doi.org/10.1177/0734371X19874396>.
- Arnett, J. J. (1998). Learning to Stand Alone: The Contemporary American Transition to Adulthood in Cultural and Historical Context. *Human Development*, 41(5/6), 295–315. <https://doi.org/10.1159/000022591>.
- Arnett, J. J. (2006). Emerging adulthood: Understanding the new way of coming of age. In J. J. Arnett, & J. L. Tanner (Eds.), *Emerging adults in America: Coming of age in the 21st century* (pp. 3–19). American Psychological Association. <https://doi.org/10.1037/11381-001>.
- Baert, B. S., Neyt, B., Omey, E., & Verhaest, D. (2022). Student work during secondary education, educational achievement, and later employment: a dynamic approach. *Empirical Economics*, 63(3), 1605–1635. <https://doi.org/10.1007/s00181-021-02172-7>.
- Blanche, P., Zareini, B., & Rasmussen, P. V. (2022). A hazard ratio above one does not necessarily mean higher risk, when using a time-dependent Cox model. *Research Methods in Medicine & Health Sciences*, 3(2), 42–48. <https://doi.org/10.1177/26320843211061288>.
- Boon, C., & Biron, M. (2016). Temporal issues in person-organization fit, person-job fit and turnover: The role of leader-member exchange. *Human Relations*, 69(12), 2177–2200. <https://doi.org/10.1177/0018726716636945>.
- Bou-Hamad, I., Larocque, D., & Ben-Ameur, H. (2011). A review of survival trees. *Statistics Surveys*, 5, 44–71. <https://doi.org/10.1214/09-SS047>.
- Cockx, B., & Picchio, M. (2012). Are short-lived jobs stepping stones to long-lasting jobs?. *Oxford Bulletin of Economics and Statistics*, 74(5), 646–675. <https://doi.org/10.1111/j.1468-0084.2011.00668.x>.
- Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62(2), 269–276. <https://doi.org/10.1093/biomet/62.2.269>.
- Cox, D. R., & Oakes, D. (1984). *Analysis of Survival Data*. Chapman & Hall.
- Das, B. L., & Baruah, M. (2013). Employee Retention: A Review of Literature. *Journal of Business and Management*, 14(2), 8–16. <https://doi.org/10.9790/487X-1420816>.

- De Rose, A., & Pallara, A. (1997). Survival trees: An alternative non-parametric multivariate technique for life history analysis. *European Journal of Population*, 13(3), 223–241. <https://doi.org/10.1023/A:1005844818027>.
- Domurath, A., Taggar, S., & Patzelt, H. (2023). A contingency model of employees' turnover intent in young ventures. *Small Business Economics*, 60(3), 901–927. <https://doi.org/10.1007/s11187-022-00629-2>.
- Eurostat. (2022). *Unemployment by sex and age – annual data*. https://doi.org/10.2908/UNE_RT_A.
- Fisher, L. D., & Lin, D. Y. (1999). Time-dependent covariates in the Cox proportional-hazards regression model. *Annual Review of Public Health*, 20(1), 145–157. <https://doi.org/10.1146/annurev.publhealth.20.1.145>.
- Gadomska-Lila, K. (2015). Pokolenie Y wyzwaniem dla zarządzania zasobami ludzkimi. *Zarządzanie Zasobami Ludzkimi*, (1), 25–39.
- Główny Urząd Statystyczny. (2018). *Zeszyt metodologiczny. Badanie Aktywności Ekonomicznej Ludności*. https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5818/3/1/1/zeszyt_metodologiczny_bael.pdf.
- Główny Urząd Statystyczny. (2019). *Aktywność ekonomiczna ludności Polski I kwartał 2019 r.* <https://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-bezrobotni-bierni-zawodowo-wg-bael/aktywnosc-ekonomiczna-ludnosci-polski-i-kwartal-2019-roku,4,33.html>.
- Główny Urząd Statystyczny. (2021). *Wybrane aspekty rynku pracy w Polsce. Aktywność ekonomiczna ludności przed i w czasie pandemii COVID-19*. <https://stat.gov.pl/obszary-tematyczne/rynek-pracy/zasady-metodyczne-rocznik-pracy/wybrane-aspekty-rynku-pracy-w-polsce-aktywnosc-ekonomiczna-ludnosci-przed-i-w-czasie-pandemii-covid-19,11,1.html>.
- Grzenda, W. (2019). *Modelowanie karier zawodowej i rodzinnej z wykorzystaniem podejścia bayesowskiego*. Wydawnictwo Naukowe PWN.
- Hassan, M. M., Jambulingam, M., Alagas, E. N., Uzir, M. U. H., & Halbusi, H. A. (2023). Necessities and ways of combating dissatisfactions at workplaces against the Job-Hopping Generation Y employees. *Global Business Review*, 24(6), 1276–1301. <https://doi.org/10.1177/0972150920926966>.
- Horáčková, L., & Kopáček, M. (2018). Generation Y on the labour market: regional analysis of the Visegrad Group countries. *Geografické informácie / Geographical Information*, 22(1), 118–130. <https://doi.org/10.17846/GI.2018.22.1.118-130>.
- Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. John Wiley & Sons. <https://doi.org/10.1002/9780470258019>.
- Ignaczak, L. (2014). A Birth Cohort Analysis of First Employment Spells. *Applied Mathematics*, 5(11), 1651–1671. <https://doi.org/10.4236/am.2014.511159>.
- Ignaczak, L., Raffestin, L., & Voia, M. (2022). Do the determinants of employment duration vary across employment spells?. *Applied Economics*, 54(9), 1011–1029. <https://doi.org/10.1080/00036846.2021.1972084>.
- Immervoll, H., & Knotz, C. (2018). *How demanding are activation requirements for jobseekers?* (OECD Social, Employment and Migration Working Papers No. 215). <https://doi.org/10.1787/2bdfecca-en>.

- Kalbfleisch, J. D., & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley & Sons. <https://doi.org/10.1002/9781118032985>.
- Klein, J. P., & Moeschberger, M. L. (2006). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer Science & Business Media. <https://doi.org/10.1007/b97377>.
- Lewandowski, P., & Magda, I. (2018). The labor market in Poland, 2000–2016. *IZA World of Labor*, (426). <https://doi.org/10.15185/izawol.426>.
- McGinnis Johnson, J., & Ng, E. S. (2016). Money Talks or Millennials Walk: The Effect of Compensation on Nonprofit Millennial Workers Sector-Switching Intentions. *Review of Public Personnel Administration*, 36(3), 283–305. <https://doi.org/10.1177/0734371X15587980>.
- Musinszki, Z., Vallasek, M., Mélypataki, G., Horváthné Csolák, E., & Lipták, K. (2020). Workaholism and a New Generation – Labour Market Survey among Hungarian and Romanian Youth. *Amfiteatru Economic*, 22(14), 1227–1242. <https://doi.org/10.24818/EA/2020/S14/1227>.
- Ng'andu, N. H. (1997). An empirical comparison of statistical tests for assessing the proportional hazards assumption of Cox's model. *Statistics in Medicine*, 16(6), 611–626. [https://doi.org/10.1002/\(SICI\)1097-0258\(19970330\)16:6<611::AID-SIM437>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1097-0258(19970330)16:6<611::AID-SIM437>3.0.CO;2-T).
- Pandey, D. L. (2019). Job Hopping Tendency In Millenials. *NCC Journal*, 4(1), 41–46. <https://doi.org/10.3126/nccj.v4i1.24733>.
- Portugal, P., & Varejão, J. (2022). Why do firms use fixed-term contracts?. *Portuguese Economic Journal*, 21(3), 401–421. <https://doi.org/10.1007/s10258-022-00215-0>.
- Robak, E. (2017). Expectations of generation Y connected with shaping the work-life balance. The case of Poland. *Oeconomia Copernicana*, 8(4), 569–584. <https://doi.org/10.24136/oc.v8i4.35>.
- Shimokawa, A., Kawasaki, Y., & Miyaoka, E. (2015). Comparison of Splitting Methods on Survival Tree. *The International Journal of Biostatistics*, 11(1), 175–188. <https://doi.org/10.1515/ijb-2014-0029>.
- Steenackers, K., & Guerry, M. A. (2016). Determinants of job-hopping: an empirical study in Belgium. *International Journal of Manpower*, 37(3), 494–510. <https://doi.org/10.1108/IJM-09-2014-0184>.
- Wawrzonek, A. (2014). Młodzi dorośli wobec zmian na rynku pracy. *Studia Edukacyjne*, (33), 305–320. <https://doi.org/10.14746/se.2014.33.16>.
- Żarczyńska-Dobiesz, A., & Chomątowska, B. (2014). Pokolenie „Z” na rynku pracy – wyzwania dla zarządzania zasobami ludzkimi. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (350), 405–415. <https://doi.org/10.15611/pn.2014.350.36>.

Impact of Hofstede's cultural dimensions on insurance demand¹

Norbert Duczkowski,^a Adam Śliwiński,^b Lubomir Słowik^c

Abstract. Factors determining the demand for different types of insurance and consumer behaviours have been the subject of many studies. The analysed factors included cultural dimensions, i.e. certain aspects of culture that could be measured, which allows determining the position of a given culture in relation to others. However, the research results were very diverse, which motivated the authors to conduct their own research. The aim of the study discussed in the article was to verify the relationship between Hofstede's cultural dimensions and insurance demand in selected European countries. It covered various types of insurance: life insurance, property and casualty insurance, motor insurance, general liability insurance and health insurance, as well as all types of insurance in total. The study used secondary data relating to 27 countries and the years 2004–2020 presented on the Insurance Europe and Geert Hofstede websites. Six Hofstede's cultural dimension indices: the power distance index, individualism vs. collectivism, masculinity vs. femininity, the uncertainty avoidance index, long-term normative orientation vs. short-term normative orientation and indulgence vs. restraint, as well as insurance density (total premiums per inhabitant) were used as variables in the models. The research confirmed the lack of or a very weak relationship between most Hofstede's cultural dimensions and the demand for insurance, except the indulgence dimension. The relationship between this factor and insurance demand proved significant, which is important not only from a methodological point of view (little interest devoted to this indicator in the literature), but also from a business perspective (e.g. from a marketing communication perspective).

Keywords: insurance, demand, cultural dimensions, Geert Hofstede

JEL: G22, G40

¹ Artykuł został opracowany na podstawie referatu wygłoszonego na pierwszej ogólnopolskiej Konferencji Naukowej Doktorantów SGH z cyklu „Nauka odpowiedzią na niepewność” pn. „Interdyscyplinarne podejście do sytuacji kryzysowej w Europie”, która odbyła się w dniach 24–25 marca 2023 r. w Warszawie, oraz na 2023 Second Annual Conference of the Eastern European Risk and Insurance Association, która odbyła się w dniach 24–26 września 2023 r. w Poznaniu / The article is based on a paper delivered at the first national SGH Doctoral Student Conference in the series 'Science – the answer to uncertainty' entitled 'Interdisciplinary approach to a crisis situation in Europe', held on 24–25th March 2023 in Warsaw, and EERIA Second Annual Conference, held on 24–26th September 2023 in Poznań.

^a Szkoła Doktorska, Szkoła Główna Handlowa w Warszawie, Polska / Doctoral School, SGH Warsaw School of Economics, Poland. ORCID: <https://orcid.org/0000-0003-4543-2678>. Autor korespondencyjny / Corresponding author, e-mail: nd110753@doktorant.sgh.waw.pl.

^b Szkoła Główna Handlowa w Warszawie, Kolegium Zarządzania i Finansów, Instytut Ryzyka i Rynków Finansowych, Polska / SGH Warsaw School of Economics, Collegium of Management and Finance, Institute of Risk and Financial Markets, Poland. ORCID: <https://orcid.org/0000-0002-7817-0101>. E-mail: asliwin@sgh.waw.pl.

^c Szkoła Doktorska, Uniwersytet Jagielloński w Krakowie, Polska / Doctoral School, Jagiellonian University in Kraków, Poland. ORCID: <https://orcid.org/0000-0002-9817-9576>. E-mail: lubomir.slowik@doctoral.uj.edu.pl.

Wpływ wymiarów kulturowych Hofstedeego na popyt na ubezpieczenia

Streszczenie. Czynniki determinujące popyt na różnego rodzaju ubezpieczenia i zachowania konsumentów były przedmiotem licznych badań. Do analizowanych czynników należały m.in. wymiary kulturowe, czyli pewne aspekty, które można zmierzyć i dzięki temu określić pozycję danej kultury w stosunku do innych. Jednak wyniki badań były bardzo zróżnicowane, co skłoniło autorów do podjęcia własnego badania. Celem badania omawianego w artykule jest weryfikacja zależności między wymiarami kulturowymi Hofstedeego a popytem na ubezpieczenia w wybranych krajach europejskich. Obejmuje ono różne rodzaje ubezpieczeń: na życie, majątkowe i osobowe, komunikacyjne, odpowiedzialności cywilnej i zdrowotne oraz wszystkie rodzaje ogółem. Wykorzystano dane wtórne za lata 2004–2020, które dotyczyły 27 krajów i zostały zaczerpnięte z serwisu Insurance Europe oraz ze strony Geert Hofstede. Jako zmienne w modelach przyjęto wskaźniki dotyczące sześciu wymiarów kulturowych Hofstedeego: dystansu władzy, indywidualizmu i kolektywizmu, męskości i kobiecości, unikania niepewności, orientacji długoterminowej i krótkoterminowej oraz pobożania i powściągliwości, a także gęstość ubezpieczeń (składki ogółem na mieszkańca). Badanie wykazało brak korelacji lub bardzo słabą korelację większości analizowanych wymiarów kulturowych i popytu na ubezpieczenia z wyjątkiem powściągliwości. Związek tego wymiaru z popytem na ubezpieczenia okazał się istotny, co jest znaczące nie tylko z metodologicznego (małe zainteresowanie tym wskaźnikiem w literaturze), lecz także biznesowego punktu widzenia (np. z perspektywy komunikacji marketingowej).

Słowa kluczowe: ubezpieczenia, popyt, wymiary kulturowe, Geert Hofstede

1. Introduction

Culture plays an important role in all spheres of human life. Therefore, economic analyses increasingly often take into account cultural factors, as they play a significant role in the sphere of business, economics, finance and management (Kostro, 2009). Although it is more and more common to state that the role of cultural factors in economic processes will decrease as globalisation gains momentum, this role will not be completely marginalised. Despite the convergence of the economic environments around the world based on the emergence of a universal system of values, beliefs and attitudes, cultural differences will continue to exist (Kostro, 2009).

One of the breakthroughs in the study of culture came with the cultural dimensions theory, formulated in 1980 by a Dutch management expert, Geert Hofstede. It is presented in a variety of papers, including those of Hofstede (1983, 1995, 2001). This concept was established as a framework for cross-cultural communication. The theory is based on factor analysis and shows the effects of a society's culture on the values of its members and how these values relate to their behaviour. Hofstede (2001) included six key aspects of national culture presented on country comparison scales: the power distance index (PDI), individualism vs. collectivism (IDV), masculinity vs. femininity (MAS), the uncertainty avoidance index

(UAI), long-term normative orientation vs. short-term normative orientation (LTO) and indulgence vs. restraint (IVR). The creation of cultural indices made it relatively easy to incorporate cultural factors into management, economic and finance research, including insurance.

