ISSN 2392-1641
e-ISSN 2450-0097

Economics and Business Review

Volume 10 (2) 2024

CONTENTS

Editorial introduction
Monika Banaszewska, Joanna Lizińska, Konrad Sobański

ARTICLES

Proposal for a comprehensive retirement insurance solution (CRIS) to mitigate retirement risk based on theory of change
Krzysztof Łyskawa, Kamila Bielawska

Examining the performance of Shari'ah-compliant versus conventional stock indexes: A comparative analysis pre-, during, and post-COVID-19
Ahmad M. Abu-Alkheil, Nizar M. Alsharari, Walayet A. Khan, Sara R. Ramzani, Phungmayo Horam

Taxation of public pensions in European Union countries
Maciej Cieślukowski

Labour productivity in Italian regions: A gravitational model approach
Katarzyna Filipowicz, Oleksij Kelebaj, Tomasz Tokarski

Personal bankruptcy prediction using machine learning techniques
Magdalena Brygała, Tomasz Korol

Enhancing garbage fee compliance: Insights from a Slovak municipality
Anetta Caplanova, Eva Sirakovova, Estera Szakadatova

Growth prospects for the silver economy in the market segment of residential care services provided to dependent elderly people
Rafał Iwański

Poznań University of Economics and Business Press

Paper based publication

https://doi.org/10.18559/ebr.2024.2

ISSN 2392-1641
e-ISSN 2450-0097

POZNAŃ UNIVERSITY OF ECONOMICS AND BUSINESS PRESS
ul. Powstańców Wielkopolskich 16, 61-895 Poznań, Poland
phone +48 61 854 31 54, +48 61 854 31 55
https://wydawnictwo.ue.poznan.pl, e-mail: wydawnictwo@ue.poznan.pl
postal address: al. Niepodległości 10, 61-875 Poznań, Poland

Printed and bound in Poland by:
Poznań University of Economics and Business Print Shop

Circulation: 80 copies

sciendo

Economics and Business Review

Volume 10 (2) 2024

CONTENTS

Editorial introduction
Monika Banaszewska, Joanna Lizińska, Konrad Sobański 3

ARTICLES

Proposal for a comprehensive retirement insurance solution (CRIS) to mitigate retirement risk based on theory of change
Krzysztof Łyskawa, Kamila Bielawska 7

Examining the performance of Shari'ah-compliant versus conventional stock indexes: A comparative analysis pre-, during, and post-COVID-19
Ahmad M. Abu-Alkheil, Nizar M. Alsharari, Walayet A. Khan, Sara R. Ramzani, Phungmayo Horam 31

Taxation of public pensions in European Union countries
Maciej Cieślukowski 60

Labour productivity in Italian regions: A gravitational model approach
Katarzyna Filipowicz, Oleksij Kelebaj, Tomasz Tokarski 92

Personal bankruptcy prediction using machine learning techniques
Magdalena Brygała, Tomasz Korol 118

Enhancing garbage fee compliance: Insights from a Slovak municipality
Anetta Caplanova, Eva Sirakovova, Estera Szakadatova 143

Growth prospects for the silver economy in the market segment of residential care services provided to dependent elderly people
Rafał Iwański 165

sciendo Economics and Business Review Vol. 10 (2), 2024: 3–5
https://www.ebr.edu.pl

Editorial introduction

As the complexity of economic phenomena increases, so do the expectations placed on researchers in economics and finance. Research should not only be methodologically sound and contribute to the existing body of knowledge, but also provide guidance to stakeholders outside academia. The current issue of the Economics and Business Review addresses these needs, as the research findings can be useful to various groups, such as legislators, governmental bodies, financial advisors and investors. It contains seven papers by seventeen affiliated scholars from Poland, Slovakia and the USA. They use a variety of methodological approaches, both conceptual and empirical, to investigate key research questions in the fields of public economics, financial economics, labour economics and behavioural economics.

The opening article **Proposal for a comprehensive retirement insurance solution (CRIS) to mitigate retirement risk based on theory of change**, by Krzysztof Łyskawa and Kamila Bielawska, provides new insights into public policies to mitigate pension risks. This is of special importance nowadays, as population ageing is one of the most significant trends shaping socio-economic life in the 21st century. The basis of the authors' proposal is the assumption that separate pension products such as annuities and long-term care insurance, as advocated by many economists, are not the right direction for developing pension policy. Accordingly, the article presents a three-layer retirement insurance solution that can be flexibly adapted to the needs of individuals during the accumulation of funds and in retirement. The authors supplement their deliberations with a SWOT analysis for the proposed solution.

The second article, entitled **Examining the performance of Shari'ah-compliant versus conventional stock indexes: A comparative analysis pre-, during, and post-COVID-19**, is written by Ahmad M. Abu-Alkheil, Nizar M. Alsharari, Walayet A. Khan, Sara R. Ramzani and Phungmayo Horam. The study uses the Varmax procedure, stochastic dominance analysis and Johansen's co-integration approach to compare the performance of conventional indexes with alternative Islamic indexes from 2017 to 2023. The overall results indicate that conventional indexes outperform Islamic indexes, but the latter offer opportunities for diversification in global portfolios, especially in times of financial turmoil. The findings might be useful to analysts, and private and institutional investors.

Maciej Cieślukowski, in his article entitled **Taxation of public pensions in European Union countries**, extends the understanding of cross-country dif-

ferences in living conditions both at working age and after retirement. Using a multi-criteria comparative analysis and the agglomeration method, the article analyses whether taxation of wages and public pension benefits can have a significant impact on the decision to choose a country to work in the common market. The author argues that effective tax rates on wages and pensions are not important factors in deciding which country to work or retire in. The major factors that should guide an employee in the European Union are the average expenditure on net salaries and the average expenditure on net pension benefits.

The article **Labour productivity in Italian regions: A gravitational model approach**, by Katarzyna Filipowicz, Oleksij Kelebaj and Tomasz Tokarski, introduces a nuanced analysis of regional labour productivity disparities across Italy. Using a gravitational model of economic growth, which extends the classic Solow model, the authors delve into the varied economic dynamics between northern, central, and southern Italy. By calibrating the model with historical data and through rigorous numerical simulations, they illustrate how different investment rates, growth rates, and urbanisation levels might contribute to or hinder the convergence in labour productivity across regions. The study-'s findings underscore the necessity for region-specific investment strategies to balance productivity levels nationwide, particularly highlighting the need for higher investment rates in southern regions to match those of more prosperous areas The paper enriches the understanding of regional economic disparities in Italy, providing a valuable model for examining similar geographical disparities in other contexts.

The subsequent article, entitled **Personal bankruptcy prediction using machine learning techniques**, is written by Magdalena Brygała and Tomasz Korol. The study examines the usefulness of six machine learning methods, namely, support vector machine, random forest, adaptive boosting, extreme gradient boosting, light gradient boosting machine, and categorical boosting, in forecasting personal bankruptcy. The research is based on two samples of households (learning and testing) from the Survey of Consumer Finances conducted in the United States. The models include such variables as income, refusal to grant credit, delays in the repayment of liabilities, the revolving debt ratio, and the housing debt ratio. The findings of the study might be used by financial institutions to make credit decisions on consumer loans.

In **Enhancing garbage fee compliance: Insights from a Slovak municipality**, the authors Anetta Caplanova, Eva Sirakovova and Estera Szakadatova employ a randomised controlled trial to scrutinise the effectiveness of behavioural interventions on garbage fee compliance in Hlohovec, Slovakia. The research ingeniously categorizes households into three groups: one receives a leaflet promoting social norms, another a deterrence message, and a control group gets only the standard invoice. The outcomes reveal intriguing dynamics: while deterrent approaches unexpectedly increased non-compliance, social norm

prompts showed no significant behavioural change. This pivotal study sheds light on the intricate interplay between public policy communication and citizen behaviour, challenging the efficacy of conventional behavioural nudges and emphasising the importance of context-driven strategies in public finance reforms. This work significantly enriches the discourse on applying behavioural economics in governance and public administration.

The last study in this issue is written by Rafał Iwański and is titled **Growth prospects for the silver economy in the market segment of residential care services provided to dependent elderly people**. The author delves into the expanding demand for elderly care services amidst Poland's aging population. As family care capacities diminish, the silver economy emerges as a critical arena for addressing the care needs of the elderly. The paper offers a comprehensive examination of the various factors influencing the development of residential care services within Poland's silver economy. By integrating statistical and financial data from several sources, the study identifies key obstacles to growth, such as labour shortages and insufficient funding, which underscore the urgent need for innovative solutions in long-term care provisioning. Through a meticulous analysis, this research not only maps the current landscape but also projects future needs, providing insights for policymakers, investors, and service providers engaged in the silver economy.

Monika Banaszewska
Joanna Lizińska
Konrad Sobański
Lead Editors

sciendo Economics and Business Review https://www.ebr.edu.pl Vol. 10 (2), 2024: 7–30 https://doi.org/10.18559/ebr.2024.2.1008

Proposal for a comprehensive retirement insurance solution (CRIS) to mitigate retirement risk based on theory of change

Krzysztof Łyskawa[1] Kamila Bielawska[2]

Abstract

The aim of the paper is to propose a new comprehensive retirement insurance solution (CRIS) that, by offering appropriate modules, can be flexibly adapted to customers' needs during the accumulation of funds and entitlements and during retirement. Technically, the product is life-insurance-based and includes insurance for sickness and incapacity, long-term care (LTC), work activation expenses, hospital stays, and tontine and Luxembourg policies. Due to consumers' changing expectations and needs, the technical dimension of this solution is based on a three-layer insurance product in which individual parts of the protection are supplemented by several additional benefits (types of assistance) that improve the quality of life of insurance participants and allow the ongoing use of the product. The basis for considering such a new insurance product is the theory of change (ToC), which makes it possible to build considerable flexibility into such a solution. A SWOT analysis was used to position the proposed solution in relation to other insurance products and social security offered by the state.

Keywords

- modular insurance
- life insurance
- tontine
- long-term care insurance
- public-private partnership

JEL codes: H55, G22, I13, D14, G28

Article received 23 October 2023, accepted 13 June 2024.

This research was financed by the National Science Centre, Poland, within the project 'Retirement risk in the light of the forecasted changes of the demand for consumption goods and services of seniors in Poland' (UMO-2016/23/B/HS5/03768).

[1] Department of Insurance, Institute of Finance, Poznań University of Economics and Business, al. Niepodległości 10, 61-875 Poznań, krzysztof.lyskawa@ue.poznan.pl, https://orcid.org/0000-0002-7409-8624.

[2] Department of Banking and Finance, Faculty of Management, University of Gdansk, ul. Armii Krajowej 101, 81-824 Sopot, Poland, corresponding author: kamila.bielawska@ug.edu.pl, https://orcid.org/0000-0003-1558-0194.

Suggested citation: Łyskawa, K., & Bielawska, K. (2024). Proposal for a comprehensive retirement insurance solution (CRIS) to mitigate retirement risk based on theory of change. *Economics and Business Review*, *10*(2), 7–30. https://doi.org/10.18559/ebr.2024.2.1008

Introduction

Increased life expectancy and social and economic changes, along with associated public policies, mean that satisfactory solutions should be sought to meet people's retirement needs. The literature and financial market analyses have indicated (Greenwood & Vissing-Jorgensen, 2018) that insurance products should supplement financial resources in retirement. However, shortages are still evident today due to insufficient supply, as well as the demand for specific insurance products. In recent years, insurance products, especially those based on unit-linked life insurance, were criticised for their mis-selling or high costs (Gatzert et al., 2011; Gupta, 2012). Additionally, the risk of not achieving the appropriate rates of return necessary to supplement retirement benefits remains with the insured in unit-linked insurance, which often coincides with high costs deteriorating the value of capital (Bernard et al., 2017). Additionally, the lack of awareness of longevity risk means that people do not seek appropriate products that could mitigate this risk, such as annuities (Brown, Mitchell et al., 2001) or long-term care (LTC) insurance or tontines (Milevsky & Salisbury, 2015). Therefore, it became necessary to comprehensively consider the issue of securing access to benefits in cash and in kind to maintain satisfactory living standards in retirement. The threat is that the growing population of retirees will otherwise be exposed to retirement risk, defined herein as the inability to cover *individual needs* the old age (Łyskawa, 2004). The coverage of retirement risk can come either from income growth or access to services. Bielawska and Kozłowski (2024), based on the subjective approach to the retirement risk proposed by Łyskawa (2004), operationalized the retirement risk and measured it for Polish retirees. The results of their research indicate a significant share among retirees', households for whom the pension from the public system, taking into account access to the system of other benefits (including in-kind benefits), is not sufficient to mitigate the retirement risk. The share of such households with materialised retirement risk is especially high in one-person households (above 50%).

The limited possibilities for expanding cash benefits from the public pension system and the lack of adequate quantity and quality of benefits from social

welfare systems collide with the increase in life span. This makes it necessary to build flexible products that use the full palette of solutions from existing instruments. But combining them appropriately, introducing the variability of individual benefits within a single product makes it possible to manage the retirement risk, providing the adequate benefits in retirement with respect to the changing needs of the individual over time.

We aimed to develop a comprehensive retirement insurance solution (CRIS) based on a technical modular design, the goal of which is to provide services tailored to customers' needs at reduced cost and time (Dörbecker & Böhmann, 2013). The product design is based on a tiered solution, with basic coverage surrounded by layers of additional benefits. The first layer consists of several basic types of insurance coverage (e.g., life insurance, hospital stays, or professional adaptation). The second-layer benefits include those that affect the use of base insurance coverage or insured forms of assistance benefits (e.g., the ability to change the scope of benefits purchased during the insurance period, with guaranteed benefits). In the final layer, benefits are triggered that allow participants to use insurance products on an ongoing basis (including during the savings period) based on a retirement benefit calculator, financial education, etc. Treating multiple modules as a single financial product provides tax advantages by transferring accumulated funds and entitlements to other types of benefits—there is no benefit payment, which means that there is no consequent income tax accrual. A new solution to meet insurance participants' future pension needs must answer the following question: How should products in the financial market be shaped to meet the changing needs of people in retirement?

Theory of change (ToC) can be helpful in this regard because, in simple terms, it can be used to identify long-term goals and thereafter support 'backwards planning' (Brest, 2010), which means identifying the conditions that need to be met to achieve the goal and deciding what actions need to be taken to ensure that these conditions are met. In terms of insurance coverage for retirement risks, this means acting and making adjustments during and after working life that allow insurance products to be used to achieve an ultimate goal, which inevitably involves a need for more flexible financial products. Such measures are in line with the assumptions of modular products already in use in the insurance market.

The paper uses the method of analysis and logical construction with respect to the theory of change (ToC). The issue of the optimal combination of benefits during the period of retirement benefits was divided into smaller factors and the properties of each factor were examined individually. The effects of this activity were combined again as a result of logical construction. Consequently, this process makes it possible to find a new solution which would be flexible. The SWOT analysis has been carried out to deliver the qualitative analysis of this solution.

The paper is organised as follows. In Section 1, we briefly discuss the goals of retirement system in the context of retirement risk. Section 2 provides a brief overview the possibilities and deficiencies of life insurance products in the covering the retirement risk. In Section 3, we propose a comprehensive retirement insurance solution (CRIS) and describe the specific modules it can provide. The final section presents conclusions.

1. Pension system goals versus retirement risk

The most generally accepted purpose of a pension system is to prevent poverty in old age and to protect people's standards of living (at least partially) by providing income for all those covered by the system for the entire period after retirement (Myles, 2002; Turner, 2010). Social development has led to a situation in which pension benefits are expected to satisfactorily meet needs, come from a variety of sources, and have 'sustainable' payment characteristics (Amaglobeli et al., 2019). In most developed countries, the main source of retirement income is the state old-age pension; however, the value of such a pension is likely to decrease as pension reforms more often concentrate on sustainability rather than adequacy (Hagemejer, 2018; Szczepański et al., 2022). Within private pension plans, there has been a shift from defined-benefit (DB) to defined-contribution (DC) plans, which means transferring to participants the risk of ensuring the amount of pension assets (Mitchell & Utkus, 2012). The first wave of pension reforms in 1990s concentrated on the accumulation of funds for retirement purposes (James & Vittas, 2000), but interest in the decumulation phase is increasing, as the proliferation of defined contribution plans necessitates decisions regarding the decumulation of assets. Moreover, increasing life spans require the provision of LTC, which is not a common social security component. For example, in the United States (US), an estimated 70% of people aged 65+ will require long-term services and support at some point in their lives (Super et al., 2022). Additionally, senior citizens often regret not having adequately provided for their retirement needs. Hurwitz and Mitchell (2022) indicated that many retirees regret having too few savings and/or not obtaining appropriate insurance products to cover their retirement needs. They also found that making Americans aware of objective survival probabilities increased their levels of regret for not securing LTC insurance (i.e. it increased by more than 2.4 times their regret for not having purchased lifetime income support (Hurwitz & Michell, 2022). In Poland, a country where senior citizens' social security is almost entirely dependent on public spending, a study was conducted on the representative sample that asked retirees aged 65 and over what types of insurance from a lifetime perspective they find useful (Bielawska

& Łyskawa, 2019). The respondents' answers clearly indicated that expectations of recognised and available insurance changed with the age of retirees. Those aged 65–74 years indicated that the most relevant insurance products would cover health benefits, but 63% of the oldest senior citizens (85 years and older) indicated a need for LTC insurance. These responses point to the public system's failure to meet retirement needs in long-term care services, since the market in Poland, as in many other countries, does not provide adequate insurance products for retirees.

Needs in retirement change with age, but also with many other individual factors. Therefore, we considered retirement risk through the lens of seniors' subjective perceptions of their needs, which encompassed not only access to cash, but also in-kind benefits. We perceive the holistic approach to the assessment of the senior citizens' material security as an inevitable in the era of demographic and economic changes.

2. Life insurance for covering retirement needs

Voluntary saving for retirement requires reducing consumption by the periodic setting aside of money throughout one's working life up to retirement age. Market practices have developed many forms of saving and investment products, including life insurance, which can be categorised as relating to the accumulation (pre-retirement) or relating to the decumulation phase (used to provide an income stream in retirement). The first group includes traditional participating life insurance, unit-linked, universal life, and dynamic hybrid products. To provide a stream of income during retirement, the following types of annuities may be used: immediate or deferred, fixed-term or lifelong, traditional (with profit), unit-linked, variable, or equity-indexed. The decision to choose life insurance depends on the future retiree's awareness of the need to supplement the benefits and services provided by the social security system, his/her propensity to save and financial literacy, and many other factors. However, Crawford and O'Dea (2020) pointed out that the actual replenishment of pension benefits is influenced by the accumulation of assets for as long as possible, at the expense of other needs, from the end of working life or during the first years of retirement (e.g., to counteract digital exclusion or poorer health).

Despite calls for the transparency of benefits for pension system participants, frequent changes in the design of such systems make it difficult for individuals to assess what their financial situations might be on retirement. In many countries, employees have become accustomed to the assurance of having the means to continue living after retirement delivered by the state,

which makes it difficult to attract them to traditional life insurance products, since savings need to be set against needs that may arise in 20–30 years' time. Simply making people aware of their needs is not enough. It is necessary to build flexible mechanisms that can be adapted to the defined goals of the insurance participant.

Savings insurance programmes are based on customers' high-end capital and, when transformed into periodic annuity payment programmes, allow for an 'additional pension' and supplement the benefits of the basic pension system. There is a general rule in life insurance that the person starting the insurance contract should be free to choose its main function. A lack of awareness of the need to build adequate capital for an additional pension causes young people to prefer benefits in the event of an unfortunate accident over saving for retirement. The design of such products is clear: the greater the benefits for adverse events in an insured person's life, the smaller the available final capital for a retirement annuity (Carlson, 2016). When taking out insurance, it is important to be aware that insurance companies—at the contract stage—cannot determine the precise value of a future annuity. They present only conservative estimates, while making the insured individuals aware that no specific amount is guaranteed. Years later, the benefits may differ significantly from the estimates (Milevsky, 2006); hence, it is necessary to keep analysing the value of insurance policies and the benefits, if any, that will be paid. However, it is extremely difficult for individual participants to assess the valuations of these instruments, and such assessments depend on their levels of education and numerical skills (Brown, Kapteyn et al., 2017). Annuities are proposed to be the main tool for decumulation of retirement savings in private pension plans. Annuities are a necessity for retirees to maintain an adequate standard of living (Antolin & Stewart, 2009), but the low interest rates of recent years have undermined the profitability of annuities, especially for life insurers (Antolin et al., 2011).

Insurers are developing products with alternative return systems and moving away from fixed interest-rate guarantees (Beer & Gnan, 2015). However, there is a great need to develop rules for the funding and solvency of pension plans (especially defined benefits) so that they remain countercyclical. Thus, the effectiveness of annuities in covering longevity risk should be monitored on an ongoing basis (Wettstein et al., 2021). Limitations on interest in these forms of savings and accessibility for all participants in the pension system are also vital (Lambregts & Schut, 2020).

3. The concept of a CRIS

Fundamental Assumptions in New Insurance Product Development

The assumptions listed below form the basis for the creation of a new CRIS, which should do the following:

1. Assume the continuation of the social security system as it currently operates, which does not adequately cover retirement needs for all its participants (i.e. assuming no policy change).
2. Based on a modular design (that is, insurance products providing cash and in-kind benefits in retirement as both related and independent solutions), it should include separate demands and negotiations by participants, but simultaneously not be available for purchase outside CRIS.
3. Consider the dual nature of benefits and compensation: overall, the retirement product should allow the payment of certain benefits in cash, as well as covering the cost of certain services and even goods (in line with the retiree's basic requirements). In addition, the product should include extensively developed assistance packages triggered in certain situations (triggers). This is the basis for the idea of a three-tier product.
4. Be materializable: due to the long-term nature of fund accumulation and the lack of or limited nature of CRIS, it is necessary to introduce solutions that allow participants to record acquired entitlements or financial assets on an ongoing basis. This will make it possible to keep in contact with pension product holders and introduce educational or preventive elements, especially regarding health situations.

In terms of methodology, ToC was used to prepare the CRIS and implement the necessary changes (Dhillon & Vaca, 2018; Mayne, 2015). This theory requires the identification of long-term goals to level the effects of retirement risks. Thus, it is necessary to identify the conditions that must be met to achieve the assumed level of retirement risk levelling. In operational terms, it is also necessary to identify the actions that need to be taken to achieve the desired results. ToC requires constant adjustment of actions based on strong theoretical assumptions (Reinholz & Andrews, 2020), and experience in applying ToC, mainly in education, has shown that it is vital for those responsible for change, such as insurance underwriters or pension system researchers, to use a common language (Connolly & Seymour, 2015).