Since Hofstede's (1995) publication in *The Geneva Papers on Risk and Insurance*, which opened the door to research on the influence of cultural factors on the insurance market (Outreville, 2013), many widely cited papers have been published (Chui & Kwok, 2009; Park et al., 2002; Park & Lemaire, 2011a, 2011b; Treerattanapun, 2011). Nevertheless, it is surprising that this subject has not yet been thoroughly explored (Outreville, 2013). Furthermore, the results presented in research works are very diverse, which is a direct motivation of the research presented in this paper.

The aim of this study is to verify the relationship between Hofstede's cultural dimensions and insurance demand in selected European countries. Hofstede's cultural dimensions indices as well as insurance density (total premiums per inhabitant in EUR) in constant exchange rates have been used as variables in the models. The relationship between culture and insurance was studied for insurance in general and for its particular types: life insurance, property and casualty insurance (P&C), motor insurance, property insurance, general liability (GL) insurance and health insurance.

2. Hofstede's cultural dimensions

Hofstede's cultural dimensions were first identified during a sizable study carried out throughout the 1960s and 1970s that examined value variations among the different divisions of IBM, a worldwide computer manufacturer (Hofstede, 2001).

Over 100,000 workers from 50 different nations participated in this survey which concerned three business areas of the firm. Using a particular statistical technique known as factor analysis, Hofstede defined the first four value dimensions: IDV, PDI, UAI, and MAS. Later Chinese sociologists discovered the fifth dimension in their investigation, i.e. LTO (Bond, 1991). Finally, a replication of Hofstede's research (2001) carried out in 93 different countries, confirmed the validity of the five dimensions and discovered a sixth, referred to as IVR. The dimensions of culture established by Hofstede are briefly characterised in Table 1. The exact method of each cultural index calculation can be found on the geerthofstede.com website (in the 'values survey module 2013' and 'research and values survey module' section), as well as in Hofstede (2013), where detailed manuals and methodological notes are provided.

Table 1. Hofstede’s cultural dimensions definitions

Dimension	Definition
Power distance index (PDI)	‘The extent to which the less powerful members of organizations and institutions (like the family) accept and expect that power is distributed unequally’ is how the PDI is defined. A higher PDI means that a society’s hierarchy is undeniably established and followed without hesitation or justification. A lower value of the index indicates that individuals challenge authority and want to share power.
Individualism vs. collectivism (IDV)	IDV investigates how well members of a society are integrated into organisations. Individualistic cultures tend to have weak relationships that frequently connect a person only to his or her immediate family. Instead of emphasizing ‘us’, they focus on ‘I’. Its opposite, collectivism, defines a society where closely interwoven ties arrange people into in-groups, including extended families. These in-groups have an undeniable bond of loyalty and stand by one another when there is a dispute with another in-group arises.
Uncertainty avoidance index (UAI)	UAI is referred to as ‘a society’s tolerance for ambiguity’, and it measures how willingly or reluctantly individuals accept or avoid an occurrence that is unexpected, uncharted or inconsistent with the <i>status quo</i> . Societies with high UAIs tend to have strict rules of conduct, regulations and a general belief in absolute truth or the idea that there is one single truth that governs everything and that everyone is aware of it. A lower UAI indicates a greater tolerance for opposing viewpoints. The environment is more free-flowing, ambiguity is more common and the society tends to impose few restrictions.
Masculinity vs. femininity (MAS)	Men are classified as being masculine if they ‘prefer achievement, heroism, aggressiveness, and pecuniary incentives for success in society’. A propensity for collaboration, modesty, helping the weak and quality of life is represented by its opposite, i.e. femininity. Women tend to exhibit distinct ideals in their respective communities. They hold the same humble and compassionate ideals as males in feminine civilizations. On the other hand, women are relatively forceful and competitive in a society that values masculinity, but far less so than males. In other words, women still acknowledge the discrepancy between male and female values. In communities that are predominately masculine, this distinction is typically seen as forbidden.
Long-term normative orientation vs. short-term normative orientation (LTO)	This dimension links the relationship between the past and present-day actions and obstacles. A lesser degree of this indicator (short-term) suggests that constancy is prized while customs are respected and upheld. Societies with a high LTO (over the long term) see adaptability and pragmatic, contextual problem-solving as essential. While long-term-oriented nations continue to rise to a level of wealth, short-term-oriented impoverished nations often have little or no economic progress.
Indulgence vs. restraint (IVR)	This dimension is the extent to which societal standards permit individuals to pursue their human aspirations. According to the definition of indulgence, it is ‘a culture that permits virtually unrestricted satisfaction of fundamental and normal human wants to be connected to enjoying life and having pleasure’. Its opposite (restraint) is described as ‘a society that controls gratification of needs and regulates it by means of strict social norms’.

Source: authors’ work based on Hofstede (2001).

Hofstede’s book on the cultural dimensions model, published in 1980 became one of the most widely referenced works in the sphere of social sciences (Green, 2016). The cultural dimensions theory has sparked a great deal of empirical research over the past 40 years, broadening the field of study from pure psychology to many other areas, making the study of its use and effects a challenging endeavour. Some studies have, nevertheless, focused on the application and impact of Hofstede’s cultural

model within the context of this research stream in order to recommend future research options (Beugelsdijk et al., 2017; Kirkman et al., 2006; Taras et al., 2010).

It is worth noting, however, that there is a noticeable variability in terms of researchers' interest shown in each of Hofstede's dimensions. A review of the literature on the subject indicates that the PDI, IDV and MAS cultural dimensions are underrepresented relative to the remaining values. However, according to the table below, the IVR dimension has been the least occurring theme in Hofstede-inspired research (Leonavičienė & Burinskienė, 2022). Although Table 2 concerns research on internationalisation, it provides an interesting view of the popularity of each cultural dimension among researchers.

Table 2. Extent of Hofstede's cultural dimensions research

Year	Number of papers	Hofstede's cultural dimensions					
		PDI	IDV	MAS	UAI	LTO	IVR
Total	12,147	66	157	16	1,303	2,779	15
1994–1998	360	2	3	–	6	10	–
1999–2003	487	2	2	1	103	141	3
2004–2008	1,790	7	7	3	245	453	2
2009–2013	2,040	40	58	4	287	843	3
2014–2018	3,870	7	79	4	329	656	3
2019–2021	3,600	8	8	4	333	676	4

Source: Leonavičienė and Burinskienė (2022).

Research shows that since the initial publishing of Hofstede's theory in 1980, the cultural dimensions framework has been incorporated in more than 1,000 studies (Kirkman et al., 2006; Zhou & Kwon, 2020). Authors of the meta-analysis of Hofstede-related inquiries indicate that the publications primarily cover psychology, sociology and general social sciences, therefore there is a theoretical and application gap in the context of Hofstede-inspired research in the economy and insurance sector (Zhou & Kwon, 2020). This paper addresses not only this application gap but also focuses on the least covered cultural dimension, i.e. IVR.

Table 3 provides an exhaustive literature review of Hofstede-inspired research in the following areas: change management, human resource management (HRM), leadership & entrepreneurship, work-related attitudes, negotiation, group processes and personality, entry modes, foreign direct investment, joint venture characteristics and performance, innovation and research and development, societal outcomes (e.g. wealth, national accounting systems, number of intellectual property violations) and insurance (Kirkman et al., 2006). A literature review devoted specifically to the area of insurance is presented in Section 3 of the paper.

Table 3. Summary of the subjects covered in Hofstede's cultural dimensions research papers

Covered subject	Research papers
Change management	Firican (2022), Seidenfuss & Storm (2022)
Human resource management	Newman and Nollen (1996), Roth & O'Donnell (1996), Ryan et al. (1999)
Leadership & entrepreneurship	Offerman & Hellmann (1997), Thomas & Mueller (2000)
Work-related attitudes	Gong et al. (2001), Peterson & Smith (1997), Robie et al. (1998), Schneider & De Meyer (1991), Van de Vliert & Van Yperen (1996)
Negotiation	Shen (2023), Smith et al. (1998)
Group processes and personality	Krug & Nigh (1998), Oyserman et al. (2002)
Entry modes	Anand & Delios (1997), Arora & Fosfuri (2000), Barkema & Vermeulen (1998), Brouthers & Brouthers (2000, 2001), Chang & Rosenzweig (2001), Erramilli (1991, 1996), Erramilli et al. (1997), Erramilli et al. (2002), Harzing (2002), Hennart & Larimo (1998), Kim & Hwang (1992), Kogut & Singh (1988), Nachum (2003), Pan (1996, 2002), Shane (1992, 1994)
Foreign direct investment	Andrianova (2022), Benito & Gripsrud (1992), Habib & Zurawicki (2002), Kallunki et al. (2001), Li & Guisinger (1992), Loree & Guisinger (1995)
Joint venture characteristics and performance	Barkema et al. (1996), Barkema et al. (1997), Barkema & Vermeulen (1997), Datta & Puia (1995), Glaister & Buckley (1999), Gómez-Mejía & Palich (1997), Håkanson & Nobel (2001), Li & Guisinger (1991), Luo (2001a, 2001b), Luo & Park (2001), Luo & Peng (1999), Luo et al. (2001), Merchant & Schendel (2000), Park & Ungson (1997), Pothukuchi et al. (2002)
Innovation and research and development	Richards & De Carolis (2003), Shane et al. (1995)
Societal outcomes (e.g. wealth, national accounting systems, number of intellectual property violations)	Diener & Diener (1995), Diener et al. (1995), Diener et al. (2000), Franke et al. (1991), Husted (1999), Riahi-Belkaoui (1998), Salter & Niswander (1995)
Insurance	Chui & Kwok (2009), Park et al. (2002), Park & Lemaire (2011a, 2011b), Treerattanapun (2011)

Source: authors' work based on Kirkman et al. (2006), Leonavičienė and Burinskienė (2022), and Outreville (2013).

3. Literature review

As insurance plays a significant role in the economy, factors influencing the demand for insurance have been the subject of many studies (Duczowski, 2022). These factors were analysed across different perspectives such as geographical, social or economic and were also considered in terms of their impact on the different types of insurance, including life and non-life insurance (Duczowski, 2022; Śliwiński, 2019). The multitude of studies and publications has also resulted in the availability of several papers that summarise the results of other research, by authors such as Jaspersen (2016), Śliwiński (2016, 2019), or Zietz (2003). These publications point to the role and importance of cultural factors in the demand for insurance.

The most important publications include those indicated in the Introduction (Chui & Kwok, 2009; Park et al., 2002; Park & Lemaire, 2011a, 2011b; Treerattanapun, 2011). The study published by Park et al. (2002) was one of the first

to examine the relationship between culture and insurance pervasiveness in various countries. Using data from a representative sample of 37 countries across the globe, the authors of the aforementioned paper found that certain cultural and socio-political variables can significantly influence the level of insurance pervasiveness. Specifically, it turned out that Hofstede's MAS dimension of national culture had statistically significant effects (among other variables such as aggregate income, socio-political stability and government regulation).

A classic research position describing the impact of culture on the demand for insurance is also presented in two papers published by Park and Lemaire (2011a, 2011b). In the first one, they used regression techniques to an unbalanced panel data set relating to 68 countries observed over a ten-year period to explore the factors that affect non-life insurance demand. The authors discovered that non-life insurance consumption is adversely impacted in countries where a large fraction of the population holds Islamic beliefs. Moreover, three of Hofstede's dimensions turned out to be statistically significant: PDI, IDV and UAI (Park & Lemaire, 2011a). Another important finding the work shows is that culture impacts non-life insurance in affluent countries, with an adjusted R -square coefficient increasing by 11.7%, more than in developing countries, where the R -square coefficient increase due to cultural impacts is only 1.2%. These results have implications for multinational insurers seeking to enter a new market. *Ceteris paribus*, these insurers should target locations and population segments within countries with a low PDI and high IDV and UAI.

In their other paper, Park and Lemaire (2011b) extend Chui and Kwok's work (2009) by analysing the fifth Hofstede cultural dimension, i.e. LTO. After building a database that includes values of 17 variables for 27 countries over a period of 9 years, they apply an unbalanced panel GLS regression model to prove that LTO has a strong positive influence on life insurance demand. Their findings were then confirmed by several robustness tests.

The study by Chui and Kwok (2009) is also of great importance to the analysis of the impact of culture on insurance demand. It involved cross-disciplinary research that examined how national culture and cultural practices affect cross-country variations in life insurance consumption. The authors use the refined measure of the GLOBE project, which includes several additional cultural dimensions, not presented in Hofstede's analysis, as the proxy for national culture dimensions. Using 1966–2004 data relating to 38 countries, the analysis reveals a strong relationship between life insurance consumption and the practice scores of IDV and PDI. A strong relationship was observed between cultural variables and the insurance demand, which continues to hold even after controlling for other country-level variables, such as national income, expected inflation rate, banking sector development, investor protection index, dependency ratio, life expectancy and religion.

Recent research on the impact of culture on non-life insurance consumption is presented in a paper by Treerattanapun (2011). Various economic, institutional and cultural variables regarding 82 countries across a 10-year period were considered in order to build the most efficient and most parsimonious regression model. This research found that high non-life insurance consumption is observed among nations with a low PDI, and a high IDV and UAI. The empirical results suggest that consumers may respond to insurance solicitations according to their cultural beliefs, not only economic rationality.

Table 4 summarises the cited studies where Hofstede’s dimensions of culture were found to be significant.

Table 4. Selected research papers describing the impact of Hofstede’s cultural dimensions on insurance demand

Research paper	Hofstede’s cultural dimensions considered significant
Park et al. (2002)	MAS
Chui & Kwok (2009)	PDI, IDV
Park & Lemaire (2011a, 2011b)	PDI, IDV, UAI, LTO
Treerattanapun (2011)	PDI, IDV, UAI

Source: authors’ work.

It is worth noting that the IDV dimension is underrepresented in the literature not only pertaining to the subject of insurance but also in general (as presented in Table 2).

Not only the diversity of the results shown in the table above is worth emphasising, but also the fact that the researchers did not always obtain the same results. An example of this is the publication by Esho et al. (2004), who highlighted that the demand for property-liability insurance is not significantly affected by cultural factors. In conclusion, the impact of Hofstede’s cultural dimensions on insurance demand remains unexplored (Outreville, 2013).

4. Research method

In order to conduct the empirical part of this study, widely available secondary data have been used. Hofstede’s cultural dimensions indices come from a website dedicated to this project (Geert Hofstede, n.d.) and the insurance density (EUR, total premiums per inhabitant) in constant exchange come from Insurance Europe (n.d.). The data used to prepare the regression models are presented in Tables 5 and 6. The study was based on statistics from 27 countries covering the years 2004–2020.