Modular product design

The modern market for financial products is characterised by mechanisms for interactive product selection, which is particularly evident in the insurance field. Usually, this means scaling individual insurance coverage and choosing among different providers. In many cases, the insurance products offered are package products, with a single basic product and a set of optional extras but no ability to modify the scope of coverage. The shortcomings of traditional bundled products are supposed to be offset by the 'brick' (modular) design of the insurance product (Schmidt-Jochmann & Gröbner, 2012).

Insurance products constructed in this form consist of separate components ('bricks') that can be either sold as part of a modular product or marketed independently. The individual 'bricks' thus function as separate insurance products, which are standardised based on the parameters of the product design. The advantage of modular construction from the consumer's point of view is that modules can be selected or configured based on a combination of customer requirements and risk characteristics. Consequently, a modular product allows for flexible combinations of individual components. Such a design meets two seemingly contradictory goals: 1) from the customer's point of view, individualisation of insurance, and 2) from the insurance company's point of view, standardisation of the most important benefit elements. Each

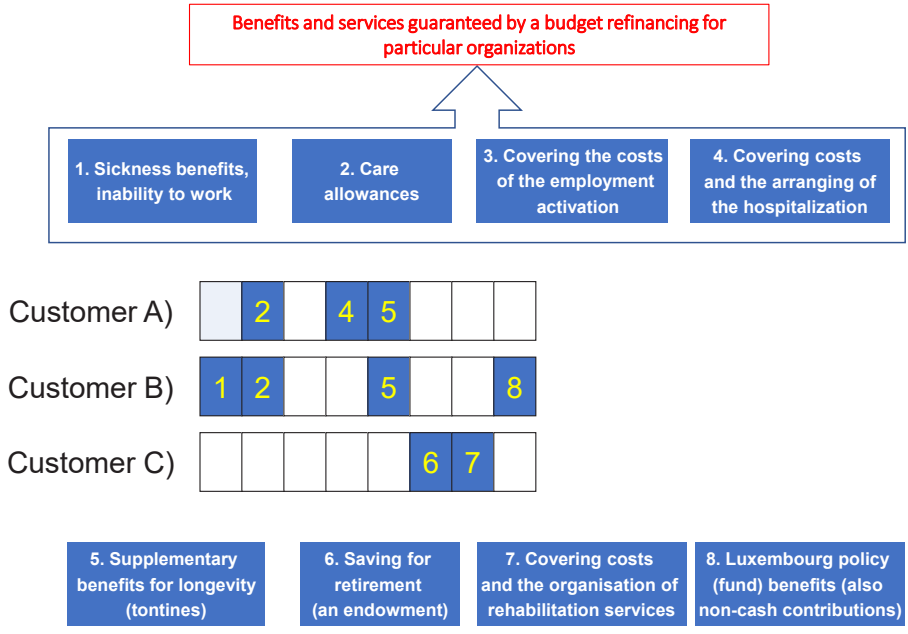

Figure 1. Proposed coverage for integrated modular pension insurance or CRIS

Source: own work.

of the proposed modules covers different consequences of approaching or actual retirement age (Figure 1). Therefore, the details of the selected product elements are discussed in the following sections.

At this point, it is necessary to determine the number of feasible options for an individual. In purely technical terms, the number of possible combinations is provided by the following formula: the number of possible module choices is $2\hat{}n$ minus the number of combinations after (min–1*) of n (*assuming that each module is only a 'take it or leave it' choice without any further variants). Bearing in mind that even a single module can be selected as the end result, this solution (8 modules) allows 255 different types of pension product combinations to be built.

In the case of personal insurance products relating to health status or future LTC expectations, pure product configuration mechanisms are unsatisfactory because for each selected configuration, products from multiple providers may perform similar functions (Stolze et al., 2000). In such a situation, it is necessary to build a product configuration mechanism (Felfernig et al., 2007) based on a defined and continuously calculated retirement risk. This means that before entering into an insurance contract or during its term of validity, each participant should have the ability to estimate the income and benefits already available. Each insurance company is obliged to offer an appropriate pension calculator and configurator to provide an optimal solution for a given participant.

The design of an insurance product for covering retirement risk should aim to reach a particular age cohort. The differences between age groups are large enough to determine the need for certain modules for each age cohort (e.g. those related to LTC at a sufficiently high level). Building appropriate distribution channels requires analysing the habits and possible technological exclusions of individual cohorts (Ayuso et al., 2021).

The product base for the proposed solution is traditional life and endowment insurance, the goal of which is to create financial security for insured individuals and enable them to maintain a certain standard of living at retirement age, while providing material support for the family in the event of their premature death (Nurittamont, 2021). Any contract of this kind, unless broken by the policyholder, ends with the payment of a benefit. Premiums for endowment insurance are paid throughout the insurance period and, most often, due to their considerable cost, in instalments. According to Skipper and Black (2000), the net premium (which creates a fund to cover the insurance company's future liabilities) for life and endowment insurance is based on the calculation of a term of life insurance (protection only in the event of death during the assumed period) and endowment insurance (premium calculated in such a way that after the assumed period of payment of premiums, a benefit of specified value is paid). Benefits are paid as a single payment or in the form of an annuity based on the current sum insured. In practice, its

value corresponds to the value of the premium paid for life insurance plus the interest rate on that premium. Within CRIS, the accumulated funds in life insurance are transferred to purchase the other modules, according to the need of the individual. Moreover, the addition of the tontine module makes it possible to obtain additional benefits for the members of a given cohort based on demographic changes (e.g., increased mortality). The period of participation in the programme is closely linked to the year of birth. It is necessary to construct uniform age cohorts, possibly differentiated further by place of residence. Endowment products include a minimum participation period (Huebner & Black, 1982); therefore, it is necessary to build a mechanism that will guarantee sufficiently long participation in the proposed solution. The possibility of changing the number of proposed modules can also be a response to negative regulatory changes for the insured regarding the basis of the pension or social benefit system. From the point of view of ToC, it is necessary to create a strategy for changing systems. In addition, it is necessary to keep an eye on the impact of the environment on participants' decisions to change individual modules. In the face of significant social change, it is also important to analyse the involvement of other family members in financing retirees' needs or selecting specific products. Institutionally, the proposed solution should be based on the operations of life insurance companies. Due to the benefit guarantees specified hereafter, these entities should have the ability to implement public–private partnerships, as is already done in the field of LTC (Cohen et al., 2018).

It should be emphasised that the modular solution described above and the large number of combinations are intended to show how many possibilities this type of insurance offers. However, it is essential to note that, from the point of view of the objective of retirement risk management, a separate study should be carried out with the question: Which of these proposed modules are most important for future retirees? On the one hand, political pressure means that more and more benefits that could be placed in such an insurance solution are being shifted to social assistance and provided by the state or local government. On the other hand, however, it is not possible to count on the needs in a given area being met in full, due to the shortages in the public funds. We are convinced that the use of a modular product, with different types of insurance that cover the retirement risk, would give impetus to the launch of a risk management mechanism by pension scheme participants.

Scope of protection of individual CRIS modules

Sickness benefits and inability to work

Insurance companies can, through their products and their skilful supply, contribute to the co-funding of the state healthcare system, which is in constant flux (Böheim & Leoni, 2017). Their role should be to supplement rather than replace the public system. According to Wang et al. (2012), a product offered to meet this objective should:

– cover risks relevant to people of retirement age (referring to actual and uncovered risks),
– be accessible to all interested parties,
– complement other general insurance programmes,
– be characterised by a high degree of flexibility in terms of individual types of disease events,
– provide protection for all or part of the actual cost of treatment (moving away from lump sum rules).

Care allowances

LTC policies provide security for the assets of an insured person who requires LTC. The market for this type of product is automatically limited by the following:

– health status,
– age (only people up to a certain age limit are accepted for insurance),
– the cost of the realised LTC.

Insurance products of this type are usually purchased by people of advanced age, but this trend is changing, especially in group insurance. Individual LTC products in the 1990s were purchased by customers whose average age was 72 years. For those who participated in group plans, this average age was 69 years, while for employee insurance, it was 43 years (Gelder & Johnson, 1997). LTC insurance policies offered today include coverage for nursing care (at various levels of intensity and expertise), care for the aged, and care for the disabled. Modern policies expand the scope of the benefits covered. As a result, the design of LTC insurance has changed, and it is now increasingly focused on solutions based on the number of payments (depending on the level of disability assessment or actual expenses incurred) and the provision of specific services (Denuit et al., 2019). From another perspective, Hieber and Lucas (2022) have indicated that the coverage presented is attractive to potential customers, and possible difficulties with the solvency of implemen-

ted insurance programmes can be overcome by using the tontine mechanism. When discussing LTC insurance, it is important to consider the conditions that determine the start of benefit payments. In earlier types of policies, insurance companies primarily required a hospital stay, a doctor's call (an extremely casual approach), or necessary actions in case of illness or disability (standard medical benefits). The most common system for issuing LTC insurance payouts is one based on disability ratings (Chen & Xu, 2020).

Covering the costs of employment activation

The decision to leave the labour force after the statutory retirement age is determined by a number of factors. Rapid social changes (related, e.g., to the development of information technology) and transformations in education systems (stronger specialisation and focused training, particularly for young workers) frequently worsen the position of the elderly in the labour market relative to other social groups. If these factors are further reinforced by unemployment or deteriorating health, there may be a strong desire to retire. However, it is also important to have the right competencies and opportunities to work before retirement age. The CRIS module presented here addresses both periods of employability: 1) until retirement age and 2) after retirement age. Age discrimination in the labour market is common in almost every country. Of course, performing a paid job at an older age allows for increased pension benefits from the basic pension system. Ribeiro et al. (2018) proved that working at an older age is associated with better physical health and social relations. The proposed features of the insurance module 'Retirement Insurance Product for Job Activation' are necessary for the preparation. The rule of this insurance is to reflect the current demand for training, services, or financial support related to vocational activation. Therefore, there is a need to constantly monitor the labour market and consider the various factors that influence the exclusion of older people.

Covering costs and arranging hospitalisation

There have been significant annual increases in the number of patients treated in hospitals because the increasing incidences of diseases and mental changes have forced people to make greater use of their insurance entitlements under the state health care system. In the case of the elderly, there are additional age-specific conditions, such as frailty (sarcopenia) (Zhang et al., 2018), the risk of malnutrition (Fávaro-Moreira et al., 2016), and even dementia, for which a stay in hospital is not the best form of care (Cowdell, 2010). Despite hospitals treating larger numbers of patients more promptly, there are still many unmet needs. A huge number of patients are waiting for

healthcare, which, in addition to their unnecessary suffering, creates a huge social burden, especially for the elderly.

The proposed design of this insurance module primarily involves the organisation of a hospital stay, the ability to choose the unit (ward), and even the attending physician and nursing care. The insurance company's coverage of the costs of medical procedures is based on a formula for listed surgical units (named 'perils') and disease units. The level of coverage for surgeries can be 100% or via specified lump sums. When several operations are carried out simultaneously, the insurer covers the full cost of the highest-classified operation, a maximum of 50% of each subsequent operation in a different operating group, and a maximum of 25% in the same operating group. This kind of deductible makes it possible, on the one hand, to reduce the cost of insurance coverage (premiums) and, on the other, allows the use of funds accumulated in other modules.

It is also important to bear in mind the phenomenon of supply-induced demand (SID), which has been observed in many countries (Yu et al., 2020). The keys to explaining this phenomenon are information asymmetry and the idea of moral hazard. Providers / professionals, based on their status as experts, sometimes generate strings of in-house or outsourced services. This occurs more frequently if tests and medical procedures are paid for by the patient individually or the contract with the payer provides reimbursement of all costs incurred. Patients (especially the elderly) with little knowledge of actual health needs or how to meet those needs tend to rely on the knowledge of the doctors who decide the actual course of treatment. Thus, the peculiarity of the market for medical services is that the demand curve for medical services reflects the preferences of doctors (the providers of services) rather than patients (the consumers) (Bickerdyke et al., 2002). Insurance companies may seek to establish their own hospitals (envisioning a steady stream of newly insured potential recipients of such services) and have additional quality and cost control (Burns & Pauly, 2018).

Supplementary benefits for longevity (tontines)

Tontines have a long history, yet modern tontines have little in common with loans to the Treasury in exchange for lifetime annuities. They have been handed over to private institutions—insurance companies. However, the essence of a tontine, distinguishing it from a life annuity or, today, from classic life insurance, is still valid: in the absence of payment of the contracted benefit to the insured person or another person designated by him/her, the accumulated funds pass to the other tontine participants in the same class.

CRIS assumes that each tontine has a certain size (e.g., a minimum of 2,000 members in a given age cohort based on 5-year stages), the profits of which are entirely reserved for members of that class who live to the designated

age. Considering the purpose of the modular insurance in question, this age should be set well above the retirement age limit for the country. This would facilitate improved investment returns for those who accumulate additional savings for retirement and are not concerned about the eventual financing of their needs. However, during the entire duration of this association, savings are unavailable.

There are two types of tontine associations:

1. Tontines with regular contributions (transferred monthly, quarterly, semi-annually, or annually while allowing the regular accumulation of savings), necessarily with a minimum contribution period determined on the basis of appropriate actuarial calculations.
2. Tontines with a one-time contribution (only one transfer during the contribution period, defined as a minimum, with the intention of multiplying the initial capital).

At this point, it should be noted that single-premium tontines can be financed through another module associated with the product presented: the Luxembourg insurance (funds) module. Tontines reflect the idea of community and solidarity, not only in terms of financing but also in terms of an ideology that unites a group of people (Kemayou et al., 2011). Tontines resemble ordinary investments, with the major difference that tontine investments are usually irrevocable (Chen & Rach, 2022). In the proposed solution, insurance companies take on the role of managing individual tontines.

Saving for retirement (endowment)

The wide variety of available retirement investment programmes requires basic systematisation. In modern pension systems, which attach much greater importance to the foresight of their participants than in the past, two basic types of additional retirement savings can be distinguished:

1. Systems organised by companies or socio-professional groups.
2. Individual savings, the purpose of which is to increase income after retirement age, which should be separated from short-term savings or those with a purpose that differs from increasing the level of pension benefits.

In the proposed solution (CRIS), we deal exclusively with individually initiated and paid-for products. Technically, it is difficult to identify optimal insurance products. The development of financial services is so dynamic that new products appear as often as several times a year, providing the possibility of using individual instruments or combining them with others. Therefore, the task of the operators of the proposed modular product is to ensure sufficient flexibility of the proposed solutions to expand, e.g., the package of proposed investment directions on an ongoing basis or to modify the rules of partici-

pation (Sohn, 2017). It is also necessary to respond dynamically to changing investment expectations (Mahayni & Muck, 2017). If there is no confidence in capital markets (e.g., due to low interest rates), insurers must consider other options that can satisfy the expectations of policyholders. At present, this means directing investments to companies actively involved in shaping the 'green revolution' and pursuing sustainable development (Nguyen et al., 2018). At all times, however, it is necessary to guarantee participants the protection of already acquired rights (Zelizer, 2017).

One of the most difficult tasks facing pension systems based on defined contribution principles is protecting system participants with low incomes (and, as a result, low pension contributions) (Rajasekhar et al., 2017). An example of such a measure is a mechanism that provides a guaranteed minimum pension paid from the system or certain social allowances (Butler, 2016). This brings the sum of the pension and social allowances received closer to the established level of the minimum pension. People with low incomes have little opportunity to voluntarily generate additional retirement savings (regardless of the quantity and quality of available solutions). From the state's point of view, in such a situation, it is necessary to consider what is more cost-effective in the long run in terms of overall social goals: creating special subsidised programmes, co-funding contributions during the accumulation period, or paying additional benefits during the pension realisation period. The proposed modular solution, especially in the context of solvency guarantees for insurance companies, aims to provide greater accessibility for people with low incomes by setting a better-than-standard cost structure at or a lower value (Benish et al., 2016).

Luxembourg policy (fund) benefits (with the option of non-cash contributions)

A Luxembourg policy (fund) is an arrangement that uses owned or accumulated assets to invest in a certain way, and the purpose of investments is to generate funds to cover expenses defined in other modules of the product under review (most often, a minimum value of assets under management). Investment funds can operate in two main legal forms:

1. undertakings for collective investment in transferable securities (UCITS),
2. alternative investment funds (AIFs) (Kofoworola et al., 2019; Rokas & Siafarika, 2019).

The proposed product is based on an investment in an insurance solution. The starting point is payment to the insurance company. Investments in this regard are not significantly different from those of the 'saving for retirement' module. However, notably, the 'deposit' can also have a property or business value converted into a monetary value. In such a situation, a valuation of the

asset is made at the time the contract is concluded, and the equivalent in a specific currency is transferred to the account of the managing institution.

Luxembourg funds benefit from several conditions, the imposition of which should safeguard the interests of insured people and be a *de facto* supplement to pension benefits in the country where such an instrument is launched. In conclusion, it should be stressed that a Luxembourg policy in its basic form is no different from common insurance with a savings element. But at CRIS, we would like to focus primarily on the additional option in this type of policy, where the insurance premium does not necessarily have to be paid in cash. The premium can also be an asset: housing, land, paintings, valuable movables.

Three-layer CRIS design

Modern three-layer financial products (including life insurance) are not only tools for covering specific contingencies (living to a certain age or the death of the insured person) but are often implemented as contracts for specific goods and services. Figure 2 shows the construction of the three layers of the proposed CRIS. The proposed solution assumes the use of additional insurance products in layer 2. In this regard, the modification of selected modular solutions is envisaged (changing the number of modules or the extent of coverage of individual products). This layer also provides the possibility of the payment of a premium of the policyholder's children (one-time or over a certain number of instalments). Consequently, the capital accumulated by parents can be transferred to the products they need in given situa-

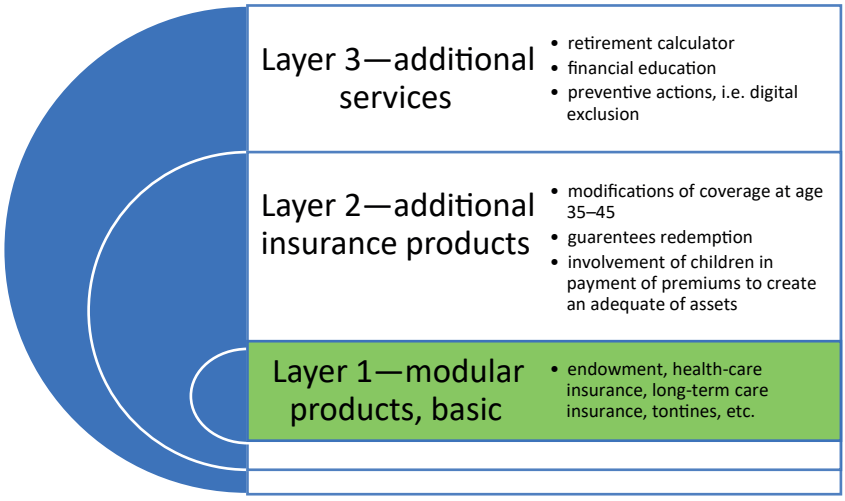

Figure 2. The three-layer design of CRIS

Source: own elaboration.

tions (e.g., the module 'Covering Costs and the Arranging for Hospitalisation' or 'Care Allowances').

Guaranteed realisation of selected benefits

The design of the multimodule pension product was assumed to cover the basic needs that arise during retirement. Despite changes in expenditures and the length of life in retirement, certain types of expenditures are crucial for preserving the health and lives of elderly people. Therefore, guided by social responsibility for this social group, it is necessary to create a state budget mechanism to guarantee the financing of certain types of benefits in a modular product, directed towards life insurance companies that sell and service modular insurance. The guarantee would cover modules related to: 1) sickness benefits and inability to work, 2) care allowances, 3) covering the costs of employment activation, and 4) covering the costs of arranged hospitalisation.

This means that quasi-reinsurance solutions currently operate in many countries (e.g. insurance for the agricultural sector). In addition, reinsurance mechanisms can be applied to social risks (Dror & Preker, 2002), but this requires that governments allocate part of their budgets in order to equalise the loss ratios of given insurance companies. Within CRIS, this means that a large risk materialisation in one year is financed from budget resources in the next year. Such a mechanism can be financed by reducing the necessary budgetary measures to provide these social welfare benefits. However, it is necessary to ensure the stability of such a solution. Life insurance companies and CRIS participants must be confident that the refinancing mechanism will function over a multi-year period and will not be subject to ad hoc decisions. Therefore, the premiums charged for each module should be verified by national actuaries. However, there is also a need for a public reinsurer entity, which must be guaranteed by the national treasury if its solvency is threatened. In the event of a need for disbursements in excess of the funds accumulated in the system, the security and liquidity of the proposed solution is ensured by the treasury guarantee. The reinsurer makes demands on the relevant state funds for recapitalisation (or infusion of funds depending on the legal form), and the entity responds with an appropriate issue of bonds or other risk securitisation instruments. The reinsurer's obligations to insurance companies to offer guaranteed benefits and services are unconditional, so the reinsurer must have state treasury guarantees.

The qualitative analysis of CRIS

Evaluating the effectiveness of the different parts of the new solution (CRIS) requires appropriate quantitative research both among people in working life

Table 1. SWOT analysis of CRIS

Strengths	Weaknesses
– hyper-personalisation: possibility to choose different risk coverage options at the time of contract conclusion and to change the choice during the contract period – the ongoing adaptation of benefit variants in individual products to a changing environment (e.g., additional in-kind public benefits, increased or decreased cash benefits) thanks to the application of the theory of change to the management of this product – competitiveness in comparison to traditional product solutions, as it incorporates the positive features of health insurance or unit-linked insurance, while allowing the use of tontine as a method of spreading risk in the population	– the need to build appropriate long-term investment solutions (e.g. change of legal regulations and introduction of tax incentives in other solutions) – portability between products can create a moral hazard for participants to use funds in the most expensive services – lack of an appropriate participant scale may destroy the insurance nature of CRIS (failure to spread the burden of benefits across a large group of participants) – an increase in the use of non-insurance benefits (layer 3) may disrupt the cost-effectiveness of the overall CRIS construction (particularly relevant for insurance and financial coverage in layers 1 and 2)
Opportunities	**Threats**
– an increase in competition between publicly funded objectives (e.g. energy transition or ESG) will reduce the interest of successive governments in increasing benefits for pensioners or building new types of benefits – technological developments are allowing more individualisation to be built into financial, investment product choices (CRIS meets these expectations) – the propensity of people, of all ages, to gamble may be the basis for engagement with the tontine mechanism	– a change in the level of benefits from the universal system or an increase in health and long-term care assistance will reduce interest in CRIS – the expansion of other financial products may blur the distinction between CRIS and other insurance and investment products – continued or worsening lack of interest in building capital on their own for use in retirement

Source: own work.

and among pensioners. At the conceptual stage, however, a qualitative assessment of the proposed solutions can be made. A SWOT analysis was used in this respect (Table 1), which is a technique for organising and analysing information related to the implementation of a new solution (Valentin, 2001).

The implementation of a new solution (CRIS) requires, above all, a willingness to change. This solution can take effect without a major revolution in the pension system or profound changes in legal regulations. In some countries, starting to offer tontines requires an adaptation of legal regulations to those enshrined in the directive for the operation of life insurance. However, unless there is a clear willingness to improve, and to do so in an iterative, i.e.

continuous improvement, there will be no chance of achieving sustainable results in improving the funding of retirees' needs.

As part of further work, appropriate tools and schedules would need to be prepared to manage this type of change. Including cost estimation and performance measurement in the analysis presented, which is the authors' intention as part of further work, would enable a holistic view of the economic elements, including the question of the value of this complex product to the customer. However, this requires further research in relation to the CRIS product as a whole, as well as to its individual components. The theory of change proposed in the paper allows resources to be used efficiently, effectively and to create sufficient value to justify the use of those resources (King, 2021).

Conclusions

The demographic shift that is driving a sustained increase in the proportion of elderly people in the population requires the adaptation of public policies and solutions offered by the private sector to mitigate retirement risk. The coverage of this risk can come from both income (cash benefits) and access to services. In a situation where ongoing pension reforms transfer the responsibility for pension benefits to participants, combined with increasing life expectancy, it is necessary to provide comprehensive solutions to reduce retirement risks.

For years, researchers have advocated the development of the annuity market and the development of LTC insurance. We are convinced that the focus on separate insurance products is inadequate. Therefore, we propose a new comprehensive view of retirement risk protection that can be flexibly adapted to the changing needs of insured people. CRIS is based on a three-layer modular product. The first layer combines insurance products that can accommodate policyholders' needs solely during retirement (as tontines or LTC insurance), solely during working life (to cover the cost of employment activation), or in both periods (pre-retirement and retirement). The second layer provides additional insurance products (i.e. guarantees), and the third layer offers different additional services to enhance policyholders' understanding of retirement risk and support their decisions regarding changes in modular products to better fulfil their needs. The modular insurance product should provide wealth creation and be part of the pension system. Its operation will reduce the social welfare system's obligations to individual retirees and should therefore not be taxed. The transfer of entitlements between modules allows individual resources and wealth to be built. However, the effects will be felt after the end of working life or will be transferred to the next generation (through inheritance).

The proposed solution, based on a ToC mechanism of insurance product creation and evolution, makes it possible to constantly review the solutions used, modify them internally, or add new modules. Only such an approach will provide an ongoing response to the changing mix of goods and services that future retirees will require. Simultaneously, the application of a three--tier solution will permit insurance participants to make changes in their selected products both during the period of accumulation of funds and vesting, as well as during retirement and the partial use of funds. The SWOT analysis carried out indicates that there are many opportunities in front of this new product solution due to the need to individualise product solutions and retirement risk management. But similar constraints as in existing pension products remain: the lack of propensity to build up additional savings for retirement and the state's commitment to raising benefits or providing additional care and health benefits for retirees.

References

Amaglobeli, M. D., Chai, H., Dabla-Norris, M. E., Dybczak, M. K., Soto, M., & Tieman, A. F. (2019). *The future of saving: The role of pension system design in an aging world*. International Monetary Fund.

Antolin, P., Schich, S., & Yermo, J. (2011). The economic impact of protracted low interest rates on pension funds and insurance companies. *OECD Journal: Financial Market Trends*, *1*, 237–256. https://doi.org/10.1787/fmt-2011-5kg55qw0m56l

Antolin, P., & Stewart, F. (2009). *Private pensions and policy responses to the financial and economic crisis*. OECD Working Papers on Insurance and Private Pensions, 36. https://doi.org/10.1787/224386871887

Ayuso, M., Bravo, J. M., Holzmann, R., & Palmer, E. (2021). Automatic indexation of the pension age to life expectancy: When policy design matters. *Risks*, *9*(5), 96. https://doi.org/10.3390/risks9050096

Beer, C., & Gnan, E. (2015). Implications of ultra-low interest rates for financial institutions' asset liability management—a policy-oriented overview. *Monetary Policy and the Economy*, *2*(Q2), 15–53.

Benish, A., Haber, H., & Eliahou, R. (2016). The regulatory welfare state in pension markets: Mitigating high charges for low-income savers in the United Kingdom and Israel. *Journal of Social Policy*, *46*(2), 313–330. https://doi.org/10.1017/s0047279416000593

Bernard, C., Cui, Z., & Vanduffel, S. (2016). Impact of flexible periodic premiums on variable annuity guarantees. *North American Actuarial Journal*, *21*(1), 63–86. https://doi.org/10.1080/10920277.2016.1209119

Bickerdyke, I., Dolamore, R., Monday, I., & Preston, R. (2002). *Supplier induced demand for medical services*. Productivity Commission Staff Working Paper, Canberra.

Bielawska, K., & Kozłowski, A. (2024). A proposal for retirement risk measurement based on subjective assessment of income: An empirical study. *Social Indicators Research*, *172*, 1–28. https://doi.org/10.1007/s11205-023-03295-3

Bielawska, K., & Łyskawa, K. (2019). Wykorzystanie produktów ubezpieczeniowych i inwestycyjnych przez obecnych emerytów – wyniki badań. In I. Kwiecień & P. Kowalczyk-Rólczyńska (Eds.), *Ubezpieczenia: wyzwania rynku* (pp. 11–24). C.H. Beck.

Böheim, R., & Leoni, T. (2017). Sickness and disability policies: Reform paths in OECD countries between 1990 and 2014. *International Journal of Social Welfare*, *27*(2), 168–185. https://doi.org/10.1111/ijsw.12295

Brest, P. (2010). *The power of theories of change: Stanford social innovation review*. Springer.

Brown, J. R., Kapteyn, A., Luttmer, E. F. P., & Mitchell, O. S. (2017). Cognitive constraints on valuing annuities. *Journal of the European Economic Association*, *15*(2), 429–462. https://doi.org/10.1093/jeea/jvw009

Brown, J. R., Mitchell, O. S., Poterba, J. M., & Warshawsky, M. J. (2001). *The role of annuity markets in financing retirement*. The MIT Press. https://doi.org/10.7551/mitpress/6007.001.0001

Burns, L. R., & Pauly, M. V. (2018). Transformation of the health care industry: Curb your enthusiasm? *The Milbank Quarterly*, *96*(1), 57–109. https://doi.org/10.1111/1468-0009.12312

Butler, M. (2016). Insights from Switzerland's pension system. In O. S. Mitchell & R. C. Shea (Eds.), *Reimagining pensions: The next 40 years* (pp. 247–273). https://doi.org/10.1093/acprof:oso/9780198755449.003.0013

Carlson, R. C. (2016). *The new rules of retirement: Strategies for a secure future*. John Wiley & Sons.

Chen, A., & Rach, M. (2022). A note on financial fairness in tontines with mixed cohorts. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4009770

Chen, L., & Xu, X. (2020). Effect evaluation of the Long-Term Care Insurance (LTCI) system on the health care of the elderly: A review. *Journal of Multidisciplinary Healthcare*, *13*, 863–875. https://doi.org/10.2147/jmdh.s270454

Cohen, M., Feder, J., & Favreault, M. (2018). *A new public-private partnership: Catastrophic public and front-end private LTC insurance*. Urban Institute and LTSS Center @uMass Boston. https://www.umb.edu/media/umassboston/content--assets/gerontologyinstitute/Public-Catastrophic-Insurance-Paper-for-Bipartisan-Policy-Center-1-25-2018.pdf

Connolly, M. R., & Seymour, E. (2015). *Why theories of change matter*. WCER Working Paper, 2015-2. https://wcer.wisc.edu/docs/working-papers/Working_Paper_No_2015_02.pdf

Cowdell, F. (2010). The care of older people with dementia in acute hospitals. *International Journal of Older People Nursing*, *5*(2), 83–92. https://doi.org/10.1111/j.1748-3743.2010.00208.x

Crawford, R., & O'Dea, C. (2020). Household portfolios and financial preparedness for retirement. *Quantitative Economics*, *11*(2), 637–670. https://doi.org/10.3982/qe725

Denuit, M., Lucas, N., & Pitacco, E. (2019). Pricing and reserving in LTC insurance. In E. Dupourqué, F. Planchet, & N. Sator (Eds.), *Actuarial aspects of long term care* (pp. 129–158). Springer.

Dhillon, L., & Vaca, S. (2018). Refining theories of change. *Journal of MultiDisciplinary Evaluation*, *14*(30), 64–87. https://doi.org/10.56645/jmde.v14i30.496

Dörbecker, R., & Böhmann, T. (2013). *The concept and effects of service modularity—a literature review*. 2013 46th Hawaii International Conference on System Sciences. https://doi.org/10.1109/hicss.2013.22

Dror, D. M., & Preker, A. S. (Eds.). 2002. *Social reinsurance: A new approach to sustainable community health financing*. World Bank Publications.

Fávaro-Moreira, N. C., Krausch-Hofmann, S., Matthys, C., Vereecken, C., Vanhauwaert, E., Declercq, A., Bekkering, G. E., & Duyck, J. (2016). Risk factors for malnutrition in older adults: A systematic review of the literature based on longitudinal data. *Advances in Nutrition*, *7*(3), 507–522. https://doi.org/10.3945/an.115.011254

Felfernig, A., Isak, K., Kreutler, G., Kruggel, T., & Teppan, E. (2006). Knowledge representations for the interactive selling of financial services. *Information Systems and E-Business Management*, *5*(2), 143–166. https://doi.org/10.1007/s10257-006-0042-9

Gatzert, N., Huber, C., & Schmeiser, H. (2011). On the valuation of investment guarantees in unit-linked life insurance: A customer perspective. *The Geneva Papers on Risk and Insurance – Issues and Practice*, *36*(1), 3–29. https://doi.org/10.1057/gpp.2010.35

Gelder, S., & Johnson, D. (1997). *Long-term care insurance: A market update*. Health Insurance Association of America.

Greenwood, R. M., & Vissing-Jorgensen, A. (2018). *The impact of pensions and insurance on global yield curves*. Harvard Business School Finance Working Paper, 18–109. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3196068

Gupta, A. (2012). Unit Linked Insurance Products (ULIPs)—insurance or investment? *Procedia – Social and Behavioral Sciences*, *37*, 67–85. https://doi.org/10.1016/j.sbspro.2012.03.276

Hagemejer, K. (2018). The adequacy of reformed pension systems. *Social Insurance. Theory and Practice*, *136*(1), 21–47.

Hieber, P., & Lucas, N. (2022). Modern life-care tontines. *ASTIN Bulletin*, *52*(2), 563–589. https://doi.org/10.1017/asb.2022.6

Huebner, S. S., & Black, K. (1982). *Life insurance* (10th ed.). Prentice-Hall.

Hurwitz, A., & Mitchell, O. (2022). *Financial regret at older ages and longevity awareness*. NBER Working Paper, 30696. https://doi.org/10.3386/w30696

James, E., & Vittas, D. (2000). *Annuity markets in comparative perspective: Do consumers get their money's worth?* Policy Research Working Papers 2493. https://doi.org/10.1596/1813-9450-2493

Kemayou, L., Guebou Tadjuidje, F. & Madiba, M. (2011). Tontine et banque en contexte camerounais. *La Revue des Sciences de Gestion*, *3–4*(249–250), 163–170. https://doi.org/10.3917/rsg.249.163

King, J. (2021). Expanding theory-based evaluation: Incorporating value creation in a theory of change. *Evaluation and Program Planning*, *89*, 101963. https://doi.org/10.1016/j.evalprogplan.2021.101963

Kofoworola, O., Dodd, N., Boyano, A., Konstantas, A., Wolf, O., Alessia, L., & Ossola, E. (2019). *Development of EU ecolabel criteria for retail financial products*. European Commission Joint Research Center.

Lambregts, T. R., & Schut, F. T. (2020). Displaced, disliked and misunderstood: A systematic review of the reasons for low uptake of long-term care insurance and life annuities. *The Journal of the Economics of Ageing*, *17*, 100236. https://doi.org/10.1016/j.jeoa.2020.100236

Łyskawa, K. (2004). Starość a ryzyko emerytalne w systemie zabezpieczenia emerytalnego. In H. Worach-Kordus (Ed.), *Ubezpieczenia społeczne i na życie. Stan i perspektywy* (pp. 29–50). Wyższa Szkoła Humanistyczno-Ekonomiczna w Łodzi.

Mahayni, A., & Muck, M. (2017). The benefit of life insurance contracts with capped index participation when stock prices are subject to jump risk. *Review of Derivatives Research*, *20*(3), 281–308. https://doi.org/10.1007/s11147-017-9131-9

Mayne, J. (2015). Useful theory of change models. *Canadian Journal of Program Evaluation*, *30*(2), 119–142. https://doi.org/10.3138/cjpe.30.2.142

Milevsky, M. A. (2006). T*he calculus of retirement income: Financial models for pension annuities and life insurance*. Cambridge University Press.

Milevsky, M. A., & Salisbury, T. S. (2015). Optimal retirement income tontines. *Insurance: Mathematics and Economics*, *64*, 91–105. https://doi.org/10.1016/j.insmatheco.2015.05.002

Mitchell, O. S., & Utkus, S. P. (2012). Lessons from behavioral finance for retirement plan design. In O. S. Mitchell and S. P. Utkus (Eds.), *Pension design and structure: New lessons from behavioural finance*. Oxford University Press.

Myles, J. (2002). A new social contract for the elderly? In G. Esping-Andersen (Ed.), *Why we need a new welfare state* (pp. 130–172). https://doi.org/10.1093/0199256438.003.0005

Nguyen, H. T., Nguyen, H., Nguyen, N. D., & Phan, A. (2018). Determinants of customer satisfaction and loyalty in Vietnamese life-insurance setting. *Sustainability*, *10*(4), 1151. https://doi.org/10.3390/su10041151

Nurittamont, W. (2021). Enhancing the factors' influence on purchasing decision of endowment insurance: Case of testing mediate and moderate variables. *Journal of Management Information and Decision Sciences*, *24*(7), 1–11.

Rajasekhar, D., Kesavan, S., & Manjula, R. (2017). Are our contributory pension schemes failing the poor? *Economic and Political Weekly*, *52*(27), 77–85.

Reinholz, D. L., & Andrews, T. C. (2020). Change theory and theory of change: What's the difference anyway? *International Journal of STEM Education*, *7*(1). https://doi.org/10.1186/s40594-020-0202-3

Ribeiro, P. C. C., Almada, D. S. Q., Souto, J. F., & Lourenço, R. A. (2018). Permanência no mercado de trabalho e satisfação com a vida na velhice. *Ciência and Saúde Coletiva*, *23*, 2683–2692. https://doi.org/10.1590/1413-81232018238.20452016

Rokas, I., & Siafarika, A. (2019). The notion of insurance-based investment products. In P. Marano & I. Rokas (Eds.), *Distribution of insurance-based investment products: The EU regulation and the liabilities* (pp. 3–25). Springer.

Schmidt-Jochmann, C., & Gröbner, M. (2012). Aufbau und Betrieb von shared service centers–eine Versicherungsperspektive. *Controlling and Management*, *56*(3), 58–62. https://doi.org/10.1365/s12176-012-0644-7

Skipper, H. D., & Black, K. (2000). *Life and health insurance* (12th ed.). Prentice Hall.

Sohn, H. (2016). Racial and ethnic disparities in health insurance coverage: Dynamics of gaining and losing coverage over the life-course. *Population Research and Policy Review*, *36*(2), 181–201. https://doi.org/10.1007/s11113-016-9416-y

Stolze, M., Field, S., & Kleijer, P. (2000). *Combining configuration and evaluation mechanisms to support the selection of modular insurance products*. 8th European Conference on Information Systems, 858–865.

Super N., Burstein A., Davis J., & Servat C. (2022). Innovative strategies to finance and deliver long-term care. In O. S. Mitchell (Ed.), New *models for managing longevity risk* (pp. 122–149). Oxford University Press. https://www.doi.org/10.1093/oso/9780192859808.003.0008

Szczepański, M., Ratajczak-Leszczyńska, J., Bielawska, K., Rutecka-Góra, J., & Pieńkowska-Kamieniecka, S. (2022). Poland: A pension system under constant (re)construction. In J. Kolaczkowski, M. Maher, Y. Stevens, & J. M. Werbrouck (Eds.), *The evolution of supplementary pensions: 25 years of pension reform* (pp. 229–256). Edward Elgar Publishing.

Turner, J. A. (2010). *Pension policy: The search for better solutions*. WE Upjohn Institute.

Wang, H., Switlick, K., Ortiz, C., Zurita, B., & Connor, C. (2012). *Health insurance handbook: How to make it work* (vol. 219). World Bank Publications.

Wettstein, G., Munnell, A., Hou, W., & Gok, N. (2021). The value of annuities. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3797822

Valentin, E. K. (2001). Swot analysis from a resource-based view. *Journal of Marketing Theory and Practice*, *9*(2), 54–69. https://doi.org/10.1080/10696679.2001.11501891

Yu, J., Qiu, Y., & He, Z. (2018). Is universal and uniform health insurance better for China? Evidence from the perspective of supply-induced demand. *Health Economics, Policy and Law*, *15*(1), 56–71. https://doi.org/10.1017/s1744133118000385

Zelizer, V. A. R. (2017). *Morals and markets. The development of life insurance in the United States*. Columbia University Press. https://doi.org/10.7312/zeli18334

Zhang, X., Zhang, W., Wang, C., Tao, W., Dou, Q., & Yang, Y. (2018). Sarcopenia as a predictor of hospitalization among older people: A systematic review and meta-analysis. *BMC Geriatrics*, *18*(1), 188. https://doi.org/10.1186/s12877-018-0878-0

sciendo Economics and Business Review https://www.ebr.edu.pl Vol. 10 (2), 2024: 31–59 https://doi.org/10.18559/ebr.2024.2.1177

Examining the performance of Shari'ah--compliant versus conventional stock indexes: A comparative analysis pre-, during, and post-COVID-19

Ahmad M. Abu-Alkheil[1] Nizar M. Alsharari[2]
Walayet A. Khan[3] Sara R. Ramzani[4]
Phungmayo Horam[5]

Abstract

This study aims to conduct an empirical comparative analysis of the performance of Shari'ah and conventional stock indexes during the period 2017–2023, which includes the COVID-19 pandemic. Additionally, it aims to investigate investors' preferences and analyse the long-term relationship of these indexes, as well as exploring the potential diversification benefits. The research methodology incorporates stochastic dominance analysis, the VARMAX procedure, and Johansen's co-integration approach. The data utilized consists of 31 conventional and 31 Islamic stock indexes, specifically from the FTSE, DJ, MSCI, and S&P series.

The results show that there are no long-term co-integration links between 30 out of 31 pairs of Islamic and conventional indexes. While conventional indexes tend to outperform Islamic indexes, they also come with a higher

Keywords

- Islamic finance
- indexes
- stochastic dominance
- COVID-19

[1] GISMA University of Applied Sciences, Multidisciplinary Research Centre for Innovations in SMEs (MrciS), Konrad-Zuse-Ring 11, 14469 Potsdam, Germany, corresponding author: ahmad.abu-alkheil@gisma.com, https://orcid.org/0000-0001-7432-9544.

[2] Jackson State University ,1400 John R. Lynch St. Jackson, MS 39217, USA, nizar.m.alsharari@jsums.edu, https://orcid.org/0000-0002-4494-8130.

[3] Schroeder Family School of Business Administration, University of Evansville, 1800 Lincoln Avenue, Evansville, IN 47722, USA, wk3@evansville.edu, https://orcid.org/0000-0003-4496-7646.

[4] GISMA University of Applied Sciences, Multidisciplinary Research Centre for Innovations in SMEs (MrciS), Konrad-Zuse-Ring 11, 14469 Potsdam, Germany, sara.ramzani@gisma.com, https://orcid.org/0000-0002-5336-7223.

[5] GISMA University of Applied Sciences, Multidisciplinary Research Centre for Innovations in SMEs (MrciS), Konrad-Zuse-Ring 11, 14469 Potsdam, Germany, phungmayo.horam@gisma.com, https://orcid.org/0000-0003-1621-8736.

risk. On the other hand, Islamic indexes are considered to be less risky, offering potential diversification opportunities that may be attractive for global portfolios, particularly during periods of financial distress.

JEL codes: C14, Z120, G17

Article received 19 February 2024, accepted 12 May 2024.

Suggested citation: Abu-Alkheil, A. M., Alsharari, N. M., Khan, W. A., Ramzani, S. R., & Horam, P. (2024). Examining the performance of Shari'ah-compliant versus conventional stock indexes: A comparative analysis pre-, during, and post-COVID-19. *Economics and Business Review*, *10*(2), 31–59. https://doi.org/10.18559/ebr.2024.2.1177

Introduction

Islamic finance is grounded in Islamic law, which incorporates moral, ethical, and social dimensions. Screen-based Islamic investing, falling under the category of Socially Responsible Investing (SRI), involves Shari'ah-compliant mutual funds that avoid investing in companies involved in tobacco, gambling, war equipment, and excessive pollution. Faith-based mutual funds, such as Christian and Islamic funds, operate based on religious beliefs and value systems. Islamic mutual funds, in particular, refrain from investing in financial companies like conventional banks due to the prohibition of interest in Islam, as well as in companies involved in unethical sectors such as pornography. Additionally, speculative investing and trading involving excessive uncertainty are not permissible in Islamic investing (Grira et al., 2016; Mumtaz et al., 2015).

Islamic financial products are registering substantial growth. The Islamic finance industry has an estimated value of approximately $2.2 trillion (Standard Chartered, 2023), and industry experts predict that the asset base could grow to $ 4.94 trillion by 2025. Capital inflow into Islamic exchange traded funds is an important development triggering this growth. However, there are also other factors, including the fact that Islamic investing was relatively more immune to the global financial crisis of 2008 (Arouri et al., 2013; Ibrahim, 2015). The total assets of the Islamic finance industry grew 10.6% in 2020 and reached a significant milestone in 2021, according to the S&P Global Ratings and Refinitiv reports (Standard and Poor's, 2022). This growth occurred amidst global challenges stemming from the COVID-19 pandemic and falling oil prices.