Table 5. Hofstede's cultural dimensions indices

Country	PDI	IDV	MAS	UAI	LTO	IVR
Austria	11	55	79	70	60	63
Belgium	65	75	54	94	82	57
Bulgaria	70	30	40	85	69	16
Croatia	73	33	40	80	58	33
Czechia	57	58	57	74	70	29
Denmark	18	74	16	23	35	70
Estonia	40	60	30	60	82	16
Finland	33	63	26	59	38	57
France	68	71	43	86	63	48
Germany	35	67	66	65	83	40
Greece	60	35	57	112	45	50
Hungary	46	80	88	82	58	31
Ireland	28	70	68	35	24	65
Italy	50	76	70	75	61	30
Latvia	44	70	9	63	69	13
Malta	56	59	47	96	47	66
Netherlands	38	80	14	53	67	68
Norway	31	69	8	50	35	55
Poland	68	60	64	93	38	29
Portugal	63	27	31	104	28	33
Romania	90	30	42	90	52	20
Slovakia	104	52	110	51	77	28
Slovenia	71	27	19	88	49	48
Sweden	31	71	5	29	53	78
Switzerland	34	68	70	58	74	66
Turkey	66	37	45	85	46	49
United Kingdom	35	89	66	35	51	69
Average	51	59	47	70	56	45

Source: authors' work based on Geert Hofstede (n.d.).

Table 6. Average insurance density per inhabitant in the years 2004–2020

Country	All insurance	Life insurance	P&C insurance	Motor insurance	Property insurance	GL insurance	Health insurance
	in EUR, constant exchange rates						
Austria	19,55.54	770.69	10,48.56	360.51	295.49	89.06	211.63
Belgium	25,75.67	16,56.48	703.40	315.79	230.19	67.67	268.08
Bulgaria	125.29	18.48	107.44	76.17	17.76	2.43	3.35
Croatia	287.92	80.35	192.51	108.28	39.50	9.81	15.05
Czechia	540.86	217.39	297.48	151.11	83.98	26.86	19.15
Denmark	43,44.33	30,38.19	12,70.94	322.23	417.72	68.44	47.11
Estonia	252.66	64.59	187.19	116.29	48.22	5.57	5.39
Finland	3,733.93	2,987.38	650.40	264.13	160.84	37.31	106.36
France	3,109.79	2,011.15	917.92	295.88	229.02	90.71	207.36
Germany	2,198.85	1,070.10	707.75	287.87	206.36	88.51	421.02
Greece	399.54	189.13	204.19	122.50	28.77	7.67	7.45
Hungary	252.25	124.15	126.08	62.38	44.88	6.12	2.36
Ireland	4,121.12	3,612.84	706.02	323.75	211.73	113.83	517.36
Italy	2,375.79	1,703.31	650.86	324.83	90.07	58.53	38.29
Latvia	198.59	39.14	133.00	73.72	23.14	4.81	30.53
Malta	1,194.12	752.95	381.54	197.99	77.27	24.74	66.67
Netherlands	4,375.54	1,163.78	2,143.32	264.25	204.89	62.67	2,223.19
Norway	2,275.84	1,479.19	840.93	342.52	313.27	26.67	111.75
Poland	264.47	111.00	151.63	89.82	28.66	9.17	2.93
Portugal	1,242.25	806.70	378.62	155.26	69.95	9.94	105.47
Romania	80.59	13.72	63.93	46.72	10.26	2.03	2.90
Slovakia	335.07	186.29	177.97	110.67	42.89	11.26	14.59
Slovenia	959.79	271.46	594.60	246.01	105.77	25.60	219.46
Sweden	2,660.47	1,916.70	630.27	220.49	278.30	29.74	143.97
Switzerland	6,452.57	3,525.51	2,644.90	659.54	455.79	228.45	1,051.12
Turkey	49.26	6.64	35.37	18.86	9.24	1.06	5.65
United Kingdom ...	3,697.87	2,568.31	1,092.60	265.39	267.74	150.12	86.09
Average	1,854	1,125	631	216	148	47	220

Source: authors' work based on Insurance Europe (n.d.).

In the next step, regression models were built between the explained variable (insurance density (total premiums per inhabitant) for all insurance, life insurance, P&C insurance, motor insurance, property insurance, GL insurance and health insurance) and the explanatory variable (all Hofstede's cultural dimensions indices). This is how 42 econometric models were created. However, what is important is that the indices describing Hofstede's cultural dimensions cannot be considered individually, but only when compared to a selected point of reference (e.g. another country). That is why the final form of the estimated models can be presented according to the formula below:

$$\ln(\text{density}_{i,j} / \langle \text{density}_i \rangle) = A_i \cdot \ln(\text{HOF}_{i,j} / \langle \text{HOF}_i \rangle) + C_i, \quad (1)$$

where:

$\text{density}_{i,j}$ – density for i -th type of insurance, j -th country,

$\text{HOF}_{i,j}$ – Hofstede's index value for i -th dimension, j -th country,

$\langle \cdot \rangle$ – arithmetic mean value,

A_i (for i -th dimension) – estimated model parameters (slope coefficient),

C_i – estimated model parameters (constant).

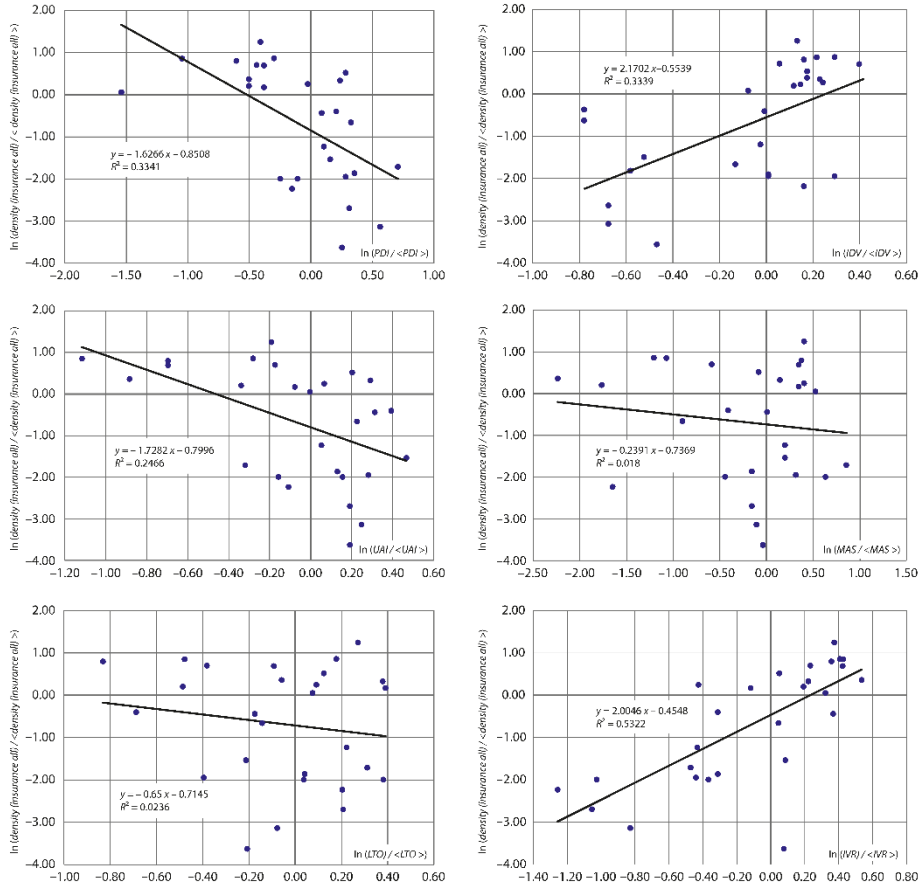
In the next step, econometric models were estimated, where all Hofstede's cultural dimensions indices were analysed in the first iteration for each type of insurance, and then statistically insignificant variables were removed. In other words, equation (2), as presented below, was calculated:

$$\ln(\text{density}_{i,j} / \langle \text{density}_i \rangle) = \sum_{i=1}^6 A_i \cdot \ln(\text{HOF}_{i,j} / \langle \text{HOF}_i \rangle). \quad (2)$$

5. Results and discussion

Figure 1 presents the results of research carried out in accordance with the procedure described in Section 4 for all insurance.

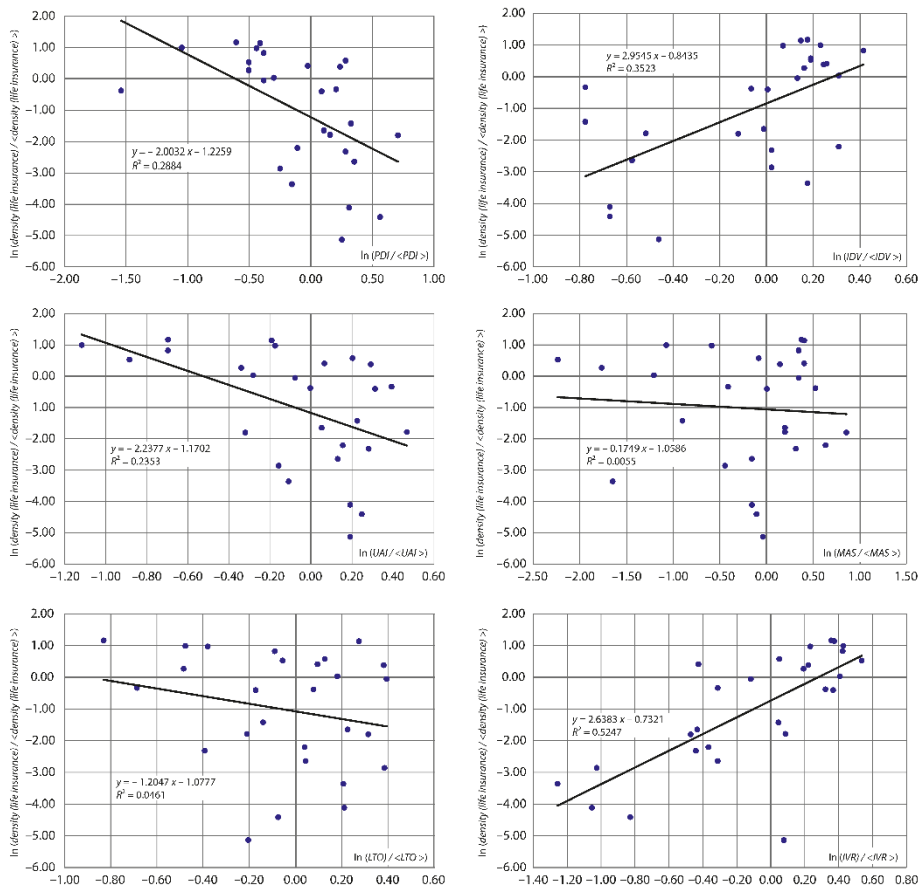
Figure 1. Results of linear regression models for all insurance density in EUR, constant exchange rates



Source: authors' work based on data from Tables 5 and 6.

Figure 2 presents the results of research carried out in accordance with the procedure described in Section 4 for life insurance.

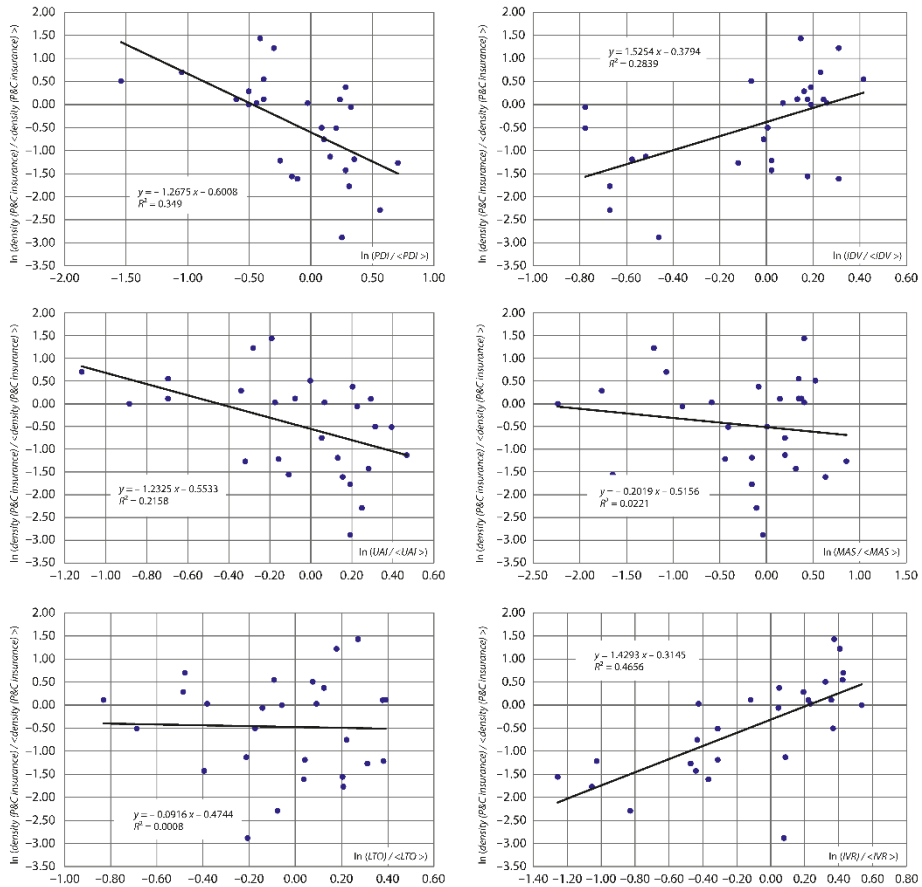
Figure 2. Results of linear regression models for life insurance density in EUR, constant exchange rates



Source: authors' work based on data from Tables 5 and 6.

Figure 3 presents the results of research carried out in accordance with the procedure described in Section 4 for P&C insurance.

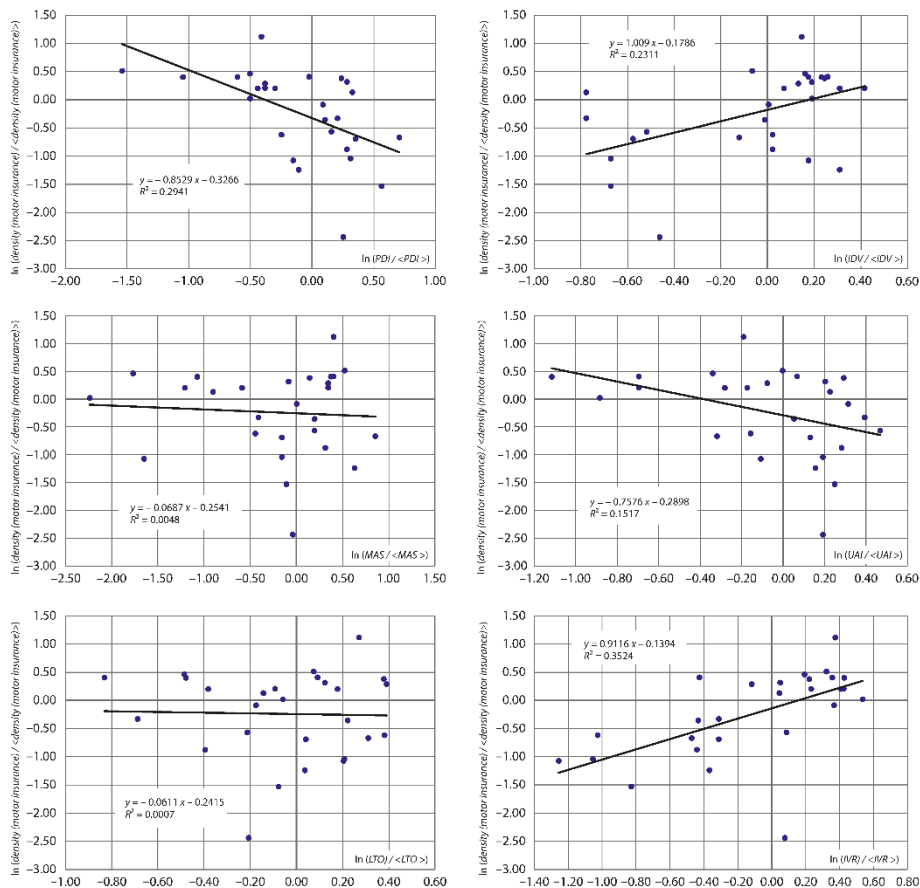
Figure 3. Results of linear regression models for P&C insurance density in EUR, constant exchange rates



Source: authors' work based on data from Tables 5 and 6.

Figure 4 presents the results of research carried out in accordance with the procedure described in Section 4 for motor insurance.

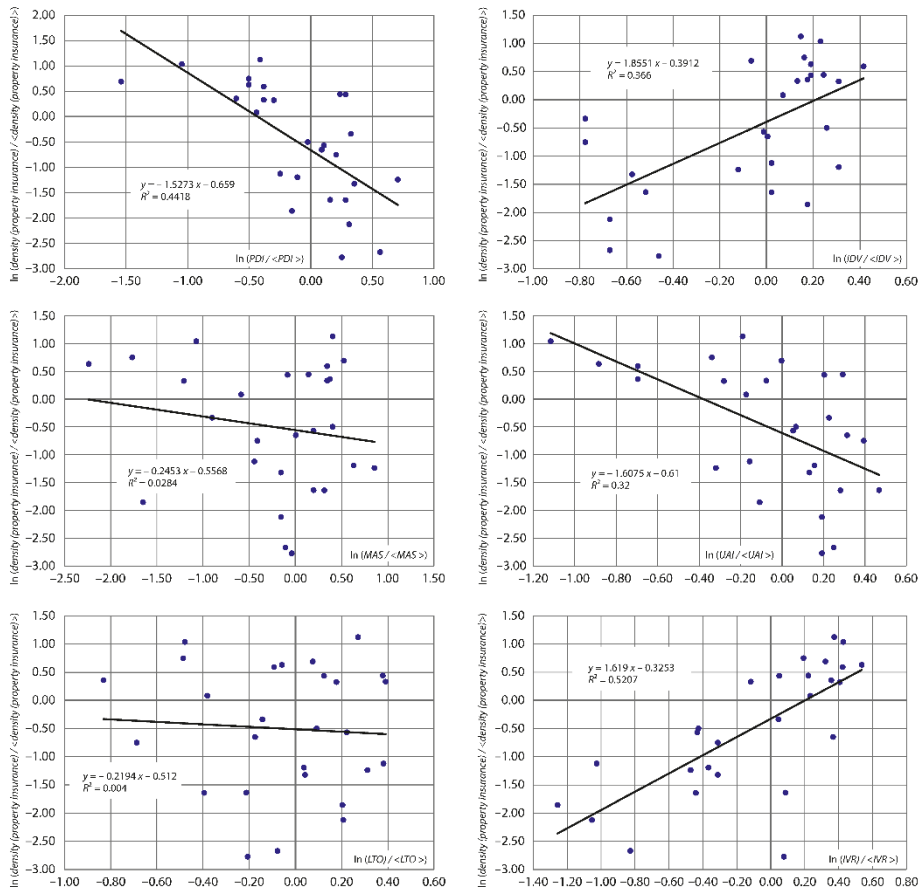
Figure 4. Results of linear regression models for motor insurance density in EUR, constant exchange rates



Source: authors' work based on data from Tables 5 and 6.

Figure 5 presents the results of research carried out in accordance with the procedure described in Section 4 for property insurance.

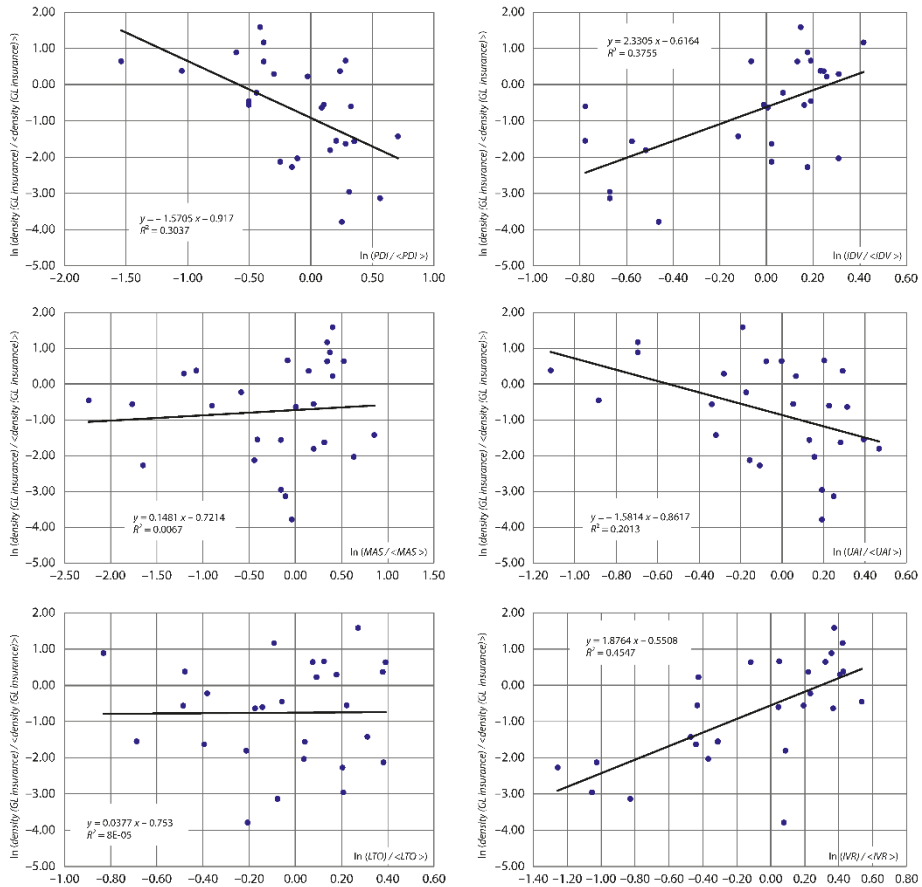
Figure 5. Results of linear regression models for property insurance density in EUR, constant exchange rates



Source: authors' work based on data from Tables 5 and 6.

Figure 6 presents the results of research carried out in accordance with the procedure described in Section 4 for general liability insurance.

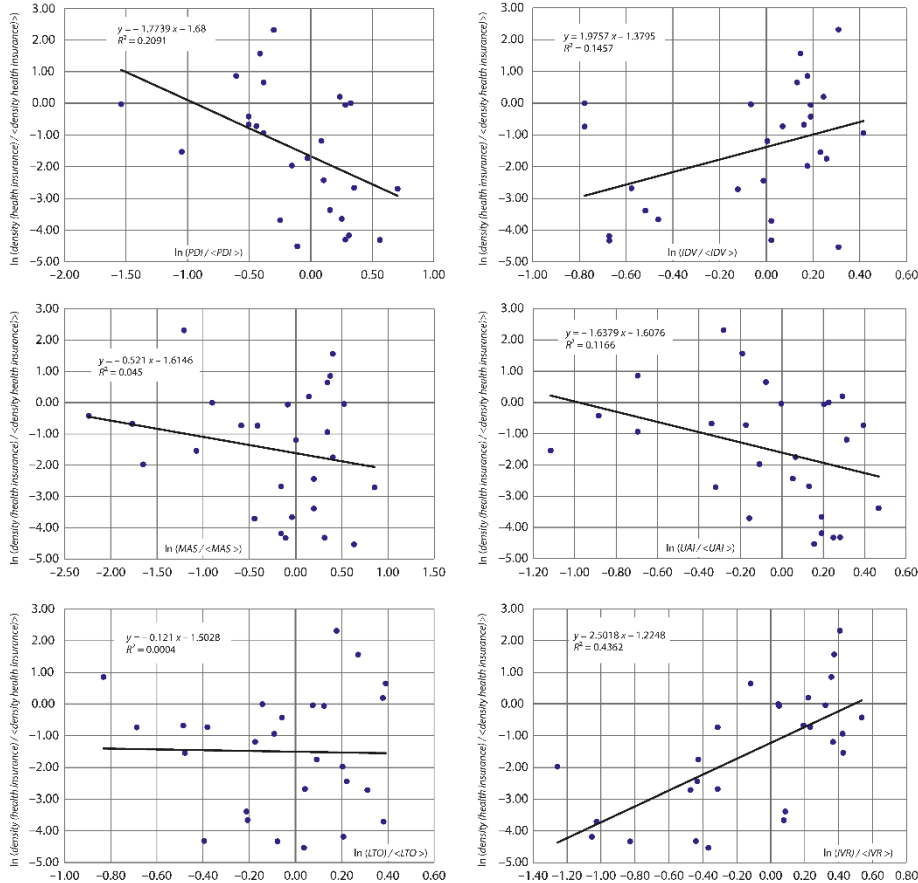
Figure 6. Results of linear regression models for GL insurance density in EUR, constant exchange rates



Source: authors' work based on data from Tables 5 and 6.

Figure 7 presents the results of research carried out in accordance with the procedure described in Section 4 for health insurance.

Figure 7. Results of linear regression models for property insurance density in EUR, constant exchange rates



Source: authors' work based on data from Tables 5 and 6.

Table 7 summarises the linear regression results presented in Figures 1–7 for all the analysed types of insurance.

Table 7. Hofstede’s cultural dimensions regression results – summary

Type of insurance	PDI		IDV		MAS		UAI		LTO		IVR													
All insurance	$y = -1.6266x - 0.8508$ $R^2 = 0.3341$ $F = 12.5406$ significance $F = 0.0016$		$y = 2.1702x - 0.5539$ $R^2 = 0.3339$ $F = 12.5342$ significance $F = 0.0016$		$y = -0.2391x - 0.7369$ $R^2 = 0.018$ $F = 0.4582$ significance $F = 0.5047$		$y = -1.7282x - 0.7996$ $R^2 = 0.2466$ $F = 8.1809$ significance $F = 0.0084$		$y = -0.65x - 0.7145$ $R^2 = 0.0236$ $F = 0.6038$ significance $F = 0.4444$		$y = 2.0046x - 0.4648$ $R^2 = 0.5322$ $F = 28.4405$ significance $F = 1.579E-05$													
	A_i	-1.6266	C	-0.8508	A_i	2.1702	C	-0.5539	A_i	-0.2391	C	-0.7369	A_i	-1.7282	C	-0.7996	A_i	-0.6500	C	-0.7145	A_i	2.0046	C	-0.4648
	SE	0.4593	SE	0.2287	SE	0.6130	SE	0.2267	SE	0.3532	SE	0.2831	SE	0.6042	SE	0.2414	SE	0.8365	SE	0.2739	SE	0.3759	SE	0.1919
	t_{stat}	-3.5413	t_{stat}	-3.7205	t_{stat}	3.5404	t_{stat}	-2.4438	t_{stat}	-0.6769	t_{stat}	-2.6025	t_{stat}	-2.8602	t_{stat}	-3.3121	t_{stat}	-0.7771	t_{stat}	-2.6083	t_{stat}	5.3330	t_{stat}	-2.4221
	$pvalue$	0.0016	$pvalue$	0.0010	$pvalue$	0.0016	$pvalue$	0.0219	$pvalue$	0.5047	$pvalue$	0.0153	$pvalue$	0.0084	$pvalue$	0.0028	$pvalue$	0.4444	$pvalue$	0.0151	$pvalue$	0.0000	$pvalue$	0.0230
Life insurance	$y = -2.0032x - 1.2259$ $R^2 = 0.2884$ $F = 10.1301$ significance $F = 0.0039$		$y = 2.9545x - 0.8435$ $R^2 = 0.3523$ $F = 13.5974$ significance $F = 0.0011$		$y = -0.1749x - 1.0586$ $R^2 = 0.0055$ $F = 0.1378$ significance $F = 0.7136$		$y = -2.2377x - 1.1702$ $R^2 = 0.2353$ $F = 7.6915$ significance $F = 0.0103$		$y = -1.2047x - 1.0777$ $R^2 = 0.0461$ $F = 1.2085$ significance $F = 0.2821$		$y = 2.6383x - 0.7321$ $R^2 = 0.5247$ $F = 27.5968$ significance $F = 1.938E-05$													
	A_i	-2.0032	C	-1.2259	A_i	2.9545	C	-0.8435	A_i	-0.1749	C	-1.0586	A_i	-2.2377	C	-1.1702	A_i	-1.2047	C	-1.0777	A_i	2.6383	C	-0.7321
	SE	0.6294	SE	0.3133	SE	0.8012	SE	0.2963	SE	0.4712	SE	0.3777	SE	0.8069	SE	0.3224	SE	1.0958	SE	0.3589	SE	0.5022	SE	0.2564
	t_{stat}	-3.1828	t_{stat}	-3.9124	t_{stat}	3.6875	t_{stat}	-2.8472	t_{stat}	-0.3712	t_{stat}	-2.8029	t_{stat}	-2.7733	t_{stat}	-3.6300	t_{stat}	-1.0993	t_{stat}	-3.0031	t_{stat}	5.2533	t_{stat}	-2.8554
	$pvalue$	0.0039	$pvalue$	0.0006	$pvalue$	0.0011	$pvalue$	0.0087	$pvalue$	0.7136	$pvalue$	0.0096	$pvalue$	0.0103	$pvalue$	0.0013	$pvalue$	0.2821	$pvalue$	0.0060	$pvalue$	0.0000	$pvalue$	0.0085
P&C insurance	$y = -1.2675x - 0.6008$ $R^2 = 0.349$ $F = 13.4049$ significance $F = 0.0012$		$y = 1.5254x - 0.3794$ $R^2 = 0.2839$ $F = 9.9116$ significance $F = 0.0042$		$y = -0.2019x - 0.5156$ $R^2 = 0.0221$ $F = 0.5647$ significance $F = 0.4593$		$y = -1.2325x - 0.5533$ $R^2 = 0.2158$ $F = 6.8795$ significance $F = 0.0146$		$y = -0.0916x - 0.4744$ $R^2 = 0.0008$ $F = 0.0202$ significance $F = 0.8882$		$y = 1.4293x - 0.3145$ $R^2 = 0.4656$ $F = 21.7798$ significance $F = 8.835E-05$													
	A_i	-1.2675	C	-0.6008	A_i	1.5254	C	-0.3794	A_i	-0.2019	C	-0.5156	A_i	-1.2325	C	-0.5533	A_i	-0.0916	C	-0.4744	A_i	1.4293	C	-0.3145
	SE	0.3462	SE	0.1723	SE	0.4845	SE	0.1792	SE	0.2687	SE	0.2154	SE	0.4699	SE	0.1878	SE	0.6450	SE	0.2112	SE	0.3063	SE	0.1564
	t_{stat}	-3.6613	t_{stat}	-3.4862	t_{stat}	3.1483	t_{stat}	-2.1175	t_{stat}	-0.7515	t_{stat}	-2.3939	t_{stat}	-2.6229	t_{stat}	-2.9467	t_{stat}	-0.1421	t_{stat}	-2.2461	t_{stat}	4.6669	t_{stat}	-2.0116
	$pvalue$	0.0012	$pvalue$	0.0018	$pvalue$	0.0042	$pvalue$	0.0443	$pvalue$	0.4594	$pvalue$	0.0245	$pvalue$	0.0146	$pvalue$	0.0069	$pvalue$	0.8882	$pvalue$	0.0338	$pvalue$	0.0001	$pvalue$	0.0552
Motor insurance	$y = -0.8529x - 0.3266$ $R^2 = 0.2941$ $F = 10.4146$ significance $F = 0.0035$		$y = 1.009x - 0.1786$ $R^2 = 0.2311$ $F = 7.5160$ significance $F = 0.0111$		$y = -0.0687x - 0.2541$ $R^2 = 0.0048$ $F = 0.1195$ significance $F = 0.7323$		$y = -0.7576x - 0.2898$ $R^2 = 0.1517$ $F = 4.4724$ significance $F = 0.0446$		$y = -0.0611x - 0.2415$ $R^2 = 0.0007$ $F = 0.0167$ significance $F = 0.8981$		$y = 0.9116x - 0.1394$ $R^2 = 0.3524$ $F = 13.6033$ significance $F = 0.0012$													
	A_i	-0.8529	C	-0.3266	A_i	1.0090	C	-0.1786	A_i	-0.0687	C	-0.2541	A_i	-0.7576	C	-0.2898	A_i	-0.0611	C	-0.2415	A_i	0.9116	C	-0.1394
	SE	0.2643	SE	0.1316	SE	0.3680	SE	0.1361	SE	0.1987	SE	0.1593	SE	0.3583	SE	0.1431	SE	0.4729	SE	0.1549	SE	0.2472	SE	0.1262
	t_{stat}	-3.2272	t_{stat}	-2.4826	t_{stat}	2.7415	t_{stat}	-1.3127	t_{stat}	-0.3458	t_{stat}	-1.5953	t_{stat}	-2.1148	t_{stat}	-2.0242	t_{stat}	-0.1293	t_{stat}	-1.5598	t_{stat}	3.6883	t_{stat}	-1.1050
	$pvalue$	0.0035	$pvalue$	0.0201	$pvalue$	0.0111	$pvalue$	0.2012	$pvalue$	0.7324	$pvalue$	0.1232	$pvalue$	0.0446	$pvalue$	0.0538	$pvalue$	0.8982	$pvalue$	0.1314	$pvalue$	0.0011	$pvalue$	0.2797