The current century has been marked by numerous global events, with COVID-19 having a significant impact on the financial and economic aspects of life. The increased uncertainty resulting from the pandemic has prompted researchers to investigate the effects of COVID-19 on different asset returns, risk parameters, and the overall economy. Several studies found a negative impact of these shocks on financial markets, stocks performance, global economy, and level of risk (Anh & Gan, 2020; Angosto-Fernández & Ferrández-Serrano, 2022; Aslam et al., 2020; Goodell, 2020; Zhang et al., 2020).

Islamic finance attracts interest from researchers due to its unique features, and researchers have been studying various aspects of Islamic finance since its inception in 1975, while comparing it with conventional finance models.[6] Most recently, COVID-19 has generated new interest among academics into whether screen-based Islamic stocks performed differently to conventional stocks. Cheong (2021) reveals that Shari'ah-compliant stocks would be safer due to several restrictions imposed on Islamic investing, which include limits on leverage and prohibition on speculative income. In a review study, Sherif (2020) found that during times of market uncertainty, like the COVID-19 pandemic, Islamic investing attracts increased attention from investors. Nevertheless, empirical research reveals conflicting evidence.

One stream of research concludes that the COVID outbreak produced similar reactions in both groups of indexes in terms of the decline in stock valuations (Hassan et al., 2021a) and the increase in volatility (Hasan et al., 2021b). However, another stream of research found higher market valuation, lower volatility, and faster recovery during post-COVID period among Shari'ah-compliant stock indexes (Chowdhury et al., 2021; Dharani et al., 2022).

The growth in Islamic financial products has led to the development of numerous new indexes to cater to the increasing interest and investment in screened Islamic stocks. The inclusion of a stock in an Islamic stock index involves a screening process that aligns with fundamental Islamic principles concerning business, financial, and accounting practices (Tahir & Ibrahim, 2020; Tanin et al., 2021).

This study intends to analyse individual Shari'ah-compliant and non-compliant (conventional) stocks from the well-known FTSE, DJ, MSCI, S&Ps and Jakarta series. The advantage of focusing on indexes is that it avoids distortion introduced by the transaction costs of the funds, the effects of managerial skill sets, and their timing activities (Ashraf, 2013). Recent research has demonstrated the value of Islamic investing.

AlKhazali et al. (2022) found the existence of size anomalies in Islamic stocks contradicting the efficient market hypothesis. They conclude that the inclusion of Islamic stocks can produce abnormal returns for the investors. Saiti

[6] The Dubai Islamic Bank became the first modern commercial Islamic bank in the world in 1975 (Shaikh, 2010).

et al. (2014) demonstrate that including Islamic stocks in conventional stock indexes produces diversification benefits stemming from the low correlation between two sets of stocks—conventional and Islamic.

In this study, robust statistical techniques, specifically stochastic dominance (SD), are employed. Unlike previous studies that primarily relied on mean-variance (MV) criteria or the capital asset pricing model (CAPM), which utilize parametric statistics and assume normality in the data (despite the presence of fat tails in return data), SD techniques offer distinct advantages. SD techniques, as proposed by Davidson and Duclos (2000), provided a more comprehensive approach to evaluating the performance of mutual funds and portfolio selections. These techniques consider higher moments of the data, capturing key information that is often overlooked by traditional approaches relying solely on the first two moments. By utilizing SD techniques, this study overcomes the limitations of previous research and provides a more robust evaluation of investment performance.

Our research contributes to the existing literature as follows. First, while many studies examined the impact of COVID-19 on different markets and assets, these studies were confined to conventional investing securities. Very few studies have addressed the impact of COVID-19 on the performance of Shari'a--compliant indexes. Second, the research design of this study involves evaluating the comparative performance of conventional indexes versus alternative Islamic indexes, encompassing the most reputable Islamic indexes and up-to-date data from 2017 to 2023. Third, to assess the impact of COVID-19, and the structural changes that occurred in this period, the time span was split into three windows, namely, before, during and after the shock periods, utilizing robust statistical techniques (SD) based on monthly returns. Fourth, the findings of this paper have important implications for investors and portfolio managers, as it examines the potential diversification opportunities that may arise from Islamic indexes. Fifth, given the unique nature of Islamic investing, our study will be of interest to investors, professional money managers, analysts, and policy makers who value the empirical evidence related to different dimensions of alternative Islamic funds.

The overall results suggest that there are no long-term co-integration links between Islamic and conventional indexes. Although conventional indexes outperform Islamic indexes, they also carry higher levels of risk. In contrast, Islamic indexes offer potential diversification opportunities that may be attractive for global portfolios, particularly during times of financial turmoil.

The remainder of this study is as follows: Section 1 discusses the relevant existing literature followed by a detailed explanation of the methodology utilized in Section 2. Section 3 critiques the results and findings of the study, while the final section presents the conclusion.

1. Literature review

The COVID-19 pandemic had a profound impact on every aspect of the economy and finance, prompting academics and practitioners to explore different dimensions of this phenomenon. The existing literature highlights several key areas of research in this field. Firstly, there are studies that analyse the performance of traditional stock indexes. Secondly, researchers have examined the comparative behaviour of returns and volatility between Islamic and conventional stock indexes, including sectoral indexes. Thirdly, there are studies that investigate the impact at the company level, focusing on company-specific characteristics. Lastly, some studies have also explored the effects on regional and country indexes.

Studies that focused on market performance and conventional stock indexes found a negative association between returns and the daily frequency of COVID cases (Al-Awadhi et al., 2020; Anh & Gan, 2020; Topcu & Gulal, 2020; Yarovaya et al., 2021). On the other hand, studies that examined the comparative performance of Islamic and conventional indexes during COVID-19 period revealed mixed results. Hasan et al. (2021a) analysed the comparative performance of a comparable pair of indexes—the Islamic Dow Jones index and FTSE index—over the period of 21st January to 27th November 2020 and found a similar impact on returns volatility. Hasan et al. (2021b) used a global data of 50 countries to evaluate the performance of MSCI conventional equity indexes versus Islamic indexes from January 1 through September 30, 2020. Overall, their results showed a similar decline in returns. Exceptions were a few Asian countries where Islamic indexes yielded superior returns. Sherif (2020), however, showed different results. He found a significant negative impact on the conventional UK stock index, while an insignificant impact was observed on the UK Dow Jones Islamic index.

Moreover, studies that examined the comparative sectoral indexes found similar results. For example, Chowdhury et al. (2021) focused on sectoral indexes data including emerging countries data covering a sample period from January 1, 2020, to August 15, 2020. They report that Islamic sectoral indexes exhibited a relatively smaller decline in returns. However, their recovery was quicker than their counterpart conventional Dow Jones indexes.

Based on a study conducted in the period of September 1, 2019–April 30, 2020, and utilizing daily data from 15 countries by Nomran and Haron (2021), it was found that the COVID-19 pandemic had a negative impact on both Islamic and conventional stock market returns. However, conventional indexes showed a negative trend throughout the sample period, whereas the returns of Islamic indexes had produced a positive trend by mid-April 2020. The study also revealed that the impact of the pandemic was weaker on Islamic stock markets compared to conventional ones. The overall findings of the studies suggest that be-

fore and during the pandemic Islamic indexes performed better, and in the post period they showed a lesser adverse impact compared to conventional ones.

Ding et al. (2020) and Heyden and Heyden (2020) ran firm level analysis. Ding et al. (2020) collected the data for 6,700 companies from 61 countries and found that the negative impact of COVID-19 was smaller for firms with a higher liquid asset base, higher profits but less leverage. Identifying factors that impact profitability can play a crucial role in formulating regulatory policies aimed at promoting the stability and sustainable performance of companies listed on Islamic indexes at the broader level. By gaining insights into these influential factors, policymakers can make informed decisions and implement effective measures to foster long-term profitability and stability within the Islamic finance industry. This approach creates a facilitating environment for companies listed on Islamic indexes to flourish, thus contributing to the overall growth and sustainability of the Islamic finance sector (Alsharari & Alhmoud, 2019). On the other hand, Heyden and Heyden (2020) found that stocks produced significantly negative returns in the COVID period. These returns, however, were a function of company-specific characteristics.

Sherif (2020) conducted a study comparing the Islamic Dow Jones Index with its conventional counterpart in the UK. The analysis used daily data from January 20 to May 20, 2020, focusing on ten industrial sectors. Results showed a significant link between the conventional stock market index performance and the COVID-19 pandemic. While the Dow Jones Islamic Index was negatively impacted, the effect was not statistically significant compared to its UK counterpart. This study provides insights into the relationship between the pandemic and the performance of Islamic and conventional stock market indexes in the UK during the specified period.

In his study, Erdoğan et al. (2020) utilized the DCC-GARCH method to analyse daily data from February 10, 2011, to September 2, 2020, focusing on the Islamic and conventional stock indexes in Turkey. The objective was to investigate the impact of COVID-19 returns on these indexes. His results showed that Islamic indexes yielded greater stability compared to the conventional indexes.

Salisu and Sikiru (2020) focused on the regional index of Asia-Pacific Islamic versus a conventional index to explore the hedging potential against uncertainty over the pre-COVID-19 sample (8.31.2010 to 12.31.2019) and the period of COVID-19 pandemic (1.09.–15.09.2020). They found the Islamic index to be a better choice for hedging purposes.

Another stream of research examined whether Islamic funds can serve as safe-haven assets due to the spillover effect from one index to another. Yarovaya et al. (2021) showed a stronger spillover from conventional stock indexes to Islamic indexes during the pandemic. However, Sukuk (Islamic bonds) were less impacted by the spillover from conventional bonds. These results show that Sukuk can be used as a hedge against conventional bonds. On the other hand, Arif et al. (2021) found that Islamic stocks exhibited safe-haven

asset characteristics during the pandemic. They examined the comparative performance of conventional and Islamic indexes of G7 countries.

By and large, past studies employed parametric methods, a short sample period, crisis, or prosperous period only, and did not fully cover the major global stock indexes. This study fills the gap in the literature by using a non-parametric methodology (free of limitations of parametric methodologies), covering the major global Islamic and conventional stock indexes. It also statistically examines indexes' performance over three sub-periods, namely pre, during and post COVID-19 pandemic.

2. Data and methods

This research employs quantitative research methodologies. A sample of 31 Islamic indexes along with 31 conventional indexes (Table 1) were selected and categorized into four primary clusters: Dow Jones, FTSE, MSCI, and S&P indexes. All related data was obtained from Data Stream.

The selection of the indexes sample was not arbitrary. For accurate comparison between Islamic and conventional stock indexes, it is important to consider factors like time dimension and exchange rate fluctuations. This study addresses time-zone bias by selecting trading days when constituent stocks for both indexes were actively traded. The sample period ranges from January 1, 2017, to October 31, 2023, enabling a thorough analysis of index performance. Additionally, using the US dollar as the common currency across all countries ensures comparability in the data set.

Furthermore, to ensure that the Islamic index is comparable to its counterpart the index provider considers each conventional index separately to make sure that it serves as a comparison benchmark for its Islamic counterpart index. The pair of indexes are developed to obtain the largest coverage by country over time. Data constructed by Data Stream enhances cross-country comparability.

Our full sample begins from January 1, 2017 and ends on October 31, 2023. It is intentionally split into three sub-periods. The first period, January 1, 2017 to February 28, 2020 represents the pre-COVID-19 pandemic period. The second period, from March 2, 2020 to April 30, 2022 corresponds to the COVID-19 period, providing an opportunity to assess the impact of the global pandemic on both the Islamic and conventional indexes.[7] The third period, from May 2, 2022 to October 31, 2023 represents the post-COVID-19 pan-

[7] On March 11, 2020, the World Health Organization (WHO) declared COVID-19 a pandemic. This designation indicated that the virus had spread globally and posed a significant threat to public health across multiple countries and continents (Mishra et al., 2020).

Table 1. List of Islamic and conventional indexes

#	Islamic index	Conventional index
1	DJ ISLAMIC ASIA/PACIFIC	DJGL ASIA PACIFIC $
2	DJ ISLAMIC Market EUROPE	DJ EUROPE TOTAL STOCK MKT
3	DJ ISLAMIC CANADIAN	DJGL CANADA DJTM CANADA DEAD
4	DJ ISLAMIC JAPAN $	DJGL JAPAN $
5	DJ ISLAMIC MKT CHINA HK TITANS 30 $	DOW JONES CHINA 88
6	DJ ISLAMIC US	DJ US TOTAL STOCK MKT
7	DJ ISLAMIC US LARGE CAP	DJ US LARGE CAP TOTAL STOCK MKT
8	DJ ISLAMIC US MID CAP	DJ US MID CAP TOTAL STOCK MKT
9	DJ ISLAMIC US SMALL CAP	DJ US SMALL CAP TOTAL STOCK MKT
10	DJISLAMIC MARKET INDEX	DJGLOBALINDEX
11	FTSE BURSA MALAYSIA HIJRAH SHARI'AH $	FTSE BURSA MALAYSIA EMAS $
12	FTSE GWA P SHARI'AH DEV $	FTSE GWA ALL-WORLD $
13	FTSE SHARI'AH DEV ASIA PACIFIC	FTSE DEV ASIA PAC $
14	FTSE SHARI'AH DEVELOPED	FTSE GWA DEVELOPED $
15	FTSE SHARI'AH JAPAN 100 $	FTSE W JAPAN
16	FTSE SHARI'AH MULT 150	FTSE MULTINATIONALS ($)
17	FTSE SHARI'AH USA $	FTSE US $
18	MSCI AC AMERICAS IS U$	MSCI AC AMERICAS U$
19	MSCI AC ASIA IS U$	MSCI AC ASIA PACIFIC U$
20	MSCI AC ASIA PACIFIC IS U$	MSCI AC ASIA U$
21	MSCI AC EUROPE IS U$	MSCI AC EUROPE U$
22	MSCI AC FAR EAST IS U$	MSCI AC FAR EAST U$
23	MSCI AC PACIFIC IS U$	MSCI AC PACIFIC U$
24	MSCI AC WORLD IS U$	MSCI AC WORLD U$
25	MSCI GOLDEN DRAGON IS U$	MSCI GOLDEN DRAGON U$
26	MSCI INDIA IS U$	MSCI INDIA
27	MSCI MALAYSIA IS	MSCI MALAYSIA
28	MSCI PAKISTAN IS U$	MSCI PAKISTAN
29	MSCI ZHONG HUA IS U$	MSCI ZHONG HUA U$
30	S&P 500 SHARI'AH $	S&P 500 COMPOSITE
31	S&P JAPAN 500 SHARI'AH $	S&P JAPAN 500

Source: (Abu-Alkheil et al., 2017).

demic period, which is utilized to examine the speed of recovery and potential long-term effects of the pandemic on the market indexes.[8]

This study employs stochastic dominance (SD) and Johansen's co-integration methods for empirical analysis purposes. The SD analysis provides insights into the preference of risk averters and risk-seekers for one index or the other. Johansen's co-integration approach enables us to examine the long-term association and efficiency of conventional indexes as well as the potential diversification benefits and portfolio optimization opportunities. Previous studies have used this approach (Abu-alkheil et al., 2017; Fan, 2003; Fong et al., 2005; Hodges & Yoder, 1996; Lean et al., 2007; Masih & Masih, 2001; Sheng & Tu, 2000) and provide insights into the different sets of investments.

The stochastic dominance method is used to evaluate the likelihood of which of the two returns distributions, in our case, conventional indexes versus Islamic indexes, will produce higher returns for a given level of risk, providing an opportunity to investors to select the optimal portfolio.

Stochastic dominance is a non-parametric method useful for investigating the risk-return characteristic of portfolios. It is free of any risk-free rate assumption, which is ingrained in standard models such as CAPM, making it particularly well-suited for "screen-based" Islamic stocks, where the concept of a risk-free rate is not applicable. Also, unlike CAPM, it does not rely on any specific distributional assumptions. In general, the use of SD methodology allows researchers to gain valuable insights into the risk and return trade-offs of different investment portfolios by overcoming the challenges of estimating asset pricing models that are distribution-specific and use a risk-free rate.

In summary, stochastic dominance (SD) is a statistical tool that enables the comparison of risk and return profiles across various investment choices, aiding investors in selecting the most suitable option based on their risk appetite and return preferences. On the other hand, Johansen's co-integration method is a statistical approach utilized to examine the enduring relationships among multiple time series variables. This method helps researchers determine the presence of a consistent relationship between these variables over time, facilitating predictive analysis and enhancing comprehension of the data's fundamental dynamics.

2.1. Co-integration analysis

Regression analysis has traditionally been a popular method for studying the relationship between financial variables. However, it is important to note that financial variables often exhibit non-stationary behaviour, and applying

[8] As of May 5, 2023, the WHO announced the end of the global Public Health Emergency (PHE) for COVID-19 (WHO, 2023).

regression analysis to non-stationary time series can lead to spurious relationships. In this study, the Engle & Granger (1987) co-integration approach is employed instead in order to investigate whether Islamic and conventional indexes are co-integrated. This approach allows for the exploration of potential diversification benefits and opportunities for portfolio optimization. By utilizing co-integration analysis, this study aims to provide a more robust and accurate understanding of the relationship between Islamic and conventional indexes.[9]

The Augmented Dickey-Fuller (ADF) test is commonly used to investigate the presence of stationarity or non-stationarity of data. The ADF test estimates a general specification controlling for autocorrelation and is based on lagged values, a time trend, the differenced variable, and a residual term. The equation is as follows:

$$\Delta x_t - \alpha + \beta_t + \gamma x_{t-1} + \sum_{i=1}^{P} (\delta_1 \Delta x_{t-1}) + \varepsilon_t \quad (1)$$

Here, X represents the variable under consideration (for stationarity), Δ denotes the difference operator, α is an intercept constant called a drift, β is the coefficient on a time trend, γ is the coefficient presenting process root, i.e. the focus of testing, p is the lag order of the first-differences autoregressive process, and ε is the residual term (Khamlichi et al., 2014). If co-integration is found, the VECM (Vector Error Correction Model) is used to assess the long-term (a stable equilibrium) relationships between variables; however, this model does not determine the causal direction between them.

To evaluate the direction of causality (from X to Y or Y to X or both), the error correction term (ECT) is analysed, which estimates how quickly variables, after being shocked, move back to their long-term equilibrium. Specifically, a statistically insignificant ECT suggests that the corresponding dependent variable is unaffected by the other variables in the system. That is, it is an exogenous variable in the long run. On the other hand, if the ECT is found to be statistically significant, it shows that the dependent variable is endogenous. That is, it is impacted by the other variables in the system.

After observing the presence of co-integration among the indexes, this study applies the VECM to determine the endogeneity or exogeneity of the variables, thus enabling us to understand the long-term relationships and dynamics between them.

[9] Islamic indexes derived from conventional ones (Saiti, 2014) may show a high correlation but not necessarily high cointegration (Khamlichi et al., 2014). Co-integration indicates long-term equilibrium between indexes, offering predictive insights. Understanding cointegration is essential for grasping the dynamics among Islamic indexes (Masih et al., 2010).

2.2. Stochastic Dominance (SD) analysis and risk preferences

Stochastic dominance (SD) offers value to ascertain the investors' risk preferences when the data exhibits non-normal distribution (which is often the case with financial time series) and/or when it is complex and challenging to define specific utility functions. The presence of a negative or positive domain of the return distribution determines the shape of utility function (Levy & Levy, 2003; Levy & Wiener, 1998; Post et al., 2008; Wong & Chan, 2007). If stochastic dominance is convex, it shows risk-seeking behaviour, whereas if it is concave, it shows risk-averse behaviour. If the SD utility function is S-shaped, it demonstrates risk-seeking behaviour below the desired level of return/risk, but it exhibits risk-aversion behaviour beyond the aspired level of return/risk. On the other hand, a reverse S-shaped utility function suggests a particular pattern of risk preferences. In the positive domain, where outcomes are favourable, individuals exhibiting a reverse S-shaped utility function would display risk-seeking behaviour. This means they are more inclined to take on higher levels of risk in pursuit of potentially greater gains. On the other hand, in the negative domain where outcomes are unfavourable, individuals with a reverse S-shaped utility function would exhibit risk-aversion. This implies a preference for avoiding or mitigating losses by opting for less risky alternatives.

Overall, the shape of the utility function influences risk preferences and decision-making behaviour in different domains (Clark et al., 2015; Denuit et al., 2013). For any investment level, if the expected return of X is greater than or equal to the expected return of Y, this implies it is first-order stochastic dominance (FSD). This would mean that stock X is expected to yield at least as great a return as stock Y.

This study compares the entire return distribution of X with Y to determine second-order stochastic dominance (SSD). If the cumulative distribution function (CDF) of X lies above the CDF of Y, regardless of the level of investment, it would indicate that X stochastically dominates Y and thus has a higher probability of achieving higher returns than Y.

In the third-order stochastic dominance (TSD), this study focuses on the entire distribution of returns together with the higher moments of the returns, as this series imposes a more stringent criteria. Under TSD, X stochastically dominates Y if two conditions are met: For any investment level (1), the CDF of X lies above the CDF of Y (2) and the higher moments of X, such as skewness and kurtosis, are more favourable than those of Y. The performance of different portfolios can be compared with the use of these three orders of stochastic dominance. Portfolio X is considered to perform better than Y if it stochastically dominates portfolio Y in all three orders, and vice versa.

The comparative performance of equity portfolios can be determined by assessing the stochastic dominance of different pairs of return series. For as-

sessment purposes, this study uses the nonparametric SD test proposed by Davidson and Duclos (2000). As discussed earlier, this method offers a robust and distribution-free approach, while considering higher moments of the return distribution (Cheong et al., 2007) to evaluate portfolio performance:

$$X \succ_1 Y \Rightarrow P(X \leq A) \leq P(Y \leq A) \tag{2}$$

In the given context, the notation $P(X \leq A)$ represents the cumulative distribution function (CDF) of variable X, while $P(Y \leq A)$ represents the CDF of variable Y. It is important to note that first-order stochastic dominance (FSD) does not rely on any assumptions about the decision-maker's utility function under uncertainty. FSD assumes that the CDF is a continuous function that monotonically increases with an increasing X (non-decreasing) (Whitmore & Findlay, 1978).