Table 7. Hofstede's cultural dimensions regression results – summary (cont.)

Type of insurance	PDI		IDV		MAS		UAI		LTO		IVR													
Property insurance	$y = -1.5273x - 0.659$ $R^2 = 0.4418$ $F = 19.7854$ significance $F = 0.0002$		$y = 1.8551x - 0.3912$ $R^2 = 0.366$ $F = 14.4347$ significance $F = 0.0008$		$y = -0.2453x - 0.5568$ $R^2 = 0.0284$ $F = 0.7309$ significance $F = 0.4007$		$y = -1.6075x - 0.61$ $R^2 = 0.32$ $F = 11.7643$ significance $F = 0.0022$		$y = -0.2194x - 0.512$ $R^2 = 0.004$ $F = 0.1012$ significance $F = 0.7530$		$y = 1.619x - 0.3253$ $R^2 = 0.5207$ $F = 27.1646$ significance $F = 2.156E-05$													
	A_i	-1.5273	C	-0.6590	A_i	1.8551	C	-0.3912	A_i	-0.2453	C	-0.5568	A_i	-1.6075	C	-0.6100	A_i	-0.2194	C	-0.5120	A_i	1.6190	C	-0.3253
	SE	0.3434	SE	0.1709	SE	0.4883	SE	0.1805	SE	0.2869	SE	0.2299	SE	0.4687	SE	0.1873	SE	0.6897	SE	0.2259	SE	0.3106	SE	0.1586
	t_{stat}	-4.4481	t_{stat}	-3.8555	t_{stat}	3.7993	t_{stat}	-2.1665	t_{stat}	-0.8549	t_{stat}	-2.4214	t_{stat}	-3.4299	t_{stat}	-3.2574	t_{stat}	-0.3182	t_{stat}	-2.2669	t_{stat}	5.2120	t_{stat}	-2.0511
$pvalue$	0.0002	$pvalue$	0.0007	$pvalue$	0.0008	$pvalue$	0.0400	$pvalue$	0.4007	$pvalue$	0.0230	$pvalue$	0.0021	$pvalue$	0.0032	$pvalue$	0.7530	$pvalue$	0.0323	$pvalue$	0.0000	$pvalue$	0.0509	
GL insurance	$y = -1.5705x - 0.917$ $R^2 = 0.3037$ $F = 10.9034$ significance $F = 0.0029$		$y = 2.3305x - 0.6164$ $R^2 = 0.3755$ $F = 15.0349$ significance $F = 0.0007$		$y = 0.1481x - 0.7214$ $R^2 = 0.0067$ $F = 0.1696$ significance $F = 0.6840$		$y = -1.5814x - 0.8617$ $R^2 = 0.2013$ $F = 6.3023$ significance $F = 0.0189$		$y = 0.0377x - 0.753$ $R^2 = 8e-05$ $F = 0.0019$ significance $F = 0.9652$		$y = 1.8764x - 0.5508$ $R^2 = 0.4547$ $F = 20.8470$ significance $F = 0.0001$													
	A_i	-1.5705	C	-0.9170	A_i	2.3305	C	-0.6164	A_i	0.1481	C	-0.7214	A_i	-1.5814	C	-0.8617	A_i	0.0377	C	-0.7530	A_i	1.8764	C	-0.5508
	SE	0.4756	SE	0.2368	SE	0.6010	SE	0.2222	SE	0.3597	SE	0.2883	SE	0.6299	SE	0.2517	SE	0.8572	SE	0.2807	SE	0.4110	SE	0.2098
	t_{stat}	-3.3020	t_{stat}	-3.8729	t_{stat}	3.8775	t_{stat}	-2.7738	t_{stat}	0.4118	t_{stat}	-2.5019	t_{stat}	-2.5104	t_{stat}	-3.4237	t_{stat}	0.0440	t_{stat}	-2.6828	t_{stat}	4.5659	t_{stat}	-2.6254
$pvalue$	0.0029	$pvalue$	0.0007	$pvalue$	0.0007	$pvalue$	0.0103	$pvalue$	0.6840	$pvalue$	0.0193	$pvalue$	0.0189	$pvalue$	0.0021	$pvalue$	0.9653	$pvalue$	0.0128	$pvalue$	0.0001	$pvalue$	0.0146	
Health insurance	$y = -1.7739x - 1.68$ $R^2 = 0.2091$ $F = 6.6091$ significance $F = 0.0164$		$y = 1.9757x - 1.3795$ $R^2 = 0.1457$ $F = 4.2628$ significance $F = 0.0495$		$y = -0.521x - 1.6146$ $R^2 = 0.045$ $F = 1.1774$ significance $F = 0.2882$		$y = -1.6379x - 1.6076$ $R^2 = 0.1166$ $F = 3.2985$ significance $F = 0.0814$		$y = -0.121x - 1.5028$ $R^2 = 0.0004$ $F = 0.0108$ significance $F = 0.9182$		$y = 2.5018x - 1.2248$ $R^2 = 0.4362$ $F = 19.3454$ significance $F = 0.0002$													
	A_i	-1.7739	C	-1.6800	A_i	1.9757	C	-1.3795	A_i	-0.5210	C	-1.6146	A_i	-1.6379	C	-1.6076	A_i	-0.1210	C	-1.5028	A_i	2.5018	C	-1.2248
	SE	0.6900	SE	0.3435	SE	0.9569	SE	0.3538	SE	0.4802	SE	0.3849	SE	0.9018	SE	0.3603	SE	1.1666	SE	0.3820	SE	0.5688	SE	0.2904
	t_{stat}	-2.5708	t_{stat}	-4.8908	t_{stat}	2.0646	t_{stat}	-3.8987	t_{stat}	-1.0851	t_{stat}	-4.1952	t_{stat}	-1.8162	t_{stat}	-4.4612	t_{stat}	-0.1037	t_{stat}	-3.9338	t_{stat}	4.3983	t_{stat}	-4.2180
$pvalue$	0.0165	$pvalue$	0.0000	$pvalue$	0.0495	$pvalue$	0.0006	$pvalue$	0.2882	$pvalue$	0.0003	$pvalue$	0.0814	$pvalue$	0.0002	$pvalue$	0.9182	$pvalue$	0.0006	$pvalue$	0.0002	$pvalue$	0.0003	

Note. The colours used in the table correspond to the level of the R^2 coefficient and thus the level of the explanation of the dependencies between the analysed variables: $R^2 \in [0,0-0,1)$ – no dependencies, grey colour; $[0,1-0,4)$ – weak explanation, green colour; $[0,4-0,6)$ – clear explanation, blue colour. SE – standard error.

Source: authors' work based on data from Tables 5 and 6.

Based on Figures 1–7 and Table 7, it can be concluded that:

- the PDI, IDV and UAI dimensions allow for a weak explanation of insurance demand measured as insurance density in constant exchange;
- LTO does not allow for an explanation of insurance demand measured as insurance density in the constant exchange rate;
- the greatest importance (the highest R^2) allows for an explanation of insurance demand measured as insurance density in constant exchange from Hofstede's cultural dimensions; the relationship between the IVR factor and the insurance density turned out to be the most significant;
- finally, the results achieved for various insurance groups as well as for insurance in general are very similar: in each case, the significance of distinct Hofstede's cultural factors is very similar.

Summary of the linear regression results, according to equation (2), is presented in Table 8 for all analysed types of insurance.

Table 8. Results of Hofstede's cultural dimensions regression models based on equation 2

Specification	All insurance	Life insurance	P&C insurance	Motor insurance	Property insurance	GL insurance	Health insurance
Statistically significant explanatory variables (after the last iteration of model estimation)	IVR IDV	IVR IDV	IVR IDV	IVR	IVR IDV	IVR IDV	IVR
Evaluation of model parameters							
A_1 (IVR)	1.8017	2.3841	1.2852	0.9697	1.4073	1.6382	3.0129
Standard error A_1	0.3801	0.5230	0.3115	0.2425	0.2960	0.4135	0.7130
t_{stat}	4.74061	4.55833	4.12546	3.99892	4.75412	3.96166	4.22594
p_{value}	0.00007	0.00012	0.00036	0.00047	0.00007	0.00055	0.00026
If $p_{value} < 0,05?$	yes	yes	yes	yes	yes	yes	yes
A_2 (IDV)	1.5052	2.1226	1.0442	–	1.3177	1.7748	–
Standard error A_2	0.5247	0.7221	0.4301	–	0.4087	0.5709	–
t_{stat}	2.8687	2.9396	2.4277	–	3.2242	3.1089	–
p_{value}	0.0083	0.0070	0.0227	–	0.0035	0.0046	–
If $p_{value} < 0,05?$	yes	yes	yes	–	yes	yes	–
Model quality assessment							
R^2	0.6519	0.6437	0.5821	0.3808	0.6706	0.6128	0.4072
Adjusted R^2	0.5979	0.5894	0.5254	0.3424	0.6175	0.5574	0.3687
Model standard error	0.94	1.29	0.77	0.64	0.73	1.02	1.89
F-value for the variance test	23.41	22.58	17.41	15.99	25.45	19.79	17.86
Significance of F	0.00000	0.00000	0.00002	0.00050	0.00000	0.00001	0.00028
If significance of F < 0,05? ...	yes	yes	yes	yes	yes	yes	yes
Residuals randomness based on the series test	yes	yes	yes	yes	yes	yes	yes

Note. The explanatory variables used in the first iteration of model estimation: PDI, IDV, MAS, UAI, LTO, IVR. Constant: No ($C = 0$). Dependency type: linear.

Source: authors' work based on data from Tables 5 and 6.

The conclusions drawn from the combined models (presented in Table 8) are, in principle, very similar to those drawn from Figures 1–7 and Table 7:

- the IVR factor shows the greatest importance (statistical significance observed in each model) for an explanation of insurance demand measured as insurance density in constant exchange from Hofstede's cultural dimensions;
- IDV proves second most important Hofstede cultural dimensions index (statistical significance noted in five out of seven econometric models). However, from a statistical point of view, the impact of IDV is much more limited than that of the IVR factor.

The results shown in Table 8 have significant methodological and practical implications in the field of cross-cultural insurance research. Most importantly, this study sheds light on the sometimes disregarded, yet crucial significance of the IVR cultural factor in clarifying the dynamics of insurance demand across various types of insurance. This analytical addition highlights the significance of IVR a factor that has hitherto been given very limited consideration in the current body of literature. Moreover, the consistent importance of IVR across many insurance domains highlights the strength and generalisability of our results. The findings of this study suggest a strategic necessity for insurance businesses to align their product portfolios and marketing communication with the cultural norms upheld by indulgent civilizations. These principles prioritise not only the reduction of risk but also the improvement of quality of life, personal satisfaction and the quest for pleasure. The implementation of strategic alignment has the capacity to enhance market competitiveness and foster customer resonance within culturally varied environments. What is more, the aforementioned findings are supported by previous research, particularly the investigations conducted by Chui and Kwok (2009), Park and Lemaire (2011a, 2011b), and Treerattanapun (2011), which further strengthen the empirical basis of these conclusions. The study's adherence to the framework of Hofstede's cultural aspects confirms its theoretical foundation, deepening our understanding of the complex interaction between culture and insurance behaviour. In brief, this study contributes to the existing empirical information and enhances the theoretical underpinnings in the field of cross-cultural insurance research. It provides practical insights for the insurance sector stakeholders who aim to achieve relevance and competitiveness in varied worldwide markets.

How could these results be interpreted and what are the practical implications of this finding? An indulgent society is one that permits its citizens to freely satisfy their basic human wants to enjoy life (Hofstede, 2001). It is important to acknowledge that leisure preferences vary across societies, irrespective of their classification as either indulgent or constrained. The perception of leisure within society is

characterised by intricate and diverse elements, shaped by a confluence of economic, historical and individual determinants. In the field of insurance marketing, being aware of the complexities associated with the value of leisure time is crucial. Instead of assuming a direct correlation between cultural characteristics and the valuation of leisure, it is more precise to examine the intersection of cultural values with broader society-related trends and economic situations. Insurance firms must customise their approaches to the distinct expressions of cultural values within the area of leisure, acknowledging that these values might be intricate and influenced by several factors. Moreover, in the context of insurance marketing, the understanding of the subtleties of freedom of expression within different cultural and political contexts becomes crucial. Insurance companies should be mindful of the diversity of perspectives within societies and avoid presumptions that certain political affiliations or cultural orientations uniformly dictate attitudes towards freedom of expression. The vast majority of leisure activities (i.e. summer and winter sports, water sports, extreme sports, hiking, etc.) cause significant health risks as well as potential personal liability. Moreover, indulgent societies in general perceive longevity and well-being as the most important values. For this reason, insurance companies could tailor their offering (different types of insurance) to the citizens of the countries with a high IVR. This may be of practical importance for the marketing communication of insurance companies which should consider the possibility of enjoying life and maintaining current standards of living in the case of unforeseen events covered by insurance.