Furthermore, the research paper can have significant implications for the distribution of stock returns, trading volume, and demand for risky assets by considering the observed first-order risk aversion and the influence of investors' risk attitudes, which are influenced by the past performance of investments (Gomes, 2005). These implications can provide valuable insights into understanding the behaviour of investors and the dynamics of financial markets.

Moreover, the condition for a second-order stochastic dominance (SSD) constraint can be expressed by the formula:

$$X \succ_2 Y \Rightarrow \int_{-\infty}^{A} P(X \leq B)dB \leq \int_{-\infty}^{B} P(Y \leq B)dB \tag{3}$$

Second-order stochastic dominance (SSD) necessitates the presence of non-stationarity in time series. Additionally, it assumes that the decision-maker possesses a utility function that is inclined towards risk aversion. It is important to note that first-order stochastic dominance (FSD) is typically a prerequisite for second-order dominance (of X over Y). While there is limited emphasis on third-order dominance conditions (TSD) in the existing literature, it is still feasible to derive TSD.

3. Results and discussion

The descriptive statistics for the returns of the 62 indexes for the entire period are reported in Table 2. The descriptive results show that the means and standard deviations of the returns vary significantly across the 62 indexes. The returns for each sector are negatively skewed, indicating that they are flatter on the left side compared to a normal distribution. The kurtosis,

Table 2. Descriptive statistics of monthly returns for Islamic and conventional indexes in the period 2017–2023

Conventional indexes	Mean	Standard deviation	Skewness	Kurtosis	Islamic indexes	Mean	Standard deviation	Skewness	Kurtosis
DJGL ASIA PACIFIC $	0.0028	0.04066	–1.077	4.222	DJ ISLAMIC ASIA/PACIFIC	–0.001	0.0400	–1.160	4.884
DJ EUROPE TOTAL STOCK MKT	0.0010	0.04566	–0.881	3.412	DJ ISLAMIC Market EUROPE	–0.001	0.0320	–1.152	5.021
DJGL CANADA DJTM CANADA DEAD	0.0019	0.04722	–1.000	3.661	DJ ISLAMIC CANADIAN	–0.006	0.0442	–1.185	5.364
DJGL JAPAN	0.0009	0.05275	–0.954	2.766	DJ ISLAMIC JAPAN	–0.004	0.0473	–1.070	4.336
DOW JONES CHINA 88	0.0016	0.04455	–0.881	2.211	DJ ISLAMIC MKT CHINA HK TITANS 30	–0.002	0.0410	–1.142	6.441
DJ US TOTAL STOCK MKT	0.0020	0.04670	–1.001	3.541	DJ ISLAMIC US	0.0020	0.0447	–1.182	5.002
DJ US LARGE CAP TOTAL STOCK MKT	0.0027	0.05804	–0.754	2.622	DJ ISLAMIC US LARGE CAP	0.0006	0.0522	–0.901	3.876
DJ US MID CAP TOTAL STOCK MKT	0.0056	0.07122	–0.822	1.561	DJ ISLAMIC US MID CAP	0.0057	0.0657	–0.700	1.764
DJ US SMALL CAP TOTAL STOCK MKT	0.0005	0.04512	–1.321	7.333	DJ ISLAMIC US SMALL CAP	0.0005	0.0292	–1.183	6.766
DJGLOBALINDEX	0.0042	0.08221	–0.546	5.561	DJISLAMIC MARKET INDEX	0.0086	0.0603	–0.562	2.243
FTSE BURSA MALAYSIA EMAS	0.0042	0.05115	–0.541	7.342	FTSE BURSA MALAYSIA HIJRAH SHARI'AH	0.0078	0.0366	–0.644	3.226
FTSE GWA ALL-WORLD	0.0052	0.09732	–2.555	8.666	FTSE GWA P SHARI'AH DEV	0.0088	0.0703	–2.173	9.563
FTSE DEV ASIA PAC $	0.0040	0.05011	–0.891	2.400	FTSE SHARI'AH DEV ASIA P.	0.0046	0.0503	–0.884	2.555
FTSE GWA DEVELOPED	0.0036	0.04101	–0.900	1.901	FTSE SHARI'AH DEVELOPED	0.0044	0.0400	–0.843	1.848

Conventional indexes	Mean	Standard deviation	Skewness	Kurtosis	Islamic indexes	Mean	Standard deviation	Skewness	Kurtosis
FTSE W JAPAN $	0.0070	0.05299	–0.967	3.301	FTSE SHARI'AH JAPAN 100	0.0066	0.0527	–1.014	2.334
FTSE MULTINATIONALS ($)	0.0068	0.05702	–0.571	1.600	FTSE SHARI'AH MULT 150	0.0055	0.0544	–0.588	1.202
FTSE US $	0.0030	0.04886	–0.781	4.102	FTSE SHARI'AH USA	0.0044	0.0448	–1.111	3.768
MSCI AC AMERICAS U$	0.0041	0.08712	–1.212	5.522	MSCI AC AMERICAS	0.0033	0.0611	–1.155	3.425
MSCI AC ASIA PACIFIC U$	–0.0010	0.04256	–1.144	4.900	MSCI AC ASIA IS	0.0022	0.0341	–0.446	0.774
MSCI AC ASIA U$	0.0025	0.04341	–0.456	0.801	MSCI AC ASIA PACIFIC	0.0011	0.0432	–0.653	1.543
MSCI AC EUROPE U$	0.0042	0.04322	–0.761	1.823	MSCI AC EUROPE	0.0055	0.0431	–0.867	6.164
MSCI AC FAR EAST U$	0.0031	0.07011	–0.322	3.111	MSCI AC FAR EAST	0.0012	0.0306	–0.589	3.744
MSCI AC PACIFIC U$	0.0040	0.04333	–0.566	5.104	MSCI AC PACIFIC	0.0046	0.0423	–1.144	4.255
MSCI AC WORLD U$	0.0022	0.04734	–1.100	2.934	MSCI AC WORLD	0.0004	0.0415	–1.162	5.707
MSCI GOLDEN DRAGON U$	0.0010	0.04155	–0.988	4.602	MSCI GOLDEN DRAGON	0.0002	0.0338	–0.644	3.262
MSCI INDIA	0.0011	0.04622	–0.544	1.444	MSCI INDIA	0.0005	0.0418	–0.672	4.482
MSCI MALAYSIA	0.0032	0.04121	–0.871	2.403	MSCI MALAYSIA	0.0020	0.0340	–0.801	8.412
MSCI PAKISTAN	0.0005	0.01065	1.661	1.704	MSCI PAKISTAN	0.0011	0.0067	–1.001	5.154
MSCI ZHONG HUA U$	–0.0050	0.05019	–1.551	4.007	MSCI ZHONG HUA	–0.0080	0.0401	–0.833	3.441
S&P 500 COMPOSITE	0.0030	0.02891	–0.945	3.003	S&P 500 SHARI'AH	0.0017	0.0270	–1.071	7.436
S&P JAPAN 500	0.0095	0.06431	–1.401	9.012	S&P JAPAN 500 SHARI'AH	0.0158	0.0253	–1.201	6.383

Source: own work.

which measures the degree of excess, is higher than 3 for 26 Islamic indexes and 24 conventional indexes, suggesting a leptokurtic distribution. However, there are 6 Islamic indexes and 8 conventional indexes with kurtosis values lower than 3, indicating a platykurtic distribution.

Previous studies on equity markets in both developed and emerging markets have also documented excess kurtosis in stock returns. These findings suggest that the return distributions of the 31 indexes are not normal. Therefore, this study utilizes non-parametric SD analysis, as it does not rely on the assumption of normality (Worthington & Higgs, 2004).

3.1. Unit Root Test and returns behaviour

Table 3 presents the outcomes of the ADF tests for non-stationarity and first differences. It contains an example of 62 tests on 31 pairs of Islamic and

Table 3. Sample results of the ADF test for non-stationarity and first differences of each series—a Unit-Root Test over the period 2017–2023

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Dependent variable: DJ ISLAMIC ASIA/PACIFIC							
Zero mean	2	0.2083	0.642	0.5312	0.7607		
Single mean	2	–1.746	0.6831	–0.584	0.7461	0.5005	0.8734
Trend	2	–6.554	0.5547	–2.025	0.5621	2.7361	0.6076
Dependent variable: DJGL ASIA PACIFIC							
Zero mean	2	0.2275	0.604	0.5633	0.6714		
Single mean	2	–0.530	0.7152	–0.334	0.7274	0.2152	0.6678
Trend	2	–5.624	0.6435	–1.746	0.5762	3.5240	0.4672
Dependent variable: first difference DJ ISLAMIC ASIA/PACIFIC							
Zero mean	2	–56.32	<0.0001	–4.726	<0.0001		
Single mean	2	–56.55	0.0006	–4.761	0.0003	10.6754	<0.0010
Trend	2	–63.24	0.0001	–5.024	0.0005	11.445	<0.0010
Dependent variable: first difference DJGL ASIA PACIFIC							
Zero mean	2	–34.34	<0.0001	–3.554	0.0002		
Single mean	2	–54.55	0.0006	–3.726	0.0048	6.7234	<0.0010
Trend	2	–46.66	0.0001	–3.654	0.0271	6.7143	0.0246

Source: own work.

conventional indexes, using both level and first difference data. The aim of these tests is to ascertain if the time series of these indexes display stationarity or non-stationarity (Hendry & Juselius, 2001).

In every test, the study does not find sufficient evidence to reject the null hypothesis of a unit root, suggesting non-stationarity (unit root) in the time series. This is because the calculated p-values are above the 5% significance level. However, when considering the first difference, the null hypothesis of a unit root is rejected, as the computed p-values are lower than the significance level. These findings suggest that the Islamic and conventional indexes are non-stationary, but they become stationary after taking the first difference. This allows us to proceed with additional analysis, such as the Johansen tests of co-integration.

The Johansen co-integration framework is utilized to determine whether a co-integrating relationship exists, indicating a long-term equilibrium connection between the paired indexes. Based on the results of the unit root tests, it is suggested that there is no co-integrating relationship among the combinations of paired indexes over the sample period and the three subperiods, except for China and Hong Kong (MSCI Zhong Hua) (as indicated in Table 4). The exception for China and Hong Kong (MSCI Zhong Hua) implies the presence of one co-integrating vector, indicating that the series are linked by common factors. This suggests that these indexes move together in the long term, even if they may exhibit different short-term movements.

The overall findings indicate that there is no long-run equilibrium relationship among the various paired indexes, except for China and Hong Kong. This suggests that including such indexes in a portfolio may provide international diversification benefits and increase the potential for abnormal gains, apart from China and Hong Kong (MSCI Zhong Hua). These results are consistent with previous studies by (Abbes & Trichilli, 2015; Khamlichi et al., 2014; Kok et al., 2009). However, they are in contrast with the findings of Ilhan and Marih (2014).

Moreover, Table 5 shows that there is at least one co-integration equation among Islamic indices in all periods except during the financial crisis. Conversely, there is no co-integration observed among conventional indices in any period. Additionally, the absence of co-integration among conventional indexes implies that they do not move together in the long run, making it difficult to predict one index based on another. This contrasts with the Islamic indices, which exhibit a tendency to move in tandem over time. Prior to and following COVID-19, Islamic indexes remained unaffected; however, during the crisis, these indices appeared to diverge from each other. Furthermore, the results presented in Table 4 indicate a lack of co-integration between Islamic and conventional indexes, suggesting that diversification benefits increased during the crisis period.

Table 4. Johansen Cointegration Rank Test between series using trace—overall period from January 1, 2017 till October 31, 2023

Conventional index	Islamic index	Eigen-value		Trace	Conventional index	Islamic index	Eigen-value	Trace	Conventional index	Islamic index	Eigen-value	Trace
DJGL ASIA PACIFIC $	DJ ISLAMIC ASIA/ PACIFIC	0	0.0166	25.726	FTSE GWA ALL-WORLD $	FTSE GWA P SHARI'AH DEV	0.0371	2.460	MSCI AC PACIFIC U$	MSCI AC PACIFIC IS	0.0107	34.156
		1	0.0003	0.0283			0.0005	0.046			0.0057	12.016
DJ EUROPE TOTAL STOCK MKT	DJ ISLAMIC Market EUROPE	0	0.0235	21.664	FTSE DEV ASIA PAC	FTSE SHARI'AH DEV ASIA PACIFIC	0.0322	3.136	MSCI AC WORLD U$	MSCI AC WORLD IS	0.0422	38.014
		1	0.003	0.2746			0.0153	1.221			0.0008	0.0641
DJGL CANADA DJTM CANADA DEAD	DJ ISLAMIC CANADIAN	0	0.033	32.614	FTSE GWA DEVELOPED	FTSE SHARI'AH DEVELOPED	0.0522	5.406	MSCI GOLDEN DRAGON U$	MSCI GOLDEN DRAGON IS	0.0173	28.161
		1	0.005	0.3251			0.0056	0.461			0.0005	0.1206
DJGL JAPAN	DJ ISLAMIC JAPAN	0	0.016	38.046	FTSE W JAPAN $	FTSE SHARI'AH JAPAN 100	0.0672	7.322	MSCI INDIA	MSCI INDIA IS	0.0164	16.214
		1	0.004	0.5623			0.0124	0.841			0.0014	0.1405
DOW JONES CHINA 88	DJ ISLAMIC MKT CHINA HK TITANS 30	0	0.028	40.204	FTSE MULTI-NATIONALS ($)	FTSE SHARI'AH MULT 150	0.1447	10.615	MSCI MALAYSIA	MSCI MALAYSIA IS	0.0251	46.144
		1	0.001	0.3671			0.0034	0.254			0.0041	0.5224
DJ US TOTAL STOCK MKT	DJ ISLAMIC US	0	0.071	63.155	FTSE US	FTSE SHARI'AH USA	0.0533	3.377	MSCI PAKISTAN	MSCI PAKISTAN IS	0.0706	62.705
		1	0.027	22.164			0.0016	0.126			0.0001	0.0154

DJ US LARGE CAP TOTAL STOCK MKT	DJ ISLAMIC US LARGE CAP	0	0.017	28.436	MSCI AC AMERICAS U$	MSCI AC AMERICAS IS	0.0352	10.124	MSCI ZHONG HUA U$	MSCI ZHONG HUA IS	0.0604	11.132
		1	0.001	0.1433			0.0243	4.031			0.0034	0.6163
DJ US MID CAP TOTAL STOCK MKT	DJ ISLAMIC US MID CAP	0	0.030	44.587	MSCI AC ASIA PACIFIC U$	MSCI AC ASIA IS	0.0213	2.455	S&P 500 COMPOSITE	S&P 500 SHARI'AH	0.0114	13.832
		1	0.008	13.533			0.0056	0.504			0.0013	0.2357
DJ US SMALL CAP TOTAL STOCK MKT	DJ ISLAMIC US SMALL CAP	0	0.028	6.2410	MSCI AC ASIA U$	MSCI AC ASIA PACIFIC IS	0.0437	4.132	S&P JAPAN 500	S&P JAPAN 500 SHARI'AH	0.1343	38.863
		1	0.002	0.3771			0.001	0.074			0.0527	13.561
DJGLOBALINDEX	DJISLAMIC MARKET INDEX	0	0.022	54.433	MSCI AC EUROPE U$	MSCI AC EUROPE IS	0.0267	1.523				
		1	0.007	18.622			0.0016	0.130				
FTSE BURSA MALAYSIA EMAS $	FTSE BURSA MALAYSIA HIJRAH SHARI'AH	0	0.017	45.410	MSCI AC FAR EAST	MSCI AC FAR EAST IS	0.2233	6.688				
		1	0.000	0.0211			0.0002	0.020				

Notes: 1) The discrepancies in the ECM are all designated as "NOINT," suggesting the absence of a constant term in the ECM mechanism. Conversely, a constant term is present in the long-term relationship. 2) In the hypothesis testing scenario, where the null hypothesis (H0) asserts that the rank equals r, and the alternative hypothesis (H1) argues that the rank exceeds r, the index pair value is (0 = 10.32; 1 = 3.19) for all indices at a 5% critical level.

Source: own work.

Table 5. Johansen Cointegration Rank Test between series using trace—average values over the three sub-periods—all indexes

Cointegration	Pre-COVID-19 January 1, 2017–February 28, 2020		During COVID-19 March 2, 2020–April 30, 2022		Post-COVID-19 May 2, 2022–October 31, 2023	
	Max. Eigen	Trace	Max. Eigen	Trace	Max. Eigen	Trace
Islamic indexes	8.81	16.01**	10.34	10.74	14.22*	21.1***
Conventional indexes	9.25	11.28	12.26	11.34	9.26	13.4

Note: *, ** and *** significant at 1%, 5% and 10%. Lag 4 was used for both Islamic and non-Islamic indexes.

Source: own work.

Table 6 presents the Granger causality analysis results for all the indexes. It is evident that there exists a significant bidirectional causal relationship between Islamic indexes and their conventional counterparts throughout the entire period and the three sub-periods, conforming strong lead lag association between the two indexes. This implies that the movement of one market can be predicted by observing the movement of the other market. These findings contribute valuable insights to the current literature by offering new evidence on the causal connections between Islamic and conventional stock indexes. This information provides an additional avenue for investors, both domestic and international, to enhance portfolio diversification and potentially optimize their investment strategies.

Table 6. Granger causality test results for all indexes—overall period and the three sub-periods

Granger causality	Overall period January 1, 2017–October 31, 2023	Pre-COVID-19 January 1, 2017–February 28, 2020	During COVID-19 March 2, 2020–April 30, 2022	Post-COVID-19 May 2, 2022–October 31, 2023
Islamic ≠ conventional	144.1*	67.43*	44.08*	63.4*
Conventional ≠ Islamic	14.77*	463.81*	17.01*	57.48*

Note: *, ** and *** significant at 1%, 5% and 10%.

Source: own work.

3.2. Stochastic dominance results

Based on the context provided, the results from Figure 1 to Figure 4 show stochastic dominance results (SD) for three sub-periods. Specifically, for the first sub-period (January 1, 2017 to February 28, 2020) and the second sub-period (March 2, 2020 to April 30, 2022), the results indicate that the conventional indexes have a higher first-order SD compared to the Islamic indexes (Figure 1–2), respectively.

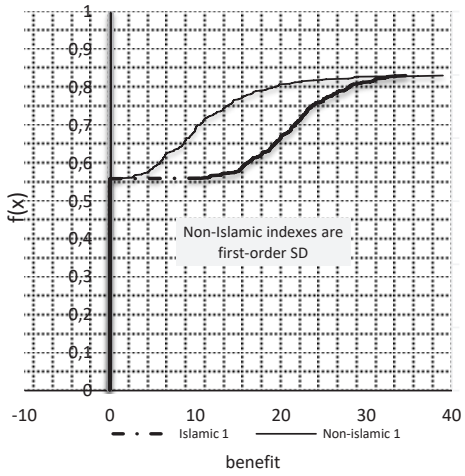

Figure 1. Stochastic dominance results of Islamic and conventional indexes over the period: January 1, 2017–February 28, 2020

Source: own work.

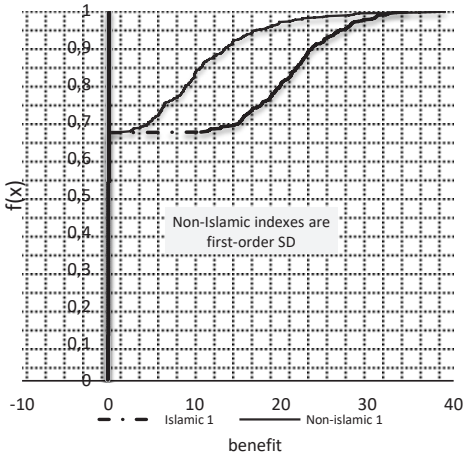

Figure 2. Stochastic dominance results of Islamic and conventional indexes over the period: March 2, 2020–April 30, 2022

Source: own work.

Islamic indexes do not invest in all sectors (industries) in the economy, due to the "screening process". They are typically concentrated in certain sectors. This may contribute to their relatively lower performance compared to conventional indexes. Nevertheless, it is important to note that the superior performance of conventional indexes does not imply that they are inherently superior in terms of their ethical or social impact for socially, ethically oriented, and faith-based investors.

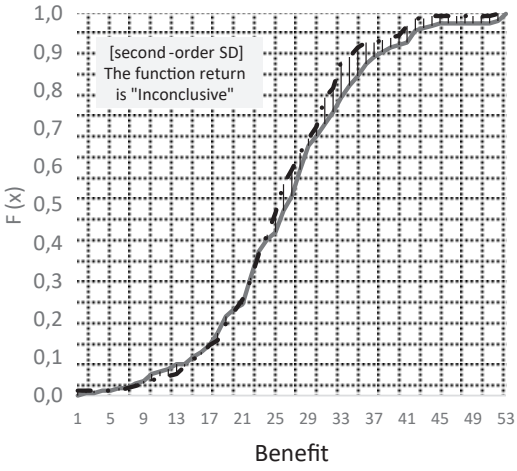

Figure 3. Stochastic dominance results of Islamic and conventional indexes over the period: May 2–October 31, 2023

Source: own work.

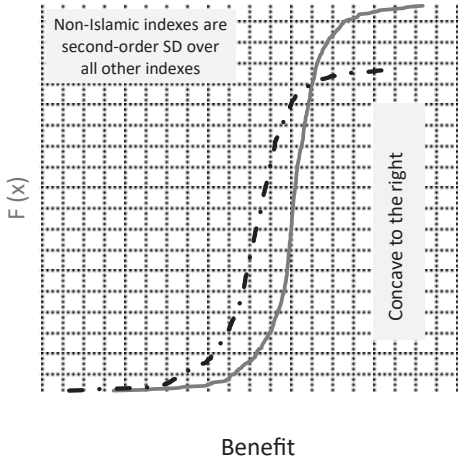

Figure 4. Stochastic dominance results of Islamic and conventional indexes over the period: January 1, 2017–October 31, 2023

Source: own work.

Moreover, during the third sub-period from May 2, 2022 to October 31, 2023, the results are inconclusive. The function returns an "Inconclusive" status because the area below the cross point is smaller than the area above it, indicating no clear pattern of dominance between Islamic and conventional indexes at the lower orders (Figure 3). Throughout the period spanning from 2017 to 2023, clear indications exist of an *S*-shaped second-order stochastic dominance of traditional indexes. This is largely due to the greater size of the area below the intersection point compared to the area above it (Figure 4).

The results mentioned suggest that conventional indexes perform better in terms of returns compared to Islamic indexes. This could be because conventional indexes include companies that may have higher levels of leverage, which can result in potentially higher returns but also higher risk. On the other hand, Islamic screening criteria prohibit use of derivative contracts and investment in interest-based institutions and companies with leverage (beyond certain minimum) as well as businesses which are mostly harmful to society and considered unethical according to Islamic principles. Also, one potential reason for the underperformance of Islamic indexes is the cost associated with implementing ethical screens. These screens require monitoring companies to ensure they comply with Islamic principles, which can lead to additional costs.

During COVID-19, it is impossible to have a safe investment. However, it is important to consider the risk-return trade-off and individual investor's preferences when evaluating the performance of different indexes. Conventional indexes may provide higher returns due to their higher risk (high volatility in stock returns proxied by the standard deviation values, as shown in Table 2), whereas Islamic indexes offer a more conservative and ethical investment methodology. Islamic indexes provided shelter when there was a collapse everywhere due to COVID-19.

Investors often engage in some level of risk-return testing to attain their desired (targeted) returns. They typically assess risks using an *S*-shaped utility function, which means they are more willing to take risks when they are below their desired level of returns (risk-seekers) and become more risk-averse once they surpass that level. This approach helps investors balance their appetite for risk and their desired outcomes (Hamo & Heifetz, 2001). When all outcomes are positive, investors with a more S-shaped profile (indicating more gains) have already secured a higher risk premium, leading to a reduced demand for the risky asset. In contrast, when all outcomes are negative, investors with a more S-shaped profile (indicating more losses) have obtained a lower risk premium, resulting in a higher demand for the risky asset.

Based on this, and since the results show a clear *S*-Shape of second–order dominance, it is suggested that risk-averse investors who exhibit a preference for leftward convexity may choose Islamic indexes over conventional indexes if the Islamic indexes surpass their predetermined desired level.

Moreover, by identifying empirically the preferences of investors, our study points out to the investors that they have diversification opportunities in either direction. Risk-seekers in conventional indexes, who have already attained the desired level (higher gains—higher S), may consider incorporating Islamic indexes to lower the risk. Moreover, conventional indexes that offer higher gains (higher S) come with various risks that should be considered due to their positive probability. Risk-averse investors in Islamic indexes, who are facing losses (many in the concave area), can diversify by adding conventional indexes to attain higher returns. Also, as Islamic indexes are less risky and offer lower return (overall), whereas conventional indexes are dominated with high return and high risk, conventional investors can add Islamic indexes to their conventional portfolios, particularly not to become vulnerable during turbulent times.

Risk-averse investors tend to avoid high-risk indexes in their portfolio, which may result in missing out on potentially higher rates of return. This cautious approach can strategically position them for the future by minimizing potential losses. On the other hand, risk-seeking investors who are "concave to the right" are more comfortable with taking risks, and may therefore prefer conventional indexes. They believe that there is a chance to recover from previous losses and potentially earn higher returns (Elowitz et al., 1980). However, investors tend to be hesitant to realize losses in the stock market and may hold onto losing investments for longer periods, hoping for a rebound (Odean, 1998).

Both Islamic and conventional indexes seem to exhibit a volatility pattern, more prominent in 2020–2022, which shows the impact of COVID-19 pandemic and its spillover to global markets. Figure 5 shows that conventional indexes are ''relatively'' riskier than the Islamic indexes, as the solid line (convention-

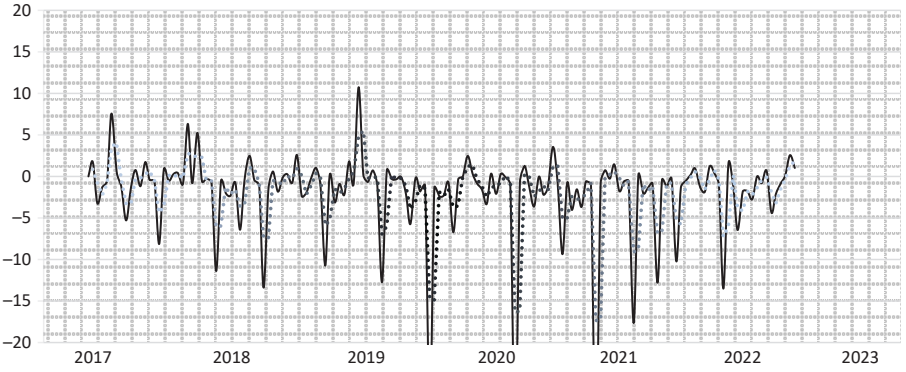

Figure 5. Volatility performance of Islamic and conventional returns over the period from January 1, 2017 till October 31, 2023 (% change)

Note: Solid line represents non-Islamic indexes.

Source: own work.

al indexes) is predominantly above the dotted line (Islamic indexes), particularly during the COVID-19 pandemic. This aligns with the common perception among investors that Islamic investments are relatively less risky due to their adherence to ethical and Sharia'h principles, such as avoiding excessive leverage. These findings are also consistent with previous studies and suggest that risk-averse investors may benefit from switching their investments from conventional to Islamic indexes to develop a low-risk portfolio (Miniaoui et al., 2015; Habib et al., 2014). This also indicates that some investors who took positions in conventional indexes may add Islamic indexes to develop an overall low-risk portfolio, particularly after learning some lessons from turbulent and volatile periods, including global financial crisis, European debt crisis, and recently, COVID-19. However, risk-seekers may continue to invest more in risky conventional indexes.

Conclusions

This study evaluates the performance of Islamic versus conventional stock indexes based on global data. It also explores the investors' preferences of Islamic or conventional indexes based on their attitude towards risk—risk aversion or risk-seeking. Given the growth and increased interest in Islamic investing globally, the performance and risk relationship of both Islamic and conventional stock indexes is a significant topic among investors, portfolio managers, index providers, and policy makers. Utilizing the VARMAX procedure, stochastic dominance analysis, and Johansen's co-integration approach, this study examines the long-term relationship between indexes and the potential benefits of diversification. The research design involves the analysis of 31 conventional and 31 Islamic stock indexes from various series, including FTSE, DJ, MSCI, S&P, and Jakarta. The sample period ranges from January 1, 2017 to October 31, 2023. The overall period of analysis was subdivided into three sub-periods: pre-, during, and post-COVID-19.

Our findings indicate that there is no long-term co-integration among 31 pairs of indexes, except for China and Hong Kong (MSCI Zhong Hua), which presents diversification opportunities. The results suggest that conventional indexes have outperformed Islamic indexes in the pre- and during COVID-19 sub-periods. However, no clear dominance order is observed in the post-pandemic period. Over the entire sample period, most conventional indexes demonstrate second-order stochastic dominance over Islamic indexes, establishing an S-Shape dominance relationship.

The research suggests that investors should consider their risk attitudes and past performance when making investment decisions. Opting for Islamic

indexes over conventional ones can be advantageous for risk-averse investors seeking lower risk. Additionally, diversification opportunities exist for both risk-seeking and risk-averse investors across various index types. Integrating Islamic indexes into conventional portfolios can help mitigate vulnerability during turbulent market conditions. While Islamic indexes offer lower risks and returns, transitioning investments from conventional to Islamic indexes may benefit risk-averse investors in uncertain market environments. However, the costs of implementing ethical screens for Islamic indexes could impact their overall performance, potentially reducing returns. Investors should weigh the benefits of ethical investing against these costs and consider both the ethical and financial implications when investing in Islamic indexes.

References

Abbes, M. B., & Trichilli, Y. (2015). Islamic stock markets and potential diversification benefits. *Borsa Istanbul Review*, *15*(2), 93–105. https://doi.org/10.1016/j.bir.2015.03.001

Abu-Alkheil, A., Khan, W. A., Parikh, B., & Mohanty, S. K. (2017). Dynamic co-integration and portfolio diversification of Islamic and conventional indices: Global evidence. *The Quarterly Review of Economics and Finance*, *66*, 212–224. https://doi.org/10.1016/j.qref.2017.02.005

Al-Awadhi, A. M., Alsaifi, K., Al-Awadhi, A., & Alhammadi, S. (2020). Death and contagious infectious diseases: Impact of the COVID-19 virus on stock market returns. *Journal of Behavioral and Experimental Finance*, *27*(1), 100326. https://doi.org/10.1016/j.jbef.2020.100326

AlKhazali, O., Lean, H. H., & Zoubi, T. (2022). The size anomaly in Islamic stock indexes: A stochastic dominance approach. *International Journal of Financial Studies*, *10*(4), 102. https://doi.org/10.3390/ijfs10040102

Alsharari, N. M., & Alhmoud, T. R. (2019). The determinants of profitability in Shariacompliant corporations: Evidence from Jordan. *Journal of Islamic Accounting and Business Research*, *10*(4), 546–564. https://doi.org/10.1108/jiabr-05-2016-0055

Angosto-Fernández, P. L., & Ferrández-Serrano, V. (2022). World capital markets facing the first wave of COVID-19: Traditional event study versus sensitivity to new cases. *Economics and Business Review*, *8*(4), 5–38. https://doi.org/10.18559/ebr.2022.4.2

Anh, D. L. T., & Gan, C. (2020). The impact of the COVID-19 lockdown on stock market performance: Evidence from Vietnam. *Journal of Economic Studies*, *48*(4). https://doi.org/10.1108/jes-06-2020-0312

Arif, M., Naeem, M. A., Hasan, M., M Alawi, S., & Taghizadeh-Hesary, F. (2021). Pandemic crisis versus global financial crisis: Are Islamic stocks a safe-haven for G7 markets? *Economic Research-Ekonomska Istraživanja*, *35*, 1707–1733. https://doi.org/10.1080/1331677x.2021.1910532

Arouri, M. E., Ben Ameur, H., Jawadi, N., Jawadi, F., & Louhichi, W. (2013). Are Islamic finance innovations enough for investors to escape from a financial downturn? Further evidence from portfolio simulations. *Applied Economics*, *45*(24), 3412–3420. https://doi.org/10.1080/00036846.2012.707776

Ashraf, D. (2013, June 30–July 3). *Genesis of differential performance between Shari'ah compliant equity indexes and conventional indexes*. 20th Annual Conference of the Multinational Finance Society. Izmir, Turkey.

Aslam, F., Mohmand, Y. T., Ferreira, P., Memon, B. A., Khan, M., & Khan, M. (2020). Network analysis of global stock markets at the beginning of the coronavirus disease (COVID-19) outbreak. *Borsa Istanbul Review*, *20*, S49–S61. https://doi.org/10.1016/j.bir.2020.09.003

Cheong, C. W. H. (2021). Risk, resilience, and Shariah-compliance. *Research in International Business and Finance*, *55*, 101313. https://doi.org/10.1016/j.ribaf.2020.101313

Cheong, M. P., Sheble, G. B., Berleant, D., Teoh, C. C., Argaud, J. Ph., Dancre, M., Andrieu, L., & Barjon, F. (2007). Second order stochastic dominance portfolio optimization for an electric energy company. Power Tech Conference, Lausanne, Switzerland.

Chowdhury, M. I. H., Balli, F., & Bruin, A. (2021). Islamic equity markets versus their conventional counterparts in the COVID-19 age: Reaction, resilience, and recovery. *International Review of Finance*, *22*(2), 315–324. https://doi.org/10.1111/irfi.12349

Clark, E., Qiao, Z., & Wong, W. K. (2015). Theories of risk: Testing investor behavior on the Taiwan stock and stock index future markets. *Economic Inquiry*, *54*(2), 907–924. https://doi.org/10.1111/ecin.12288

Davidson, R., & Duclos, J. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, *68*(6), 1435–1464. https://doi.org/10.1111/1468-0262.00167

Denuit, M., Eeckhoudt, L., Tsetlin, I., & Winkler, R. L. (2010). Multivariate concave and convex stochastic dominance. *SSRN Electronic Journal*, *29*. https://doi.org/10.2139/ssrn.1594799

Dharani, M., Hassan, M. K., Rabbani, M. R., & Huq, T. (2021). Does the COVID-19 pandemic affect faith-based investments? Evidence from global sectoral indexes. *Research in International Business and Finance*, *59*, 101537. https://doi.org/10.1016/j.ribaf.2021.101537

Ding, W., Levine, R. E., Lin, C., & Xie, W. (2020). Corporate immunity to the COVID-19 pandemic. *SSRN Electronic Journal*, *141*(2). https://doi.org/10.2139/ssrn.3578585

Elowitz, L., Janis, I. L., & Mann, L. (1978). Decision making—a psychological analysis of conflict, choice, and commitment. *Military Affairs*, *42*(2), 105. https://doi.org/10.2307/1987412

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276. https://doi.org/10.2307/1913236

Erdoğan, S., Gedikli, A., & Çevik, E.I. (2020). The effects of the covid-19 pandemic on conventional and Islamic stock markets in Turkey. *Bilimname*, 42, 89–110. https://doi.org/10.28949/bilimname.799413

Fan, W. (2003). An empirical study of cointegration and causality in the Asia-Pacific stock markets. *Social Science Research Network*. https://doi.org/10.2139/ssrn.360160

Fong, W. M., Wong, W. K., & Lean, H. H. (2005). International momentum strategies: A stochastic dominance approach. *Journal of Financial Markets*, *8*(1), 89–109. https://doi.org/10.1016/j.finmar.2004.08.001

Gomes, F. J. (2005). Portfolio choice and trading volume with loss-averse investors. *The Journal of Business*, *78*(2), 675–706. https://doi.org/10.1086/427643

Goodell, J. W. (2020). COVID-19 and finance: Agendas for future research. *Finance Research Letters*, *35*, 101512. https://doi.org/10.1016/j.frl.2020.101512

Grira, J., Hassan, M. K., & Soumare, I. (2016). Pricing beliefs: Empirical evidence from the implied cost of deposit insurance for Islamic banks. *Economic Modelling*, *55*, 152–168.

Habib, M., & Islam, K. U. (2014). Performance of Shariah compliant index: A comparative study of India and Malaysia. *Indonesian Journal of Islam and Muslim Societies*, *1*(6), 231–241. https://ssrn.com/abstract=2713219

Hamo, Y., & Heifetz, A. (2001). An evolutionary perspective on goal seeking and escalation of commitment. Working Papers 2001-4. https://econpapers.repec.org/RePEc:fth:teavfo:2001-4

Hasan, M. B., Hassan, M. K., Rashid, M. M., & Alhenawi, Y. (2021a). Are safe haven assets really safe during the 2008 global financial crisis and COVID-19 pandemic? *Global Finance Journal*, *50*, 100668. https://doi.org/10.1016/j.gfj.2021.100668

Hasan, M. B., Mahi, M., Hassan, M. K., & Bhuiyan, A. B. (2021b). Impact of COVID-19 pandemic on stock markets: Conventional vs. Islamic indexes using wavelet-based multi-timescales analysis. *The North American Journal of Economics and Finance*, *58*, 101504. https://doi.org/10.1016/j.najef.2021.101504

Hendry, D. F., & Juselius, K. (2001). Explaining cointegration analysis: Part II. *The Energy Journal*, *22*(1). https://doi.org/10.5547/issn0195-6574-ej-vol22-no1-4

Heyden, K. J., & Heyden, T. (2020). Market reactions to the arrival and containment of COVID-19: An event study. *Finance Research Letters*, *38*, 101745. https://doi.org/10.1016/j.frl.2020.101745

Hodges, C. W., & Yoder, J. A. (1996). Time diversification and security preferences: A stochastic dominance analysis. *Review of Quantitative Finance and Accounting*, *7*(3), 289–298. https://doi.org/10.1007/bf00245255

Ibrahim, M. H. (2015). Issues in Islamic banking and finance: Islamic banks, Shari'ah-compliant investment and sukuk. *Pacific-Basin Finance Journal*, *34*, 185–191. https://doi.org/10.1016/j.pacfin.2015.06.002

Ilhan, B., & Masih, M. (2014). *Do portfolio diversification opportunities exist across the Euro zone Islamic equity markets? MGARCH-DCC and Wavelet Correlation analysis*. MPRA Working Paper, 57688. http://mpra.ub.uni-muenchen.de/57688/

Khamlichi, A. E., Sarkar, K., Arouri, M., & Teulon, F. (2014). Are Islamic equity indices more efficient than their conventional counterparts? Evidence from major global index families. *Journal of Applied Business Research*, *30*(4), 1137–1150. https://doi.org/10.19030/jabr.v30i4.8660

Kok, S., Giorgioni, G., & Laws, J. (2009). Performance of Shariah-compliant indices in London and NY stock markets and their potential for diversification.

International Journal of Monetary Economics and Finance, *2*(3/4), 398–408. https://doi.org/10.1504/ijmef.2009.029071

Lean, H. H., Smyth, R., & Wong, W. (2007). Revisiting calendar anomalies in Asian stock markets using a stochastic dominance approach. *Journal of Multinational Financial Management*, *17*(2), 125–141. https://doi.org/10.1016/j.mulfin.2006.05.003

Levy, H., & Levy, M. (2003). Prospect theory and mean-variance analysis. *Review of Financial Studies*, *17*(4), 1015–1041. https://doi.org/10.1093/rfs/hhg062

Levy, H., & Wiener, Z. (1998). Stochastic dominance and prospect dominance with subjective weighting functions. *Journal of Risk and Uncertainty*, *16*, 147–163. https://doi.org/10.1023/A:1007730226688

Masih, M., Al-Sahlawi, M. A., & Lurion De Mello. (2010). What drives carbon-dioxide emissions: Income or electricity generation? Evidence from Saudi Arabia. *The Journal of Energy and Development*, *33*(2), 201–213.

Masih, R., & Masih, A. M. (2001). Long and short term dynamic causal transmission amongst international stock markets. *Journal of International Money and Finance*, *20*(4), 563–587. https://doi.org/10.1016/s0261-5606(01)00012-2

Miniaoui, H., Sayani, H., & Chaibi, A. (2015). The impact of financial crisis on Islamic and conventional indexes of the GCC countries. *Journal of Applied Business Research*, *31*(2), 357–370. https://doi.org/10.19030/jabr.v31i2.9171

Mishra, N. P., Das, S. S., Yadav, S., Khan, W., Afzal, M., Alarifi, A., Kenawy, E. R., Ansari, M. T., Hasnain, M. S., & Nayak, A. K. (2020). Global impacts of pre- and post-COVID-19 pandemic: Focus on socio-economic consequences. *Sensors International*, *1*(100042), 100042. https://doi.org/10.1016/j.sintl.2020.100042

Mumtaz, H., Shahmoradi, A., & Turk, R. (2016). An overview of Islamic finance. *Journal of International Commerce, Economics and Policy*, *7*(1), 1–28. https://doi.org/10.1142/S1793993316500034

Nomran, N. M., & Haron, R. (2021). The impact of COVID-19 pandemic on Islamic versus conventional stock markets: international evidence from financial markets. *Future Business Journal*, *7*(1), 33. https://doi.org/10.1186/s43093-021-00078-5

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of Finance*, *53*(5), 1775–1798. https://doi.org/10.1111/0022-1082.00072

Post, T., van Vliet, P., & Levy, H. (2008). Risk aversion and skewness preference. *Journal of Banking & Finance*, *32*(7), 1178–1187. https://doi.org/10.1016/j.jbankfin.2006.02.008

Saiti, B. (2014). Integration between the conventional and Islamic stock indexes with the US stock market: Evidence from cointegration analysis. *Australian Journal of Basic and Applied Sciences*, *8*(10), 18–30.

Saiti, B., Bacha, O. I., & Masih, M. (2014). The diversification benefits from Islamic investment during the financial turmoil: The case for the US-based equity investors. *Borsa Istanbul Review*, *14*(4), 196–211. https://doi.org/10.1016/j.bir.2014.08.002

Salisu, A. A., & Sikiru, A. A. (2020). Pandemics and the Asia-Pacific Islamic stocks. *Asian Economics Letters, 1*(1). https://doi.org/10.46557/001c.17413

Shaikh, S. (2010). *A brief review & introduction to practiced Islamic banking & finance*. MPRA Paper 19458. https://ideas.repec.org/p/pra/mprapa/19458.html

Sheng, H., & Tu, A. H. (2000). A study of cointegration and variance decomposition among national equity indices before and during the period of the Asian financial

crisis. *Journal of Multinational Financial Management*, *10*(3–4), 345–365. https://doi.org/10.1016/s1042-444x(00)00034-7

Sherif, M. (2020). The impact of Coronavirus (COVID-19) outbreak on faith-based investments: An original analysis. *Journal of Behavioral and Experimental Finance*, *28*, 100403. https://doi.org/10.1016/j.jbef.2020.100403

Standard and Poor's. (2022). *Islamic finance outlook*. https://www.spglobal.com/ratings/en/research/pdf-articles/islamic-finance-outlook-2022-28102022v1.pdf

Standard Chartered. (2023). *Islamic finance and the securities market outlook for 2023*. https://www.sc.com/en/news/financial-institutions-investor-groups/islamic-finance-securities-market-outlook-2023/

Tahir, M., & Ibrahim, S. (2020). The performance of Shariah-compliant companies during and after the recession period—evidence from companies listed on the FTSE All World Index. *Journal of Islamic Accounting and Business Research*, *11*(3), 573–587. https://doi.org/10.1108/jiabr-07-2016-0082

Tanin, T. I., Ahmad, A. U. F., & Muneeza, A. (2021). Shariah-compliant equities and Shariah screening: need for convergence of ethical screening of stocks with Shariah screening. *International Journal of Emerging Markets*, *18*(2), 296–315. https://doi.org/10.1108/ijoem-09-2020-1041

Topcu, M., & Gulal, O. S. (2020). The impact of COVID-19 on emerging stock markets. *Finance Research Letters*, *36*, 101691. https://doi.org/10.1016/j.frl.2020.101691

Whitmore, G., & Findlay, M. (1978). *Stochastic dominance: An approach to decision-making under risk health*. Lexington Books.