Although this study offers interesting insights into the correlation between cultural aspects and the need for insurance, it is also important to acknowledge its limitations. The findings presented in this paper are contingent upon a particular dataset and sample size, thus constraining the possibility to extend the conclusions to a broader population. It is recommended that future investigations based on this work employ datasets that are both larger in scale and more diversified in nature. Furthermore, cultural aspects, while serving a practical purpose, offer a streamlined framework for comprehending the complexities of culture. The use of qualitative research approaches and the incorporation of other cultural frameworks might provide a more detailed and comprehensive viewpoint. The present study is based on cross-sectional data, which poses difficulties in demonstrating causality. Longitudinal research and experimental methods might potentially offer a means to overcome this constraint. Moreover, it is important to delve into the issue of cultural variety among nations and the possible variances that exist within subcultures. This calls for further investigation in future study, which might be effectively conducted through the use of mixed-methods techniques. It is essential to acknowledge that cultural values have the potential to undergo transformations with time,

necessitating periodic reassessments of this study to account for the shifting cultural landscapes. These numerous features offer prospective avenues for advancing our understanding of the evolutionary process of cultural influences and their impact on the dynamics of insurance behaviour. The investigation of these dynamics encompasses not only the interconnectivity of diverse variables but also the analysis of the individual progression of each of these phenomena over time. It is crucial to establish clarity in order to thoroughly analyse the interaction between culture and insurance behaviour by precisely defining the distinctive characteristics of these dynamics. Ultimately, an examination of how insurance businesses modify their marketing and communication tactics to conform to cultural values can yield valuable practical insights for the sector. These pathways serve to expand the academic investigation into the impact of culture on insurance demand, enhancing the field's comprehension and resolving the constraints of this study. An important limitation of the study is also the analysis of the cultural aspect in isolation from other factors which undoubtedly influence the assessment of the impact of culture on the demand for insurance (e.g. the wealth of nations, as in Park and Lemaire (2011a)).

6. Conclusions

In the research, the relationship between Hofstede's cultural dimensions indices and insurance demand (insurance density – total premiums per inhabitant) has been analysed. The study was conducted for various types of insurance: all insurance, life insurance, P&C insurance, motor insurance, property insurance, GL insurance and health insurance.

The findings of the study emphasize the significant impact of the IVR cultural component on different insurance types. These findings have important implications for insurance policies and marketing tactics. Insurance firms have the potential to use this knowledge by strategically emphasizing the importance of enjoying life as a key value proposed in their marketing endeavours. This surpasses the conventional methods of reducing risk since it redefines insurance as a mechanism to facilitate leisure, individual satisfaction and overall well-being. Moreover, it emphasizes the significance of customising insurance offerings to correspond with the lives and tastes of affluent civilizations, providing coverage for risks and activities associated with leisure that have great value in these particular cultures. Marketing techniques that are successful should be consistent with the cultural norms of civilizations that prioritize indulgence. These strategies should highlight longevity and well-being as the most important values within the insurance industry. The research promotes the practice of focused market segmentation, utilising cultural characteristics such as

IVR, which allows insurance companies to pinpoint areas or nations with high IVR scores. This enables them to tailor their goods and communication strategies appropriately, thus achieving a competitive advantage in various markets. Ultimately, the research serves as a source of inspiration for product innovation, leading to the creation of insurance offerings that align with the preferences and concerns of individuals in these cultural contexts. These offerings encompass various features such as adaptable policies, comprehensive coverage for recreational pursuits and the integration of wellness initiatives within insurance packages. In conclusion, our findings serve to address methodological limitations and offer practical insights that can effectively improve customer interaction, broaden market reach and stimulate company expansion for insurance companies operating in culturally varied regions.

The indicated results are not only of methodological importance (little interest in this indicator is present in the literature) but also from the business perspective. The marketing communication of insurance companies should consider the possibility of positioning enjoyment of life and longevity as the key values in the context of their marketing strategy built for indulgent societies.

References

- Anand, J., & Delios, A. (1997). Location specificity and the transferability of downstream assets to foreign subsidiaries. *Journal of International Business Studies*, 28(3), 579–603. <https://doi.org/10.1057/palgrave.jibs.8490112>.
- Andrianova, O. A. (2022). Contextual Component in Cross-Cultural Business Interaction With Asian Partners. *Professional Discourse & Communication*, 4(4), 86–94. <https://doi.org/10.24833/2687-0126-2022-4-4-86-94>.
- Arora, A., & Fosfuri, A. (2000). Wholly Owned Subsidiary Versus Technology Licensing in the Worldwide Chemical Industry. *Journal of International Business Studies*, 31(4), 555–572. <https://doi.org/10.1057/palgrave.jibs.8490922>.
- Barkema, H. G., Bell, J. H. J., & Pennings, J. M. (1996). Foreign entry, cultural barriers, and learning. *Strategic Management Journal*, 17(2), 151–166. [https://doi.org/10.1002/\(SICI\)1097-0266\(199602\)17:2%3C151::AID-SMJ799%3E3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0266(199602)17:2%3C151::AID-SMJ799%3E3.0.CO;2-Z).
- Barkema, H. G., Shenkar, O., Vermeulen, F., & Bell, J. H. J. (1997). Working abroad, working with others: how firms learn to operate international joint ventures. *Academy of Management Journal*, 40(2), 426–442. <https://doi.org/10.5465/256889>.
- Barkema, H. G., & Vermeulen, F. (1997). What differences in the cultural backgrounds of partners are detrimental for international joint ventures?. *Journal of International Business Studies*, 28(4), 845–864. <https://doi.org/10.1057/palgrave.jibs.8490122>.
- Barkema, H. G., & Vermeulen, F. (1998). International expansion through start-up or acquisition: a learning perspective. *Academy of Management Journal*, 41(1), 7–26. <https://doi.org/10.5465/256894>.

- Benito, G. R. G., & Gripsrud, G. (1992). The expansion of foreign direct investments: discrete rational location choices or a cultural learning process?. *Journal of International Business Studies*, 23(3), 461–476. <https://doi.org/10.1057/palgrave.jibs.8490275>.
- Beugelsdijk, S., Kostova, T., & Roth, K. (2017). An overview of Hofstede-inspired country-level culture research in international business since 2006. *Journal of International Business Studies*, 48(1), 30–47. <https://doi.org/10.1057/s41267-016-0038-8>.
- Bond, M. H. (1991). *Beyond the Chinese face: Insights from psychology*. Oxford University Press.
- Brouthers, K. D., & Brouthers, L. E. (2000). Acquisition or Greenfield start-up? Institutional, cultural, and transaction cost influences. *Strategic Management Journal*, 21(1), 89–97. [https://doi.org/10.1002/\(SICI\)1097-0266\(200001\)21:1%3C89::AID-SMJ85%3E3.0.CO;2-8](https://doi.org/10.1002/(SICI)1097-0266(200001)21:1%3C89::AID-SMJ85%3E3.0.CO;2-8).
- Brouthers, K. D., & Brouthers, L. E. (2001). Explaining the National Cultural Distance Paradox. *Journal of International Business Studies*, 32(1), 177–189. <https://doi.org/10.1057/palgrave.jibs.8490944>.
- Chang, S. J., & Rosenzweig, P. M. (2001). The choice of entry mode in sequential foreign direct investment. *Strategic Management Journal*, 22(8), 747–776. <https://doi.org/10.1002/smj.168>.
- Chui, A. C. W., & Kwok, C. C. Y. (2009). Cultural Practices and Life Insurance Consumption: An International Analysis using GLOBE Scores. *Journal of Multinational Financial Management*, 19(4), 273–290. <https://doi.org/10.1016/j.mulfin.2009.02.001>.
- Datta, D. K., & Puia, G. (1995). Cross-border Acquisitions: An Examination of the Influence of Relatedness and Cultural Fit on Shareholder Value Creation in US Acquiring Firms. *Management International Review*, 35(4), 337–359.
- Diener, E., & Diener, M. (1995). Cross-Cultural Correlates of Life Satisfaction and Self-Esteem. *Journal of Personality and Social Psychology*, 68(4), 653–663. <https://doi.org/10.1037/0022-3514.68.4.653>.
- Diener, E., Diener, M., & Diener, C. (1995). Factors Predicting the Subjective Well-Being of Nations. *Journal of Personality and Social Psychology*, 69(5), 851–864. <https://doi.org/10.1037/0022-3514.69.5.851>.
- Diener, E., Gohm, C. L., Suh, E., & Oishi, S. (2000). Similarity of the relations between marital status and subjective well-being across cultures. *Journal of Cross-Cultural Psychology*, 31(4), 419–436. <https://doi.org/10.1177/0022022100031004001>.
- Duczkowski, N. (2022). Application of reduced social welfare functions for estimating household insurance expenditures in Poland. *Wiadomości Ubezpieczeniowe*, 1, 73–94. <https://doi.org/10.33995/wu2022.1.5>.
- Erramilli, M. K. (1991). The experience factor in foreign market entry behavior of service firms. *Journal of International Business Studies*, 22(3), 479–501. <https://doi.org/10.1057/palgrave.jibs.8490312>.
- Erramilli, M. K. (1996). Nationality and subsidiary ownership patterns in multinational corporations. *Journal of International Business Studies*, 27(2), 225–248. <https://doi.org/10.1057/palgrave.jibs.8490133>.
- Erramilli, M. K., Agarwal, S., & Dev, C. S. (2002). Choice Between Non-Equity Entry Modes: An Organizational Capability Perspective. *Journal of International Business Studies*, 33(2), 223–242. <https://doi.org/10.1057/palgrave.jibs.8491014>.

- Erramilli, M. K., Agarwal, S., & Kim, S. S. (1997). Are firm-specific advantages location-specific too?. *Journal of International Business Studies*, 28(4), 735–757. <https://doi.org/10.1057/palgrave.jibs.8490117>.
- Esho, N., Kirievsky, A., Ward, D. & Zurbruegg, R. (2004). Law and the Determinants of Property-Casualty Insurance. *The Journal of Risk and Insurance*, 71(2), 265–283. <https://doi.org/10.1111/j.0022-4367.2004.00089.x>.
- Firican, D. A. (2022). The Influence of Hofstede's Cultural Dimensions on Attitude Towards Change and Innovation in Organizations. *Proceedings of the International Conference on Business Excellence*, 16(1), 509–518. <https://doi.org/10.2478/picbe-2022-0049>.
- Franke, R. H., Hofstede, G., & Bond, M. H. (1991). Cultural roots of economic performance: a research note. *Strategic Management Journal*, 12(S1), 165–173. <https://doi.org/10.1002/smj.4250120912>.
- Geert Hofstede. (n.d.). *Dimension data matrix* [Data set]. Retrieved January 25, 2023, from <https://geerthofstede.com/research-and-vsm/dimension-data-matrix/>.
- Glaister, K. W., & Buckley, P. J. (1999). Performance Relationships in UK International Alliances. *Management International Review*, 39(2), 123–147.
- Gómez-Mejía, L. R., & Palich, L. E. (1997). Cultural diversity and the performance of multinational firms. *Journal of International Business Studies*, 28(2), 309–335. <https://doi.org/10.1057/palgrave.jibs.8490103>.
- Gong, Y., Shenkar, O., Luo, Y., & Nyaw, M. (2001). Role conflict and ambiguity of CEOs in international joint ventures: a transaction cost perspective. *Journal of Applied Psychology*, 86(4), 764–773. <https://doi.org/10.1037/0021-9010.86.4.764>.
- Green, E. D. (2016, May 12). *What are the most-cited publications in the social sciences (according to Google Scholar)?*. <https://blogs.lse.ac.uk/impactofsocialsciences/2016/05/12/what-are-the-most-cited-publications-in-the-social-sciences-according-to-google-scholar/>.
- Habib, M., & Zurawicki, L. (2002). Corruption and Foreign Direct Investment. *Journal of International Business Studies*, 33(2), 291–307. <https://doi.org/10.1057/palgrave.jibs.8491017>.
- Håkanson, L., & Nobel, R. (2001). Organizational Characteristics and Reverse Technology Transfer. *Management International Review*, 41(4), 395–420.
- Harzing, A. W. (2002). Acquisitions versus greenfield investments: international strategy and management of entry modes. *Strategic Management Journal*, 23(3), 211–227. <https://doi.org/10.1002/smj.218>.
- Hennart, J. F., & Larimo, J. (1998). The Impact of Culture on the Strategy of Multinational Enterprises: Does National Origin Affect Ownership Decisions?. *Journal of International Business Studies*, 29(3), 515–538. <https://doi.org/10.1057/palgrave.jibs.8490005>.
- Hofstede, G. (1983). National Cultures in Four Dimensions: A Research-based Theory of Cultural Differences among Nations. *International Studies of Management & Organization*, 13(1–2), 46–74. <https://doi.org/10.1080/00208825.1983.11656358>.
- Hofstede, G. (1995). Insurance as a Product of National Values. *The Geneva Papers on Risk and Insurance. Issues and Practice*, 20(77), 423–429. <https://doi.org/10.1057/gpp.1995.36>.
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviours, Institutions and Organisations Across Nations* (2nd edition). Sage Publications.

- Hofstede, G. (2013). Replicating and Extending Cross-National Value Studies: Rewards and Pitfalls – An Example from Middle East Studies. *AIB Insights*, 13(2), 5–7. <https://doi.org/10.46697/001c.16952>.
- Husted, B. W. (1999). Wealth, Culture, and Corruption. *Journal of International Business Studies*, 30(2), 339–359. <https://doi.org/10.1057/palgrave.jibs.8490073>.
- Insurance Europe. (n.d.). *Statistics*. Retrieved January 25, 2023, from <https://www.insuranceeurope.eu/statistics>.
- Jaspersen, J. G. (2016). Hypothetical Surveys and Experimental Studies of Insurance Demand: A Review. *The Journal of Risk and Insurance*, 83(1), 217–255. <https://doi.org/10.1111/jori.12100>.
- Kallunki, J. P., Larimo, J., & Pynnönen, S. (2001). Value Creation in Foreign Direct Investments. *Management International Review*, 41(4), 357–376.
- Kim, W. C., & Hwang, P. (1992). Global strategy and multinationals' entry mode choice. *Journal of International Business Studies*, 23(1), 29–53. <https://doi.org/10.1057/palgrave.jibs.8490258>.
- Kirkman, B. L., Lowe, K. B., & Gibson, C. B. (2006). A quarter century of *Culture's Consequences*: a review of empirical research incorporating Hofstede's cultural values framework. *Journal of International Business Studies*, 37(3), 285–320. <https://doi.org/10.1057/palgrave.jibs.8400202>.
- Kogut, B., & Singh, H. (1988). The effect of national culture on the choice of entry mode. *Journal of International Business Studies*, 19(3), 411–432. <https://doi.org/10.1057/palgrave.jibs.8490394>.
- Kostro, K. (2009). Zagadnienia kulturowe w ekonomii. *Gospodarka Narodowa. The Polish Journal of Economics*, 230(3), 27–59. <https://doi.org/10.33119/GN/101228>.
- Krug, J. A., & Nigh, D. (1998). Top management departures in cross-border acquisitions: governance issues in an international context. *Journal of International Management*, 4(4), 267–287. [https://doi.org/10.1016/S1075-4253\(98\)00015-5](https://doi.org/10.1016/S1075-4253(98)00015-5).
- Leonavičienė, E., & Burinskienė, A. (2022). Accelerating Cultural Dimensions at International Companies in the Evidence of Internationalisation. *Sustainability*, 14(3), 1–19. <https://doi.org/10.3390/su14031524>.
- Li, J., & Guisinger, S. (1991). Comparative business failures of foreign-controlled firms in the United States. *Journal of International Business Studies*, 22(2), 209–224. <https://doi.org/10.1057/palgrave.jibs.8490300>.
- Li, J., & Guisinger, S. (1992). The globalization of service multinationals in the “triad” regions: Japan, Western Europe, and North America. *Journal of International Business Studies*, 23(4), 675–696. <https://doi.org/10.1057/palgrave.jibs.8490283>.
- Loree, D. W., & Guisinger, S. E. (1995). Policy and non-policy determinants of U.S. equity foreign direct investment. *Journal of International Business Studies*, 26(2), 281–299. <https://doi.org/10.1057/palgrave.jibs.8490174>.
- Luo, Y. (2001a). Antecedents and Consequences of Personal Attachment in Cross-cultural Cooperative Ventures. *Administrative Science Quarterly*, 46(2), 177–201. <https://doi.org/10.2307/2667085>.
- Luo, Y. (2001b). Determinants of local responsiveness: perspectives from foreign subsidiaries in an emerging market. *Journal of Management*, 27(4), 451–477. <https://doi.org/10.1177/014920630102700404>.

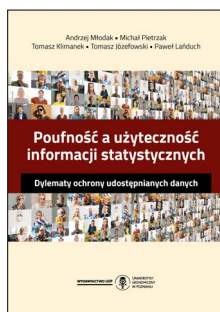
- Luo, Y., & Park, S. H. (2001). Strategic alignment and performance of market-seeking MNCs in China. *Strategic Management Journal*, 22(2), 141–155. [https://doi.org/10.1002/1097-0266\(200101\)22:2<141::AID-SMJ151>3.0.CO;2-O](https://doi.org/10.1002/1097-0266(200101)22:2<141::AID-SMJ151>3.0.CO;2-O).
- Luo, Y., & Peng, M. W. (1999). Learning to Compete in a Transition Economy: Experience, Environment, and Performance. *Journal of International Business Studies*, 30(2), 269–296. <https://doi.org/10.1057/palgrave.jibs.8490070>.
- Luo, Y., Shenkar, O., & Nyaw, M. K. (2001). A Dual Parent Perspective on Control and Performance in International Joint Ventures: Lessons From a Developing Economy. *Journal of International Business Studies*, 32(1), 41–58. <https://doi.org/10.1057/palgrave.jibs.8490937>.
- Merchant, H., & Schendel, D. (2000). How do international joint ventures create shareholder value?. *Strategic Management Journal*, 21(7), 723–737. [https://doi.org/10.1002/1097-0266\(200007\)21:7<723::AID-SMJ114>3.0.CO;2-H](https://doi.org/10.1002/1097-0266(200007)21:7<723::AID-SMJ114>3.0.CO;2-H).
- Nachum, L. (2003). Does nationality of ownership make any difference and, if so, under what circumstances? Professional service MNEs in global competition. *Journal of International Management*, 9(1), 1–32. [https://doi.org/10.1016/S1075-4253\(03\)00005-X](https://doi.org/10.1016/S1075-4253(03)00005-X).
- Newman, K. L., & Nollen, S. D. (1996). Culture and congruence: the fit between management practices and national culture. *Journal of International Business Studies*, 27(4), 753–779. <https://doi.org/10.1057/palgrave.jibs.8490152>.
- Offerman, L. R., & Hellmann, P. S. (1997). Culture's consequences for leadership behavior: National values in action. *Journal of Cross-Cultural Psychology*, 28(3), 342–351. <https://doi.org/10.1177/0022022197283008>.
- Outreville, J. F. (2013). The Relationship Between Insurance and Economic Development: 85 Empirical Papers for a Review of the Literature. *Risk Management and Insurance Review*, 16(1), 71–122. <https://doi.org/10.1111/j.1540-6296.2012.01219.x>.
- Oyserman, D., Coon, H. M., & Kimmelmeier, M. (2002). Rethinking Individualism and Collectivism: Evaluation of Theoretical Assumptions and Meta-Analyses. *Psychological Bulletin*, 128(1), 3–72. <https://doi.org/10.1037/0033-2909.128.1.3>.
- Pan, Y. (1996). Influences on foreign equity ownership level in joint ventures in China. *Journal of International Business Studies*, 27(1), 1–26. <https://doi.org/10.1057/palgrave.jibs.8490123>.
- Pan, Y. (2002). Equity Ownership in International Joint Ventures: The Impact of Source Country Factors. *Journal of International Business Studies*, 33(2), 375–384. <https://doi.org/10.1057/palgrave.jibs.8491022>.
- Park, H., Borde, S. F., & Choi, Y. (2002). Determinants of insurance pervasiveness: a cross-national analysis. *International Business Review*, 11(1), 79–96. [https://doi.org/10.1016/S0969-5931\(01\)00048-8](https://doi.org/10.1016/S0969-5931(01)00048-8).
- Park, S. C., & Lemaire, J. (2011a). *The Impact of Culture on the Demand for Non-Life Insurance* (Wharton School IRM Working Paper No. 2). <https://doi.org/10.2143/AST.42.2.2182806>.
- Park, S. C., & Lemaire, J. (2011b). Culture Matters: Long-Term Orientation and the Demand for Life Insurance. *Asia-Pacific Journal of Risk and Insurance*, 5(2), 1–23. <https://doi.org/10.2202/2153-3792.1105>.
- Park, S. H., & Ungson, G. R. (1997). The effect of national culture, organizational complementarity, and economic motivation on joint venture dissolution. *Academy of Management Journal*, 40(2), 279–307. <https://doi.org/10.5465/256884>.

- Peterson, M. F., & Smith, P. B. (1997). Does national culture or ambient temperature explain cross-national differences in role stress? No sweat!. *Academy of Management Journal*, 40(4), 930–946. <https://doi.org/10.2307/256953>.
- Pothukuchi, V., Damanpour, F., Choi, J., Chen, C. C., & Park, S. H. (2002). National and Organizational Culture Differences and International Joint Venture Performance. *Journal of International Business Studies*, 33(2), 243–265. <https://doi.org/10.1057/palgrave.jibs.8491015>.
- Riahi-Belkaoui, A. (1998). Cultural determinism and systematic risk of global stock exchanges. *International Journal of Commerce and Management*, 8(3/4), 102–108. <https://doi.org/10.1108/eb047377>.
- Richards, M., & De Carolis, D. M. (2003). Joint venture research and development activity: an analysis of the international biotechnology industry. *Journal of International Management*, 9(1), 33–49. [https://doi.org/10.1016/S1075-4253\(03\)00002-4](https://doi.org/10.1016/S1075-4253(03)00002-4).
- Robie, C., Ryan, A. M., Schmieder, R. A., Parra, L. F., & Smith, P. C. (1998). The Relation Between Job Level and Job Satisfaction. *Group and Organization Management*, 23(4), 470–495. <https://doi.org/10.1177/1059601198234007>.
- Roth, K., & O'Donnell, S. (1996). Foreign subsidiary compensation strategy: an agency theory perspective. *Academy of Management Journal*, 39(3), 678–703. <https://doi.org/10.2307/256659>.
- Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: nation and culture as explanations for variability in practice. *Personnel Psychology*, 52(2), 359–391. <https://doi.org/10.1111/j.1744-6570.1999.tb00165.x>.
- Salter, S. B., & Niswander, F. (1995). Cultural influence on the development of accounting systems internationally: a test of Gray's theory. *Journal of International Business Studies*, 26(2), 379–397. <https://doi.org/10.1057/palgrave.jibs.8490179>.
- Schneider, S. C., & De Meyer, A. (1991). Interpreting and responding to strategic issues: the impact of national culture. *Strategic Management Journal*, 12(4), 307–320. <https://doi.org/10.1002/smj.4250120406>.
- Seidenfuss, K. U., & Storm, J. (2022). Adapting project management methodologies for different cultures: teams, toolboxes and Hofstede. *Journal of Modern Project Management*, 10(1), 110–125. <https://doi.org/10.19255/JMPM02907>.
- Shane, S. A. (1992). The effect of cultural differences in perceptions of transaction costs on national differences in the preference for licensing. *Management International Review*, 32(4), 295–311. <https://doi.org/10.5465/ambpp.1992.17515164>.
- Shane, S. (1994). The effect of national culture on the choice between licensing and direct foreign investment. *Strategic Management Journal*, 15(8), 627–642. <https://doi.org/10.1002/smj.4250150805>.
- Shane, S., Venkataraman, S., & MacMillan, I. (1995). Cultural Differences in Innovation Championing Strategie. *Journal of Management*, 21(5), 931–952. <https://doi.org/10.1177/014920639502100507>.
- Shen, L. (2023). Culture and Explicitness of Persuasion: Linguistic Evidence From a 51-Year Corpus-Based Cross-Cultural Comparison of the United Nations General Debate Speeches Across 55 Countries (1970–2020). *Cross-Cultural Research*, 57(2–3), 166–192. <https://doi.org/10.1177/10693971221139523>.
- Smith, P. B., Dugan, S., Peterson, M. F., & Leung, K. (1998). Individualism: Collectivism and the handling of disagreement. A 23 country study. *International Journal of Intercultural Relations*, 22(3), 351–367. [https://doi.org/10.1016/S0147-1767\(98\)00012-1](https://doi.org/10.1016/S0147-1767(98)00012-1).

- Śliwiński, A. (2016). Popyt na ubezpieczenia na życie: przegląd badań światowych. In S. Nowak, A. Z. Nowak, & A. Sopoćko (Eds.), *Polski rynek ubezpieczeń na tle kryzysów społeczno-gospodarczych*, (pp. 110–125). Wydawnictwo Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego. http://www.wz.uw.edu.pl/portaleFiles/6133-wydawnictwo-/Polski_rynek_ubezpieczen_e-book.pdf.
- Śliwiński, A. (2019). *Rola ubezpieczeń w gospodarce*. Oficyna Wydawnicza SGH.
- Taras, V., Kirkman, B. L., & Steel, P. (2010). Examining the impact of culture's consequences: A three-decade, multilevel, meta-analytic review of Hofstede's cultural value dimensions. *Journal of Applied Psychology*, 95(3), 405–439. <https://doi.org/10.1037/a0018938>.
- Thomas, A. S., & Mueller, S. L. (2000). A Case for Comparative Entrepreneurship: Assessing the Relevance of Culture. *Journal of International Business Studies*, 31(2), 287–301. <https://doi.org/10.1057/palgrave.jibs.8490906>.
- Treerattanapun, A. (2011). The Impact of Culture on Non-Life Insurance Consumption. *Wharton Research Scholars*, 78. https://repository.upenn.edu/wharton_research_scholars/78/.
- Van de Vliert, E., & Van Yperen, N. W. (1996). Why cross-national differences in role overload? Don't overlook ambient temperature!. *Academy of Management Journal*, 39(4), 986–1004. <https://doi.org/10.5465/256720>.
- Zhou, Y., & Kwon, J.-W. (2020). Overview of Hofstede-Inspired Research Over the Past 40 Years: The Network Diversity Perspective. *SAGE Open*, 10(3). <https://doi.org/10.1177/2158244020947425>.
- Zietz, E. N. (2003). An Examination of Demand for Life Insurance. *Risk Management and Insurance Review*, 6(2), 159–191. <https://doi.org/10.1046/J.1098-1616.2003.030.x>.

Recenzja monografii naukowej Andrzeja Młodaka, Michała Pietrzaka, Tomasza Klimanka, Tomasza Józefowskiego i Pawła Lańducha *Poufność a użyteczność informacji statystycznych. Dylematy ochrony udostępnianych danych*

Review of a scientific monograph by Andrzej Młodak, Michał Pietrzak, Tomasz Klimanek, Tomasz Józefowski and Paweł Lańduch, entitled *Confidentiality vs. usefulness of statistical information. Dilemmas in protecting shared data*



Język/Language: polski/Polish

Wydawnictwo/Publisher: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu

Miejsce i rok wydania / Place and year of publication: Poznań 2023

Liczba stron / Number of pages: 276

Monografia Andrzeja Młodaka, Michała Pietrzaka, Tomasza Klimanka, Tomasza Józefowskiego i Pawła Lańducha *Poufność a użyteczność informacji statystycznych. Dylematy ochrony udostępnianych danych* jest pierwszym opracowaniem w polskiej literaturze naukowej, w którym wyczerpująco przedstawiono problematykę ochrony danych i znaczenie poufności danych statystycznych dla jakości i użyteczności badań statystycznych. Wcześniejsze publikacje na ten temat miały charakter analiz przyczynkowych dotyczących wybranych aspektów tajemnicy statystycznej, kontroli i zapewnienia poufności danych objętych tajemnicą statystyczną, a także wpływu konieczności zapewnienia poufności danych na możliwość publikowania i upowszechniania wyników danych statystycznych.

Autorzy recenzowanej książki dokładnie omówili m.in. aktualne akty prawne międzynarodowe (opublikowane przez Organizację Narodów Zjednoczonych, Unię Europejską i Międzynarodowy Instytut Statystyczny), krajowe i zagraniczne (stworzone w krajach przodujących w kształtowaniu zasad i metod badań statystycznych), a także dokumenty administracyjne istotne dla przedstawienia dylematów związanych z ochroną i udostępnianiem danych, przed którymi stają statystycy i badacze. Szczególną wartość monografii stanowi kompleksowe ujęcie zagadnień tajemnicy

statystycznej, w tym podkreślenie konieczności przestrzegania jej w konkretnych sytuacjach badawczych, a także przedstawienie konsekwencji naruszania zasad ochrony poufnych danych dla jakości badań oraz metod i narzędzi służących do kontroli i ochrony poufności. Publikacja dostarcza informacji niezbędnych do tego, aby podmiot prowadzący badania statystyczne przestrzegał etycznych i prawnych zasad ochrony danych, a respondent miał świadomość, że dane, które przekazuje, zostaną wykorzystane wyłącznie do celów określonych w badaniu statystycznym, a jego tożsamość nie zostanie ujawniona.