WHO (World Health Organization). (2023). Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic

Wong, W. K., & Chan, R. H. (2007). Prospect and Markowitz stochastic dominance. *Annals of Finance*, *4*(1), 105–129. https://doi.org/10.1007/s10436-007-0072-4

Worthington, A., & Higgs, H. (2003). Transmission of equity returns and volatility in Asian developed and emerging markets: A multivariate GARCH analysis. *International Journal of Finance & Economics*, *9*(1), 71–80. https://doi.org/10.1002/ijfe.222

Yarovaya, L., Elsayed, A. H., & Hammoudeh, S. (2021). Determinants of spillovers between Islamic and conventional financial markets: Exploring the safe haven assets during the COVID-19 pandemic. *Finance Research Letters*, *43*, 101979. https://doi.org/10.1016/j.frl.2021.101979

Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. *Finance Research Letters*, *36*, 101528. https://doi.org/10.1016/j.frl.2020.101528

sciendo Economics and Business Review https://www.ebr.edu.pl Vol. 10 (2), 2024: 60–91 https://doi.org/10.18559/ebr.2024.2.1140

Taxation of public pensions in European Union countries

Maciej Cieślukowski[1]

Abstract

The aging of society is one of the most important trends shaping the social, economic and political life of the 21st century. However, with the increasing number of people of retirement age, the problem of ensuring adequate conditions for a longer life arises. The state influences these conditions through the pension security system, including taxation of pensions. The paper attempts to answer the question whether taxation of remunerations and public pension benefits may have a significant impact on making decisions about choosing a country of work in the common market. For this purpose, Member States have been ranked in terms of two dimensions—the conditions of taxation of wages and the conditions of taxation of retirement benefits. The countries were classified using a multi-criteria comparative analysis and the agglomeration method. The study shows that taxing salaries and pension benefits is of marginal importance from the point of view of an employee's decision-making. The main factors are the average expenditure on net salaries and the average expenditure on net pension benefits.

Keywords

- public pensions
- taxes
- social security contributions
- EU countries
- labour costs

JEL codes: H2, H55, J26, J32

Article received 16 January 2024, accepted 10 May 2024.

Suggested citation: Cieślukowski, M. (2024). Taxation of public pensions in European Union countries. *Economics and Business Review*, *10*(2), 60–91. https://doi.org/10.18559/ebr.2024.2.1140

[1] Department of Public Finance, Poznań University of Economics and Business, al. Niepodległości 10, Poznań, Poland, Maciej.cieslukowski@ue.poznan.pl, https://orcid.org/0000-0001-6058-0113.

Introduction

The aging of society is one of the most important trends shaping the social, economic and political life of the 21st century. Causes include the decreasing birth rate, on the one hand, and technological progress in medicine and the healthier and more aware lifestyle among retired people on the other. However, with the increasing number of people of retirement age, the problem of ensuring adequate conditions for a longer life arises. In 2020, in EU countries, a total of 19.2% of the population over 60 years of age was at risk of poverty and social exclusion (Eurostat). However, the variation in this indicator across individual countries is enormous: in Luxembourg it was 8.3%, while in Bulgaria it was 43.5%.

The state can influence these conditions through the pension security system, including taxation of pensions. This issue is a subject of interest in the literature. The taxation of public and private pensions is examined in terms of factors such as its fiscal and redistributive effects, and its impact on the capital market and economic growth. The taxation of occupational pensions is also of interest to the European Commission in the context of the functioning of the common European market.

People of working age can choose any country in the common European market to improve their living conditions both at working age and after retirement. The paper attempts to answer the question whether taxation of remunerations and public pension benefits may have a significant impact on making such decisions. In this way, people of retirement age could reduce the potential negative effects of poverty. To the author's knowledge, such a topic has not been the subject of particular research interest so far.[2]

For this purpose, Member States have been classified in terms of two dimensions—the conditions of taxation of wages and the conditions of taxation of retirement benefits. The empirical data are of a macroeconomic nature. EU countries were ranked using multidimensional comparative analysis (MCA). In order to analyse the ranking results obtained, cluster analysis was used using the agglomeration algorithm and the Euclidean distance measure. Calculations were performed using Multivariate Statistics from the Statistica 13 package. As a result of applying the agglomeration method in the first dimension of the analysis (conditions of taxation of wages), six clusters of Member States were obtained, while in the second dimension of the analysis (conditions of taxation of pension benefits), two clusters were obtained.

[2] The impact that taxing pensions has on mobility is relatively minor and is not particularly evident in the literature. However, a related issue focusing on ongoing or planned or even expected pension reforms (caused by fiscal difficulties) is an issue analysed in the literature.

The paper is organised as follows: Section 1 provides a review of the relevant academic literature, Sections 2 and 3 provide information on the structure and taxation principles of public pensions in EU countries. Section 4 presents the demographic structures in EU countries. Section 5 presents the research methods and sources of data. The last section presents empirical findings. The paper ends with conclusions.

1. Literature review

Research on the taxation of pensions is generally conducted in terms of the taxation of private pensions (individual and occupational) and the taxation of public pensions, both from a theoretical and empirical perspective. Research on the taxation of private pensions is older and much more extensive, and is conducted in terms of the optimal taxation of capital income at the saving stage (payment of contributions), investment operations (pension funds), and payment of retirement benefits (Cremer & Pestieau, 2016). In this regard, it is possible to apply six basic taxation models: *EEE*, *EET*, *ETT*, *TTT*, *TTE*, *TEE*, where E denotes exempt and T denotes taxed. OECD reports (OECD, 2011, 2023) show that the *EET* system is in force in most European Union and OECD countries. Research generally emphasizes the advantage of the *EET* system over the competing *TEE* system in the context of the impact on savings, development of the capital market and improvement of corporate governance, the impact on economic growth and better risk-sharing between generations and between retirees and the state. Very good reviews of the literature and research achievements in this field are provided, e.g., by Cremer and Pestieau (2016) and Armstrong et al. (2015). Moreover, the *EET* system is also preferred by the European Commission in the context of the functioning of the common market and counteracting pension tax avoidance (European Commission, 2001). An overview of the effects of taxation of occupational pensions in the common market is provided by Patterson (2002) and Kluzek (2022), among other authors.

Ferrarini and Nelson (2003) point out that research on the taxation of social benefits is relatively new and dates back only to the early 1970s. Taxation of public pensions is considered at the stage of professional work and the stage of payment of pension benefits. Here we can consider four basic tax models: *EE*, *ET*, *TT* and *TE*. Initially, research was conducted separately on these two aspects (Cremer & Pestieau, 2016). In a new approach to research in this area, authors combine the design of the optimal pension level with taxation in a single model (Cremer et al., 2008; Diamond, 2009). The authors empha-

sise that in order to function over the participants' lifetimes (tens of years), the pension system must be in balance, that is, the demand for pension rights (the sum of contributions paid) and their supply (the sum of pension benefits received) must be equal. Therefore, the amount of pensions and the contribution rate are closely related and cannot be shaped freely.

Theoretical research shows that imbalances may result from situations when people of retirement age extend their working life but additional years of work do not increase their pension, or when working people take early retirement. Then the level of pensions should be adjusted through appropriate taxation. Lozachmeur (2006) and Cremer et al. (2008) indicate that taxation of income during work should be higher than the taxation of retirement benefits.

The redistributive and fiscal effects of taxes on pensions are also a relatively new subject of research. Ferrarini and Nelson (2003) point out that the redistributive effects of social benefits should be examined after they are taxed. Ignoring income taxes may lead to an overestimation of the positive impact of benefits on reducing income inequality and may lead to incorrect conclusions regarding the importance of various tax instruments in the redistribution of income.

Keenay and Whitehouse (2003) conducted a comparative analysis of average and marginal pension tax rates in 15 OECD countries in relation to the applicable working age. Their study shows that in all countries people of retirement age experience a much lower tax burden than people of working age. Similar conclusions also result from regular OECD studies (OECD, 2021, 2023).

Verbist (2007) examines the impact of taxation on pensions and unemployment benefits on redistributive effects in 15 EU countries using EUROMOD. The author finds significant differences in taxation, but points out that in all countries taxation of pensions is lower than taxation of income from work, and pensions after taxation have greater purchasing power than before taxation. Additionally, research shows that in most countries, taxation of pensions is more progressive than taxation of income from work, which translates into the considerable importance of the tax system in reducing income inequality among older people.

The use of various types of preferences in the taxation of pension contributions and benefits may lead to the creation of income savings that affect redistributive effects. In the literature on the subject, such savings are called tax expenditures (Swift, 2006). Barrios et al. (2020) estimate the fiscal and redistributive effects of tax expenditures in EU countries, across both private and public pensions, using the EUROMOD model. In particular, the authors compare the current pension taxation systems with the *ET* model system. All derogations are treated as tax expenditures. Research shows that these deviations are significant and progressive.

Jun et al. (2023) show that OECD countries, by trying to improve the redistributive effects of public pensions, introduce elements of private security into the system. As a result, tax expenditure for private pensions is increasing.

Another detailed analysis of fiscal and redistributive effects and the impact on poverty using the EUROMOD model is made by Ivaškaitė-Tamošiūnė and Thiemann (2021). The authors examine the indicated effects for 27 EU countries in scenarios of transition from the current taxation systems (*EE*, *ET*, *TE*, *TT*) to the *EE* and *TT* systems.

Genser and Holtzmann (2021) consider the effects of taxation on foreign public and private pensions in the context of the applicable rules in double taxation agreements signed by Germany. The taxation of cross-border pensions is assessed as very complex, unfair and inconsistent with the international mobility of citizens. The authors propose a new instrument in bilateral tax treaties named 'pretaxation of pensions' as a suitable economic concept for global pension taxation.

2. Pension systems in EU countries

Pension systems in today's developed countries serve three main goals: 1) helping individuals redistribute resources from work in old age, 2) protecting people emerging poverty in old age, 3) providing insurance and reducing disparities in monthly old-age income, regardless of longevity (Shi & Kolk, 2022). In traditional typologies of pension systems in OECD countries, systems are referred to as "Bismarck" are oriented towards income replacement (meeting the first and third goals), while "Beveridge" systems focus on protecting against poverty (second goal) with less emphasis on linking pensions to previous earnings (Ebbinghaus, 2021). These goals are achieved through the mechanism of transferring funds from higher income earners to lower earners as an integral part of mandatory government welfare systems.

In 2020, obligatory pension security systems in EU countries were a combination of two tiers (pillars) (Table 1). The first pillar is basic pension security. Pensions are awarded on the basis of the principle of residence (provisioning technique) or on the basis of contributions paid (insurance technique) but are not linked to earnings. In the first case, the pension is granted to residents of a given country and usually reflects basic living costs (basic pension). Another solution may be so-called targeted pensions that depend on additional residence criteria, e.g., the amount of assets owned. Then, less wealthy people receive higher pensions. The contributory system includes basic and minimum pensions. The former usually depend on the contributions paid and the

Table 1. Obligatory pension systems in EU countries in 2020

Country	First tier				Second tier	
	residence-based		contribution-based			
	basic	tar-geted	basic	mini-mum	public	private
Austria				X	DB	
Belgium				X	DB	
Bulgaria				X	DB	
Romania				X	DB	
Czech Republic			X	X	DB	
Denmark	X	X			FDC	FDC
Cyprus				X	DB	
Estonia			X		DB*/Points	FDC*
Malta				X	DB	
Finland		X			DB	
France				X	DB + Points	
Germany					Points	
Greece	X				DB + NDC	
Hungary				X	DB	
Ireland			X			
Italy					DB*+NDC	
Latvia				X	DB*/NDC + FDC	
Lithuania			X		Points	
Luxembourg			X	X	DB	
Netherlands	X					FDC
Poland				X	FDC	
Portugal				X	DB	
Slovakia				X	Points	
Slovenia				X	DB	
Spain				X	DB	
Sweden		X			NDC + FDC	FDC

* Changes applicable to people who retired in 2020.

Source: based on (Euromod, 2011–2022; OECD, 2021, 2023).

contribution period. The state may also provide, upon the meeting of certain contribution conditions, a minimum amount of pension benefit.

The second pillar is based solely on a contribution system closely related to earnings. There are three detailed pension calculation schemes: defined benefit (DB), points-based, and defined contribution (DC). The DB scheme is an expression of the pay-as-you-go principle, so pensions are calculated based on the contribution period, income earned and contributions applicable at a given time. The points system means that employees earn pension points based on their earnings. After retirement, the amount of the benefit is the product of the sum and the value of pension points.

Defined Contribution (DC) systems are based on individual pension accounts and can take two forms: Financial Defined Contribution (FDC) or Notional Defined Contribution (NDC). In FDC, contributions are transferred to an individual pension account and then invested in accordance with the adopted regulations. The accumulation of contributions and investment earnings is usually converted into a monthly pension at the time of retirement. The second type (NDC) consists of programs with individual accounts in which contributions are capitalized and for which a hypothetical rate of return is applied. In consequence, the accounts exist only in the books of the managing authority. At retirement, the accumulated nominal capital is converted into a monthly pension using a formula based on life expectancy or mortality rates.

Denmark, the Netherlands and Sweden also have mandatory private insurance pillars. They are an important complement to pension systems based mainly on a provisioning scheme. Such an additional program was also introduced in Estonia in 2020. In Ireland, there exists only the first pillar of pension security.

It should be emphasised that pension systems across EU countries also differ in terms of such factors as the principles and amounts of payment of pension contributions, indexation of pensions, sources of financing and payment of pensions, and retirement age. Some countries also have different pension systems for public and private sector employees (Euromod, 2011–2022; OECD, 2021, 2023).

For example, in Denmark and Malta, pension funds are not separated as subsectors of public finances. Contributions apply in Malta, but they go into the state budget. By contrast, in Denmark, public pensions are generally financed by taxes, and in the case of local civil servants, pensions are also paid from local budgets. In the last few decades, many countries also shifted their public pension systems towards private ones (Manor & Ratajczak, 2020).

3. Principles of taxation of public pensions in EU countries

Public pensions may be taxed at two stages of life: when working and receiving remuneration, and after retirement, i.e. when receiving a pension benefit. Taxation of a pension during the working period is considered in the context of the occurrence and tax treatment of compulsory pension insurance contributions. If it is possible to deduct all contributions from income tax (the tax base), pensions are exempt from taxation (*E* – exempt), otherwise they are fully taxed (*T* – taxed). Countries may also apply partial deductions for contributions (*t* – taxed partially). Pension contributions may be levied to varying degrees on both employees and employers. In turn, the pension benefit is one of the sources of income, and as such it may be subject to income tax (*T*), it may be subject to taxation on preferential terms (*t*), or it may be completely exempt from tax (*E*).[3]

Pension taxation systems in EU countries are diverse, and it is difficult to provide clear reasons for their construction. Taking into account the fiscal burden imposed on the employee and the pension benefit, six models of taxation of public pensions can be distinguished (Table 2).

Table 2. Classification of EU countries in terms of the model of taxation of public pensions in 2019–2022

Model	Country
EE	Bulgaria, Slovakia
ET	Croatia, Cyprus, Denmark, Estonia, Greece, Malta, Poland
Et	Austria, Belgium, Finland, Latvia, Portugal, Romania, Slovenia, Spain, Sweden, Italy
TE	Lithuania, Hungary
Tt	Czech Republic
tt	France, Germany, Ireland, Luxembourg, Netherlands

Source: based on (Euromod, 2011–2022; European Commission, n.d.; Ivaškaitė-Tamošiūnė & Thiemann 2021, pp. 10–11).

Table 2 shows that in most countries, public pensions are exempt from taxation at least at one stage of life. The most popular among EU countries is the *Et* model. In this model, an employed person has the option of fully deducting pension insurance contributions from the income tax base, and the pen-

[3] In theory, taxation of pension liabilities during the working period is also possible in systems in which the liabilities are expressed in money terms (in practice not applied).

sion benefit is taxed on preferential terms. As a rule, pensioners are granted various tax reliefs depending on the amount of their pension and their family status. In turn, the friendliest tax regulations are found in Bulgaria and Slovakia. In these countries, the entire pension insurance contribution can be deducted from the tax base, and the pension benefit is completely tax-free.

The tax burden of the pension is also influenced by the division of the pension contribution between the employee and the employer. In this respect, six models can be distinguished in EU countries (Table 3).

Table 3. Classification of EU countries in terms of the division of mandatory pension insurance contributions between the employee (*Ee*) and the employer (*Er*) in 2019–2022

Model	Country
Ee = Er	Cyprus, Germany, Luxembourg, Malta, Poland
Ee < Er	Austria, Belgium, Bulgaria, Czech Republic, Finland, France, Greece, Hungary, Ireland, Italy, Latvia, Slovakia, Spain
Ee > Er	Slovenia
Ee	Croatia, Lithuania, Netherlands, Romania
Er	Estonia, Sweden
No contributions	Denmark

Source: based on (Euromod, 2019–2022; Ivaškaitė-Tamošiūnė & Thiemann 2021, pp. 10–11).

In most EU countries, the employer pays the higher pension contribution. Particularly important differences in this respect are found in Italy, the Czech Republic, France, and Spain, among other countries. Only in Slovenia does the employee pay the greater part of the pension contribution. In four countries, only employees pay the contribution, but in Lithuania and Romania such rules have only been in force for a few years. Until the end of 2019 and 2018, respectively, the majority of the pension contribution was paid by employers. In Estonia and Sweden, pension contributions are paid exclusively by the employer, while in Denmark public pensions are financed by taxes.

In the context of the further analysis, it is significant that in Ireland and Malta compulsory social security contributions are of a general nature and also cover, in addition to pension insurance, other types of insurance (e.g., health, unemployment).

4. Demographic structure of EU countries

Pension systems, including the conditions for taxation of pension benefits, depend largely on the demographic structure of the population. Between 2010 and 2020, there was an average increase in the median age in EU countries from 40.1 to 42.7 years. This means that the EU population is aging. The median age has increased in all EU countries, except Sweden, where the rate decreased slightly from 40.7 to 40.5 years (Figure 1). The highest average median age in the study period was recorded in Germany (45.5), Italy (45.2) and Portugal (43.5), and the lowest in Ireland (36.1), Cyprus (36.8) and Slovakia (39.0).

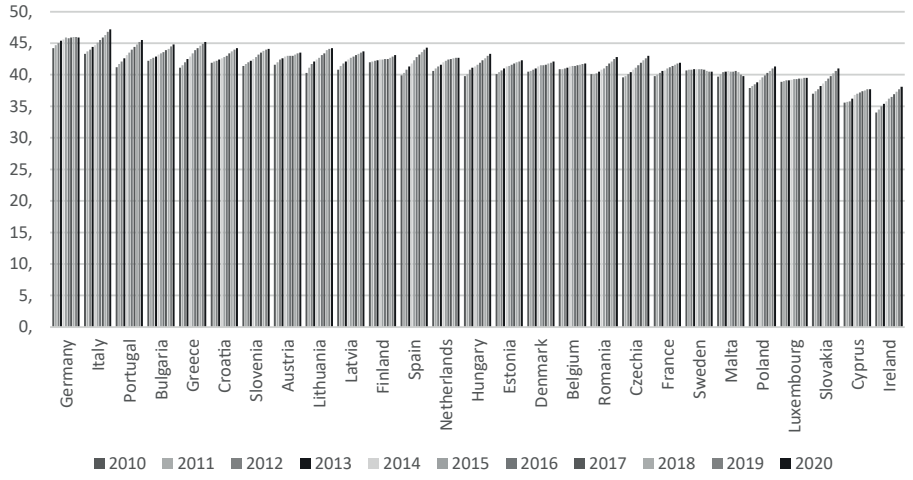

Figure 1. Median age in EU countries in 2010–2020

Source: based on Eurostat data.

As a result of aging population, the number of people of retirement age and the elderly dependency ratio are increasing. In the years 2010–2020, the total number of retirees in EU countries increased from 82.3 to almost 91.7 million people. The highest increase in the number of retirees was recorded in Luxembourg (58.8%), Ireland (45.0%) and Slovakia (44.3%). In turn, in five countries the number of retirees decreased. The largest decline was recorded in Greece (7.4%), Latvia (6.0%) and Estonia (3.2%). The number of retirees across EU countries varies greatly (Figure 2). Over 65.7% of the total number of retirees live in Germany, France, Italy, Poland and Spain.

In the years 2010–2020, the elderly dependency ratio for EU countries increased from 26.3 to 32.0%. This means that for every person aged 65 or over, the number of working age people decreased. In 2010, it was less than

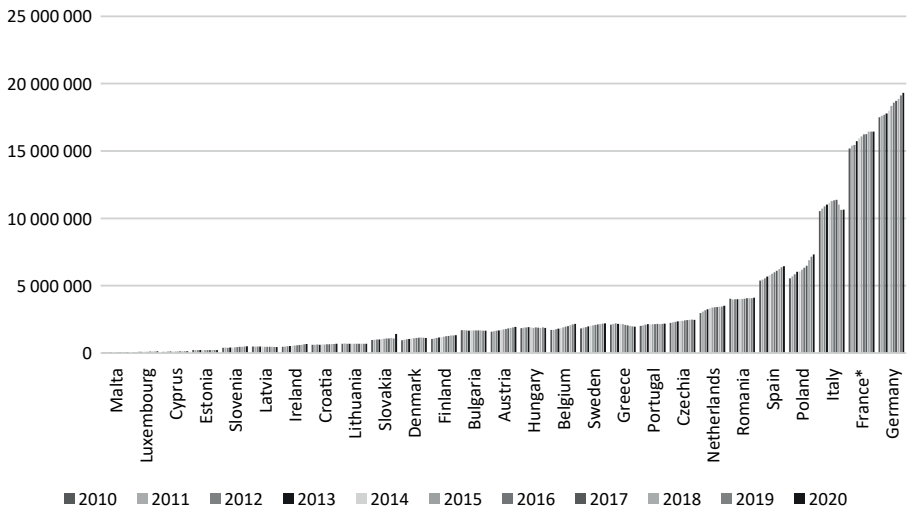

Figure 2. Number of pensioners in EU countries in 2010–2020

* No data for 2019 and 2020. Data from 2018 were used for the analysis.

Source: based on Eurostat data.

4 people, and at the end of 2020, just over 3 people. The lowest average old age dependency ratio was recorded in Ireland (19.5%), Slovakia (20.2%) and Luxembourg (20.5%), and the highest in Italy (33.7%), Germany (32.1%) and Greece (32.1%). As a result of population aging, the number of people of retirement age and the elderly dependency ratio are decreasing.

The increasing old-age dependency ratio means that fewer people work to pay one pension benefit. As a result, pension systems become financially inefficient without additional state support or changes to the system. One of the most popular reforms concerns the increase in the retirement age to order to reduce pension expenses.