Strukturę tekstu można określić jako wzorową. Praca składa się z sześciu rozdziałów. W pierwszym przedstawiono koncepcję ochrony danych i wyjaśniono podstawowe pojęcia z tego zakresu. W drugim omówiono metody pomiaru i oceny ryzyka ujawnienia informacji poufnych. W trzecim zapoznano czytelników z metodami i technikami kontroli ujawniania informacji poufnych w danych wynikowych. Rozdział czwarty – szczególnie cenny dla metodologów i projektantów systemów informatycznych w badaniach statystycznych – poświęcono zagadnieniu straty informacji jako konsekwencji stosowania ochrony danych poufnych. Należy zwrócić uwagę, że kwestia identyfikacji strat informacji ma duże znaczenie praktyczne, stanowią one bowiem czynnik skłaniający badaczy do podejmowania większego ryzyka ujawnienia danych, które powinny być szczególnie chronione. Rozdział piąty – o innym charakterze niż poprzednie – zawiera przegląd wybranych narzędzi informatycznych służących do kontroli bezpieczeństwa i ochrony danych uznanych za poufne. Jest adresowany do informatyków specjalizujących się w projektowaniu modułów statystycznych systemów informatycznych, które są wykorzystywane do kontroli bezpieczeństwa i ochrony danych poufnych w konkretnych badaniach statystycznych. W rozdziale szóstym – zawierającym informacje istotne dla statystyków projektujących i organizujących badania statystyczne – opisano praktyczne doświadczenia z organizacji kontroli dostępu do danych w różnych krajach. Ma to na celu ułatwienie zainteresowanym podmiotom wyboru rozwiązań odpowiadających potrzebom konkretnych instytucji przeprowadzających badania statystyczne albo dostosowania ich do własnych potrzeb.

Za ważne części publikacji należy uznać także obszerną bibliografię – przedstawiającą cenne pozycje polskich i zagranicznych autorów – która z pewnością okaże się przydatna naukowcom, studentom i praktykom przeprowadzającym badania statystyczne, a także słownik pojęć – pierwszy tego typu glosariusz, który spełni zapewne ważną funkcję w procesie wprowadzania i upowszechnienia terminologii związanej z ochroną danych.

Warto podkreślić staranne opracowanie redakcyjne wszystkich rozdziałów monografii. Problematyka bezpieczeństwa i ochrony informacji uznanych za poufne w badaniach statystycznych jest skomplikowana, wieloaspektowa i słabo znana. Opisanie jej w sposób zwięzły i zrozumiały dla różnych odbiorców nie jest łatwe. Udało się to autorom recenzowanej monografii, choć należy zaznaczyć, że wiele jej fragmentów może okazać się trudne w odbiorze dla czytelników o różnym przygotowaniu statystycznym i informatycznym – ich pełne zrozumienie wymaga sięgnięcia do źródeł podanych w przypisach. Trzeba zatem przyjąć, że niektóre fragmenty książki są raczej przewodnikiem po problemach niż kompletnym tekstem podręcznikowym.

Publikacja jest skierowana do wielu grup odbiorców: naukowców przeprowadzających badania statystyczne lub wykorzystujących dane statystyczne w ramach innych badań, nauczycieli akademickich prowadzących zajęcia ze statystyki i z ekonometrii, dydaktyków prowadzących szkolenia ze statystyki dla pracowników administracji publicznej i przedsiębiorstw organizujących własne badania i analizy statystyczne, pracowników statystyki publicznej, statystyków realizujących komercyjne badania statystyczne (w tym badania rynku i opinii publicznej), informatyków projektujących systemy statystyczne (w szczególności bazy danych) i administrujących nimi, doktorantów i studentów, pracowników administracji publicznej oraz banków i innych instytucji finansowych prowadzących własne badania statystyczne lub zlecających je jednostkom zewnętrznym, sędziów i prokuratorów, gestorów i administratorów systemów i baz danych administracyjnych, transakcyjnych i monitoringowych (big data), polityków pragnących korzystać w sposób rzetelny z danych i badań statystycznych, popularyzatorów badań naukowych oraz publicystów ekonomicznych.

Omawiana monografia wypełnia lukę na polskim rynku wydawniczym w zakresie problematyki dotyczącej tajemnicy statystycznej, jednej z nielicznych bezwzględnych tajemnic z mocy prawa, której przestrzeganie jest gwarantem jakości i możliwości prowadzenia badań statystycznych – zwłaszcza w naukach społecznych – a także praktycznej działalności gospodarczej, społecznej i politycznej. Warto podkreślić, że to pierwsza pozycja w polskiej literaturze naukowej obejmująca wszystkie najważniejsze aspekty tajemnicy statystycznej, podstawy metodyczne i prawne oraz narzędzia zapewniające jej przestrzeganie. Dzięki kompleksowemu ujęciu problemu zarówno pod względem metodologicznym, jak i informatycznym publikacja nie

tylko dostarczy potencjalnym zainteresowanym gruntownej wiedzy na temat poufności informacji w badaniach statystycznych, lecz także okaże się pomocna w ochronie danych źródłowych i wynikowych.

Józef Oleński

Polskie Towarzystwo Statystyczne; Urząd Statystyczny w Rzeszowie, Ośrodek Badań Obszarów Transgranicznych i Statystyki Euroregionalnej, Polska
Polish Statistical Association; Statistical Office in Rzeszów, Centre for Cross-border Areas Research and Euroregion Statistics, Poland

Jacek Kowalewski

Uniwersytet Ekonomiczny w Poznaniu, Instytut Informatyki i Ekonomii Ilościowej;
Urząd Statystyczny w Poznaniu, Polska
Poznań University of Economics and Business, Institute of Informatics
and Quantitative Economics; Statistical Office in Poznań, Poland

WYDAWNICTWA GUS. LUTY 2024 PUBLICATIONS OF STATISTICS POLAND. FEBRUARY 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:

Ludność na rynku pracy w świetle wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2021 – wersja uaktualniona ***Population in the labour market according to the results of the National Census of Population and Housing 2021 – updated version***

Zaktualizowane wydanie opracowania tematycznego przedstawiającego wyniki Narodowego Spisu Powszechnego Ludności i Mieszkań 2021.



Język: polski (przedmowa, spis treści, synteza, uwagi metodologiczne, tablice i wykresy również w języku angielskim)
Language: Polish (preface, contents, executive summary, methodological notes, tables and charts available also in English)
Seria: Spisy powszechne
Series: Censuses
Dostępne wersje: elektroniczna z tablicami w formacie Excel
Available in: electronic form with Excel tables

Ukazanie się uaktualnionej wersji publikacji wynika z ustalenia – dzięki zastosowaniu nowych rozwiązań analitycznych – poziomu wykształcenia ponad 700 tys. respondentów, dla których pierwotnie nie został on określony. Tym samym w opracowaniu podano informacje dotyczące aktywności ekonomicznej ludności według poziomu wykształcenia dla znacznie szerszej zbiorowości. Autorzy publikacji wychodzą naprzeciw oczekiwaniom odbiorców, którzy chcą uzyskać precyzyjny obraz rynku pracy, z uwzględnieniem zmian, które na nim zaszły. Co ważne, informacje zebrane podczas spisu mogą stanowić podstawę do przeprowadzenia pogłębionych analiz i dokonania bilansu zmian.

Opracowanie składa się z uwag ogólnych, części analitycznej (z mapami i wykresami), w której porównano przejawy aktywności ekonomicznej ludności w latach 2011–2021, oraz z części tabelarycznej.

W lutym br. ukazały się ponadto:

- „Biuletyn statystyczny” nr 1/2024;
- *Ceny robót budowlano-montażowych i obiektów budowlanych (grudzień 2023 r.)*;
- *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach 2000–2024 (luty 2024)*;
- *Produkcja ważniejszych wyrobów przemysłowych w styczniu 2024 r.*;
- *Przedsiębiorstwa niefinansowe powstałe w latach 2018–2022*;
- *Rocznik Statystyczny Przemysłu 2023*;
- *Struktura wynagrodzeń według zawodów za październik 2022 r.*;
- *Sytuacja społeczno-gospodarcza kraju w styczniu 2024 r.*;
- „Wiadomości Statystyczne. The Polish Statistician” nr 2/2024;
- *Zeszyt metodologiczny. Badanie budżetów gospodarstw domowych*;
- *Zeszyt metodologiczny. Produkcja roślinna*.

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 MEiN. Zgodnie z komunikatem Ministra Edukacji i Nauki z dnia 1 grudnia 2021 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych „WS” otrzymały 70 punktów.

„Wiadomości Statystyczne. The Polish Statistician” są udostępniane w następujących bazach, repozytoriach, katalogach i wyszukiwarkach: Agro, BazEkon, 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) oraz Summon.

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

1. Zgłaszanie artykułów

Prace przeznaczone do opublikowania w „WS” należy przysyłać 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 niemożności 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 oraz jego struktury i zawartości pod kątem wymogów redakcyjnych, a także zgodności tematyki z profilem czasopisma. Autor uzupełnia i poprawia artykuł stosownie do uwag redakcji, a w przypadku nieuwzględnienia danej uwagi uzasadnia swoje stanowisko. Warunkiem skierowania pracy do recenzji jest potwierdzenie oryginalności tekstu uzyskane za pomocą systemu antyplagiatoowego. W przypadku wykrycia znacznego podobieństwa do innych prac artykuł zostanie odrzucony.
2. **Ocena recenzentów**, dokonywana przez specjalistów w danej dziedzinie. Artykuł oceniają dwaj recenzenci spoza jednostki naukowej, przy której afiliowany jest autor; w przypadku pracy w języku angielskim co najmniej jeden recenzent jest afiliowany przy jednostce zagranicznej. W razie sprzecznych opinii dwóch recenzentów powoływany jest trzeci recenzent. Recenzenci kierują się kryteriami oryginalności i jakości opracowania zarówno w odniesieniu do treści, jak i formy artykułu.

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

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

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

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

4. **Opracowanie redakcyjne, autoryzacja i korekta**. Artykuł zakwalifikowany do druku jest poddawany opracowaniu merytorycznemu i językowemu. Redakcja zastrzega sobie prawo do zmiany tytułu i śródtytułów, modyfikowania tablic, wykresów i innych elementów graficznych oraz przeredagowania treści bez naruszenia zasadniczej myśli autora.

Po opracowaniu redakcyjnym artykuł jest przesyłany do autoryzacji. Tekst zatwierdzony przez autora, po składzie i łamaniu, jest poddawany korekcie i rewizji (II korekcie).

Autor dokonuje korekty autorskiej tekstu na etapie rewizji. Wykresy i inne materiały graficzne są opracowywane na podstawie plików i danych przekazanych przez autora i podawane korekcie i rewizji. Autor dokonuje ich akceptacji na etapie rewizji.

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

3. Zasady etyki publikacyjnej COPE

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

3.1. Odpowiedzialność autorów

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

Autorzy deklarują w stosownym oświadczeniu, że zgłaszany artykuł nie narusza praw autorskich osób trzecich, nie był dotychczas publikowany i nie został złożony w innym wydawnictwie oraz że jest ich oryginalnym dziełem, i określają swój wkład w opracowanie artykułu. Jeżeli doszło do zaprezentowania podobnych materiałów podczas konferencji lub

symposium naukowe, to podczas składania tekstu do publikacji w „WS” autorzy są zobowiązani poinformować o tym redakcję.

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

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

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

artykułu mogą być przekazywane przez WCN wyłącznie autorom, recenzentom, członkom RN i KR oraz wydawcy.

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

3.3. Odpowiedzialność recenzentów

1. Recenzenci przyjmują artykuł do recenzji tylko wtedy, gdy uznają, że:
 - posiadają odpowiednią wiedzę w określonej dziedzinie, aby rzetelnie ocenić pracę;
 - zgodnie z ich stanem wiedzy nie istnieje konflikt interesów w odniesieniu do autorów, przedstawionych w artykule badań i instytucji je finansujących, co potwierdzają w oświadczeniu;
 - mogą wywiązać się z terminu ustalonego przez redakcję, aby nie opóźnić publikacji.
2. Recenzenci są zobligowani do zachowania obiektywności i poufności oraz powstrzymania się od osobistej krytyki. Zawsze powinni uzasadnić swoją ocenę, przedstawiając stosowną argumentację.

3. Recenzenci powinni wskazać ważne dla wyników badań opublikowane prace, które w ich ocenie powinny zostać przywołane w ocenianym artykule.
4. W razie stwierdzenia wysokiego poziomu zbieżności treści recenzowanej pracy z innymi opublikowanymi materiałami lub podejrzenia innych przejawów nierzetelności naukowej recenzenci są zobowiązani poinformować o tym redakcję.
5. Po ukończeniu recenzji przechowywanie przesłanych przez redakcję materiałów (w jakiegokolwiek formie) oraz posługiwanie się nimi przez recenzentów jest niedozwolone.

3.4. Odpowiedzialność wydawcy

1. Materiały opublikowane w „WS” są chronione prawem autorskim.
2. Wydawca udostępni pełną treść wszystkich artykułów w internecie w trybie otwartego dostępu, tj. bezpłatnie i bez technicznych ograniczeń, od 1 stycznia 2022 r. na licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0). W przypadku artykułów zgłoszonych do „WS” od 2022 r. dozwolone jest dzielenie się artykułem (kopiowanie i rozpowszechnianie go w dowolnym medium i formacie) oraz adaptowanie go (w dowolnym celu, także komercyjnym) na warunkach określonych w tej licencji. Z pozostałych artykułów zamieszczonych w czasopiśmie można korzystać w ramach otwartego dostępu, zgodnie z ustawą o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego.
3. Wydawca deklaruje gotowość do opublikowania poprawek, wyjaśnień oraz przeprosin.

4. Wymogi redakcyjne

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

4.1. Struktura i zawartość artykułu

Wymagane elementy artykułu:

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

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

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

4.2. Przygotowanie artykułu

1. Artykuł powinien być utrzymany w formie bezosobowej.
2. Tekst należy zapisać alfabetem łacińskim. Nazwy własne, tytuły itp. oryginalnie zapisane innym alfabetem powinny być poddane transliteracji.
3. Nie należy stosować stylów; formatowanie należy ograniczyć do wymogów redakcyjnych.
4. Objętość artykułu łącznie ze streszczeniem, słowami kluczowymi, bibliografią, tablicami, wykresami i innymi materiałami graficznymi nie powinna być mniejsza niż 10 stron maszynopisu ani przekraczać 20 stron.
5. Edytor tekstu: Microsoft Word, format *.doc lub *.docx.
6. Krój czcionki:
 - Arial – tytuł, autor, 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 wycienieniach należy posłużyć się listą punktowaną z punktarami 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. \mathbf{a} , \mathbf{A} , \mathbf{w} , $\mathbf{y}(x)$, \mathbf{w}_i); macierzy – pismem prostym i pogrubionym (np. \mathbf{A} , \mathbf{a} , \mathbf{M} , $\mathbf{Y}(x)$, \mathbf{M}_i).
19. Objasnienia 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 jakiegokolwiek wątpliwości odbiorców, autor powinien wskazać stosowne źródło podawanej informacji.

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

4.3. Zasady przywoływania publikacji w treści artykułu

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

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

4.4. Przykłady opisu bibliograficznego

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

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

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

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

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

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

Pełny opis zakresu tematycznego działów: ws.stat.gov.pl/AimScope

Description of the topics covered in each section: ws.stat.gov.pl/AimScope

Studia metodologiczne / Methodological studies

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

Statystyka w praktyce / Statistics in practice

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

Studia interdyscyplinarne. Wyzwania badawcze / Interdisciplinary studies. Research challenges

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

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

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

Edukacja statystyczna / Statistical education

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

Z dziejów statystyki / From the history of statistics

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

In memoriam

- Nekrologi i artykuły wspomnieniowe

Informacje. Recenzje. Dyskusje / Discussions. Reviews. Information

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