In most countries, the retirement age for men and women is equal. Higher ages for men apply in Austria, Bulgaria, Croatia, Hungary, Lithuania, Poland and Romania. The Netherlands and Greece have the highest retirement age, at 66.3 and 67 years. The lowest age for men is 62 and applies in Italy, Luxembourg and Slovenia. In turn, the lowest age for women is 60 and applies in Austria and Poland. The average retirement age across the EU in the year in question is 64.3 years for men and 63.5 years for women. Between 2018 and 2022, the retirement age was slightly increased in France, Latvia, Lithuania, Malta, the Netherlands, Portugal, Slovenia and Slovakia. In 2020, the retirement age for men and women was equalized in Slovenia and Italy, with the latter country significantly lowering it from 67 and 66.3 to 62 years. It should also be emphasised that since 2023, most countries have significantly increased the retirement age, even to 67 years (Bulgaria, Denmark, Greece,

Italy). Only in France was the age lowered to 62.3 years. In Austria, the retirement age for women will gradually rise to that of men (i.e. 65) between 2024 and 2033 (Fratica-Dragomir, 2023).

5. Methods and materials

The classification of Member States was carried out in terms of two dimensions, i.e. divided into the conditions of taxation of remuneration for work and the conditions of taxation of pension benefits.

The conditions of taxation of remuneration from work were described using four macroeconomic variables (X_1–X_4), while the conditions for the taxation of pension benefits were described using 11 macroeconomic variables (X_5–X_{15}):

X_1 – Average annual expenditure on gross wages per person (EUR) – average data 2010–2020,
X_2 – Average expenditure on net wages per person (EUR) – average data 2010–2020,
X_3 – Implicit Tax Rate (*PIT+SSC* employee) (%) – data 2020,
X_4 – Implicit Tax Rate (*SCC* employer) (%) – data 2020,
X_5 – Average annual gross pension expenditure per person (EUR) – average data 2010–2020,
X_6 – Average annual net pension expenditure per person (EUR) – average data 2010–2020,
X_7 – Pension Effective Tax Rate (*PETR*, %) – $((X_1–X_2)/X_1)$,
X_8 – Average annual net pension expenditure per person (*PPS*) – average data 2010–2020,
X_9 – Gross Pension Replacement Rate (*GPRR*, %) – data 2022,
X_{10} – Net Pension Replacement Rate (*NPRR*, %) – data 2022,
X_{11} – Gross Pension Wealth (*GPW* men, %) – data 2022,
X_{12} – Net Pension Wealth (*NPW* men, %) – data 2022,
X_{13} – Gross Pension Wealth Pension (*GPW* women, %) – data 2022,
X_{14} – Net Pension Wealth (*NPW* women, %) – data 2022,
X_{15} – Relationship between the amount of an employee's fiscal payments and average annual gross pension expenditure per person (EUR) $((X_1–X_2)/X_5)$.

Variables X_3, X_7 *and* X_{15} are destimulants, while the remaining variables are stimulants. Some variables require additional explanation. To assess the fiscal burden of a pension at the stage of professional work, the *ITR* (*Implicit Tax Rate*) indicator was selected, covering total labour costs, i.e. both total mandatory social security contributions and personal income tax (*PIT*) paid by the employee. An objective comparison of the actual tax burden of pen-

sions at the stage of work in EU countries, resulting solely from the payment of mandatory pension contributions, is difficult. This is due to different rules on contribution to pension systems. Firstly, in several countries (Denmark, Finland, Greece, the Netherlands and Sweden), residence-based pensions constitute an important part of the system. In Denmark, pension contributions are marginal and pensions are financed mainly by taxes. Therefore, the argument that in Denmark the tax burden of a working pension is the lowest due to the lack of compulsory pension contributions would be incorrect, since the income tax on wages is one of the highest in the world. Secondly, in some countries a general social security contribution is levied without a detailed separation of the pension contribution (Ireland, Malta). As a result, the tax costs of pensions at the stage of work are considered in the article through the prism of the total fiscal burden of work. This approach is also supported by the fact that the pension contribution is one of the obligatory elements of the social security contribution system and the employee cannot choose the type of contribution payment.

The *ITR* indicator, (*PIT + SSCe + SSCr + Tr*)*/*(*We + Tr*), is the quotient of the sum of taxes and social security contributions paid by employees (*PIT, SSCe*) and employers (*Tr, SSCr*) imposed on remuneration and the sum of remuneration paid (*We*) and taxes (e.g., payroll tax) paid by employers on wages (*Tr*). This indicator is a macroeconomic approach to the tax costs of labour in the economy, showing the average tax burden of labour for all employees (European Commission, 2022, p. 42). In the analysis, this indicator was divided into the *ITR* corresponding to the employee's burden (variable X_3) and the *ITR* corresponding to the employer's burden (variable X_4). For the employee, the situation is more favorable the lower the X_3 variable and the higher the X_4 variable.

PETR shows the average tax burden imposed on pension expenditure and is calculated using the formula *PETR = GPE–NPE/GPE*. This indicator is the ratio of the difference between gross pension expenditure (*GPE*) and net pension expenditure (*NPE*) to gross pension expenditure (*GPE*).

The *GPRR* (*Gross Pension Replacement Rate*) and *NPRR* (*Net Pension Replacement Rate*) indicators show what part of the average gross and net remuneration obtained over the entire working period will constitute the gross and net pension, respectively. However, these indicators do not provide a comprehensive measure of cumulative pension payments, i.e. taking into account life expectancy, normal retirement age and indexation of pension benefits. Such measures are the *Gross Pension Wealth* (*GPW*) and *NPW* (*Net Pension Wealth*) indicators, which indicate what part of the average gross and net earnings will constitute the accumulated pension flows. These indicators are presented separately for women and men.

The last indicator (variable X_{15}) is the share of the current average amount of taxes and contributions paid by the employee on remuneration (*PIT + SSCe*) in relation to the average expenditure on pensions per person (*PE*). This indica-

tor shows the degree to which current pension expenditure is covered by the employee's mandatory labour costs. The situation is more favorable for the employee when the value of this indicator is lower. This means that the labour costs the employee must incur are lower than the amount of the benefit paid.

Diagnostic variables derived from or were developed on the basis of the following databases and studies:

- Eurostat and Euromod, country reports (variables X_1, X_2, X_5, X_6, X_7, X_8, X_{15}),
- OECD, 2023 (variables X_9–X_{14}),
- European Commission Data (n.d.) (variables X_3, X_4).

All variables are presented in Appendix A1.

Within two dimensions of the analysis, countries were ranked using multidimensional comparative analysis (MCA). The zero unitarization method was used in the MCA analysis as one of the best methods for this type of application.

Multidimensional comparative analysis is often used to classify EU countries in terms of selected variables. For example, Jankowiak (2021) compared the health care systems of EU countries using the so-called synthetic indicator. Vambol et al. (2023) used a multi-criteria analysis of municipal solid waste management in Poland compared to other EU countries. Iwacewicz-Orłowska and Sokołowska (2018) prepared a ranking of EU countries in terms of the value of environmental governance indicators in selected years.

In order to express the ranking results obtained, it was decided to additionally group the Member States from the point of view of the similarity of the variables studied. For this purpose, cluster analysis was ultimately employed, using the agglomeration algorithm and the Euclidean distance measure. The results of this type of grouping are hierarchical trees. Calculations were performed using Multivariate Statistics of the Statistica 13 package. As a result of applying the agglomeration method in the first dimension of the analysis (conditions of taxation of wages), 6 clusters of Member States were obtained, while in the second dimension of the analysis (conditions of taxation of pension benefits), 2 clusters were obtained.

An alternative cluster search algorithm was V-fold cross-validation of variables. The use of this method consisted in repeating the procedure of drawing a sample for analysis three times from the data and building a model based on it. Initial cluster centers were determined by maximizing Euclidean cluster distances. After finding the optimal number of clusters, the *k*-means clustering algorithm was used. The calculations were performed using the Data Mining module of the Statistica 13 package. Ultimately, two clusters were obtained in both dimensions of the analysis (conditions of taxation of wages and conditions of taxation of pension benefits).

The clusters in the second dimension of the analysis turned out to be identical when using two alternative clustering algorithms (agglomeration and

cross-evaluation). The paper presents detailed results of the agglomeration method, due to the fact that it resulted in more clusters in the first dimension of the analysis.

6. Results of the analysis

6.1. Wages and the fiscal burden of work

The average expenditure on gross wages per employee in the EU countries (EU-27) in 2010–2020 increased from €24,800 to €32,300. The largest increases in spending were recorded in Lithuania (133.6%), Romania (108.6%) and Bulgaria (103.6%). In Greece and Cyprus, expenditure decreased by 24.7% and 3.0%, respectively, while in Italy the increase in expenditure was only 1.7%. A particularly significant decline in spending in Greece occurred in 2013 and was a negative effect of the 2008 financial crisis (Leventi & Picos, 2019).

However, the spread of expenditure in the period being examined was very large (Figure 3). The highest figure (Luxembourg) was more than eight times higher than the lowest (Bulgaria). The classification of countries in terms of wages measured in *PPS* per person is not significantly different from the classification of countries in terms of wages measured in euro. Only the spread of salaries decreases significantly. The highest *PPS*-based figure (Luxembourg) was 3.5 times higher than the lowest (Bulgaria).

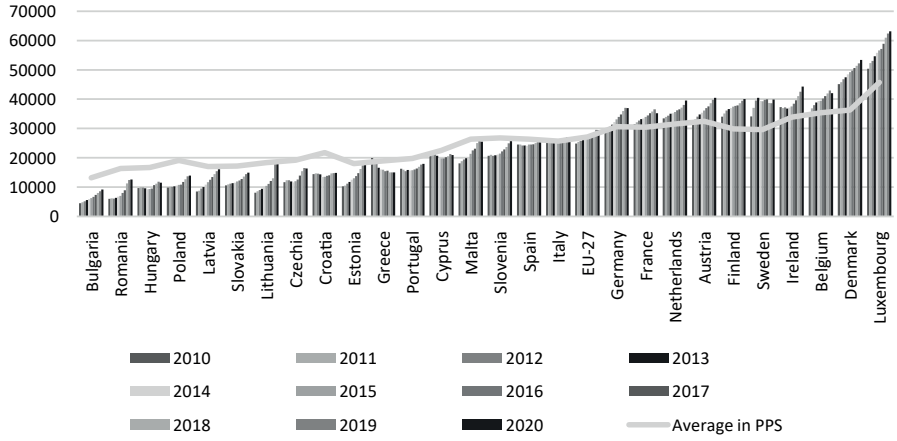

Figure 3. Average gross remuneration per person (euro and *PPS*) in EU countries in 2010–2020

Source: based on Eurostat data.

Between 2009 and 2020, the average *ITR* for the 27 EU countries increased from 36.8% to 37.8%. A significant increase in the indicator was recorded in the years 2010–2013 (by 1.4 percentage points on average) and was due to countries' reactions to the negative fiscal effects of the 2008 financial crisis (Leventi & Picos, 2019), among other factors. In subsequent years, *ITR* was very stable, and there were no significant changes in the burden in the first year of the pandemic.

ITR across EU member states varies both in terms of total weight and its structure (Figure 4). For example, in 2020, the lowest load of 23.5% was recorded in Malta, while the highest (44.1%) was recorded in Italy. On average, the largest part of *ITR* was contributions and taxes paid by employers (almost 40.2% of the total burden), followed by *PIT* (almost 34.2%), and the smallest part was contributions paid by employees (less than 25.7%).

In terms of the division of labour costs between employer and employee, the former incurs the highest costs in Italy, France and the Czech Republic, and the lowest in Lithuania, Denmark and Romania. In turn, employees in the latter three countries incur the highest labour costs, while they incur the lowest in Estonia, Cyprus and Bulgaria.

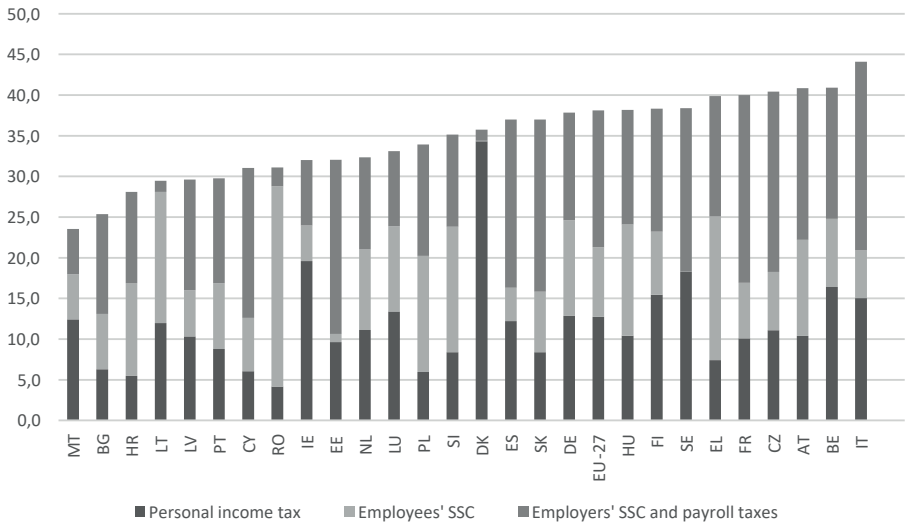

Figure 4. *ITR* on labour in EU member states in 2020

Source: based on (European Commission Data, n.d.).

The average expenditure on net wages per person resulted from the development of gross expenditure and *ITR*, which in many countries were affected by the financial crisis of 2008. In the years 2010–2020, the average expenditure on net wages per person in the EU increased from €19,800 to €23,700. The

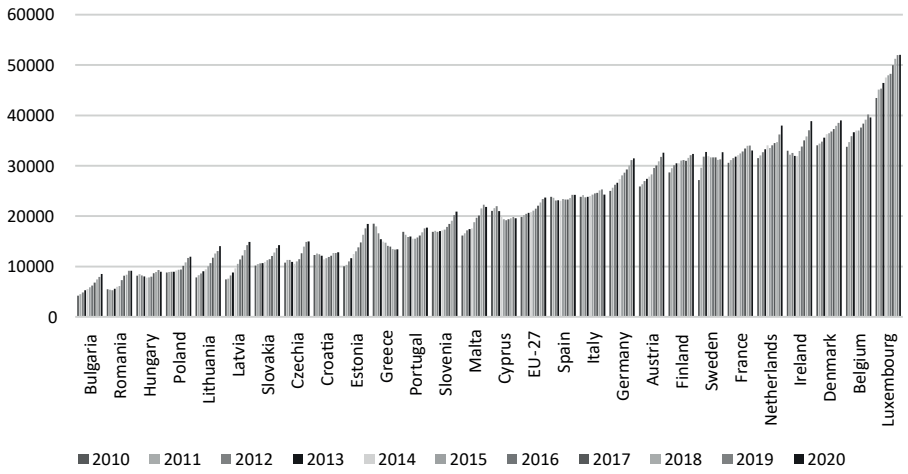

Figure 5. Average spending on net wages per employee in EU countries in 2010–2020 (euro)

Source: based on Eurostat data.

highest increases in expenditure were recorded in Bulgaria (100.8%), Latvia (99.8%) and Estonia (83.3%). A decline in salary expenditure was recorded in Greece (27.5%) and Cyprus (7.0%), while a minimal increase in expenditure was recorded in Italy (1.9%) and Spain (1.6%).

The expenditure discussed varied across EU member states (Figure 5). The lowest spending was recorded in Bulgaria, Romania and Hungary, and the highest in Luxembourg, Belgium and Denmark.

6.2. Pension expenditure and taxation of pension benefits

The average spending on pensions in the EU per person increased from €10,600 to €13,300 in the years analysed. The amount of this expenditure varies greatly across EU member states (Figure 9). In the Western EU countries and in Scandinavia, the expenditure is much higher than in the countries of Central and Eastern Europe. The highest spending was recorded in Luxembourg, Denmark and Ireland, and the lowest in Bulgaria, Romania and Lithuania. The average pension expenditure in Luxembourg was more than 14 times higher than in Bulgaria.

The classification of countries is broadly similar in terms of spending per purchasing power (*PPS*), with spending gaps across countries narrowing significantly (Figure 6). Expenditure in *PPS* turned out to be higher than expenditure in euro in the countries of Central and Eastern Europe and in Portugal,

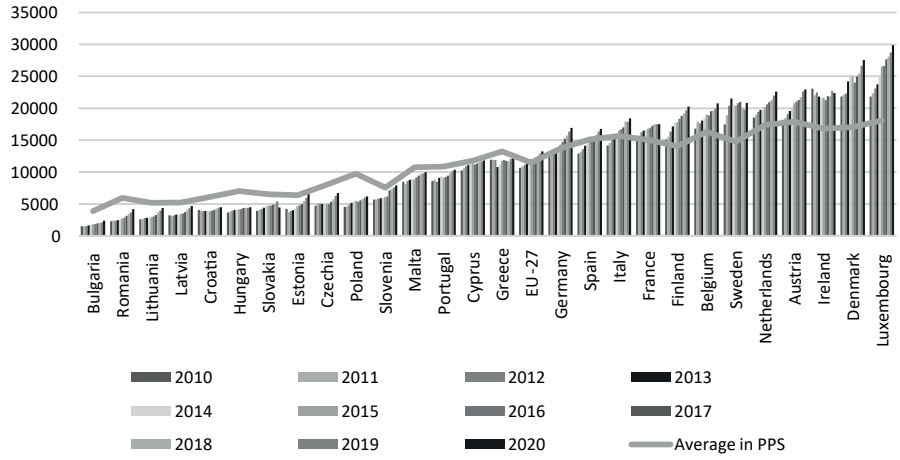

Figure 6. Classification of EU countries in terms of average gross expenditure on pensions per person in 2010–2020 (euro and *PPS*)

Source: based on Eurostat data.

Greece, Cyprus and Malta, while in the remaining countries the reverse was the case. As a result, the spending ratio between Luxembourg and Bulgaria decreased to four.

The average *PETR* in EU countries (EU-27) increased from 14.0% to 15.4% in the period under study. The largest increase in the indicator was recorded in Portugal, Cyprus, Greece and Ireland, but it occurred mainly in the years 2013–2018 and was mainly the result of the PIT reform resulting from the 2008 financial crisis. In turn, the largest decline in *PETR* was recorded in Belgium, Sweden, Poland and Malta.

PETR was very diverse across individual countries (Figure 7). A significantly lower *PETR* was recorded in the countries of Central and Eastern Europe than in Western Europe and Scandinavia. In Bulgaria, the Czech Republic, Lithuania and Slovakia, as well as in Hungary from 2014, pensions were exempt from taxation and contributions. However, the highest *PETR* was recorded in the Netherlands, Denmark and Italy.

The average net expenditure on pensions per person in the EU countries (EU-27) increased from €9,300 to €11,400 in the period in question. The largest increases in spending were recorded in Romania (81.0%), Lithuania (66.8%) and Bulgaria (60.5%). In Greece and Ireland, spending fell by 3.9% and 9.1%, respectively. Cyprus saw an 8.1% increase in spending. Expenditure varied across countries. Much higher expenditure was incurred by the countries of Western Europe and Scandinavia, mainly Luxembourg, Ireland and Denmark, and the lowest by the countries of Central and Eastern Europe, including Bulgaria, Romania and Lithuania (Figure 8). The purchasing power of the expenditure

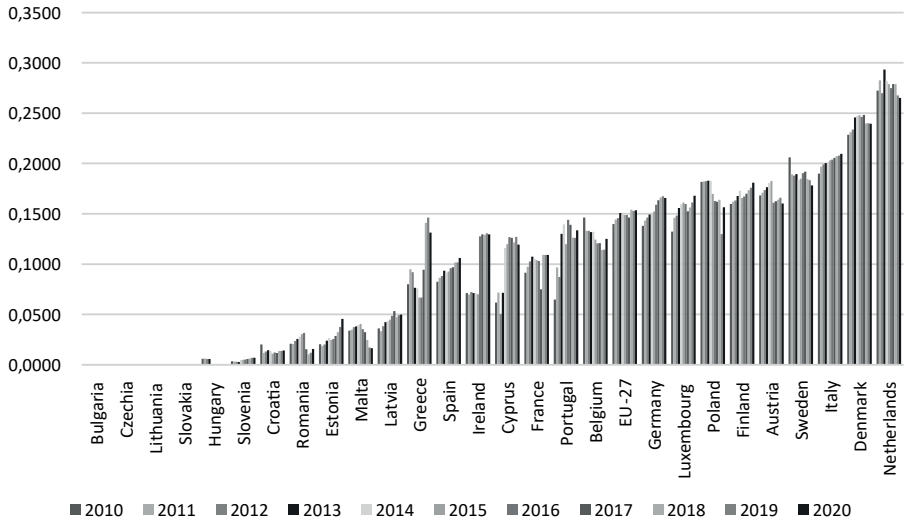

Figure 7. *PETR* in EU countries in the years 2010–2020

Source: based on Eurostat data.

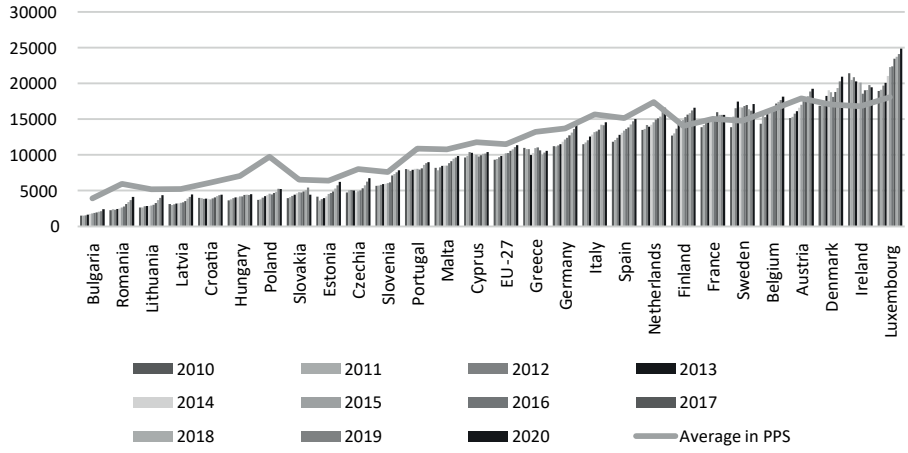

Figure 8. Average net spending on pensions per person in EU countries in 2010–2020 (euro and *PPS*)

Source: based on Eurostat data.

in question was also much higher in Western and Scandinavian countries, but in the countries of Central and Eastern Europe, expenditure expressed in *PPS* was clearly higher than expenditure expressed in euro.

6.3. Replacement rates

The average replacement rate in gross terms (*GPRR*) (before taxes and contributions for salaries and pensions) in the EU in 2022 was 54.9% (Figure 9). However, the spread of values of the indicator between countries is very large. The lowest rates were recorded in Lithuania (18.2%), Ireland (26.2%) and Estonia (28.1%), and the highest in Greece (80.8%), Spain (80.4%) and Italy (76.1%).

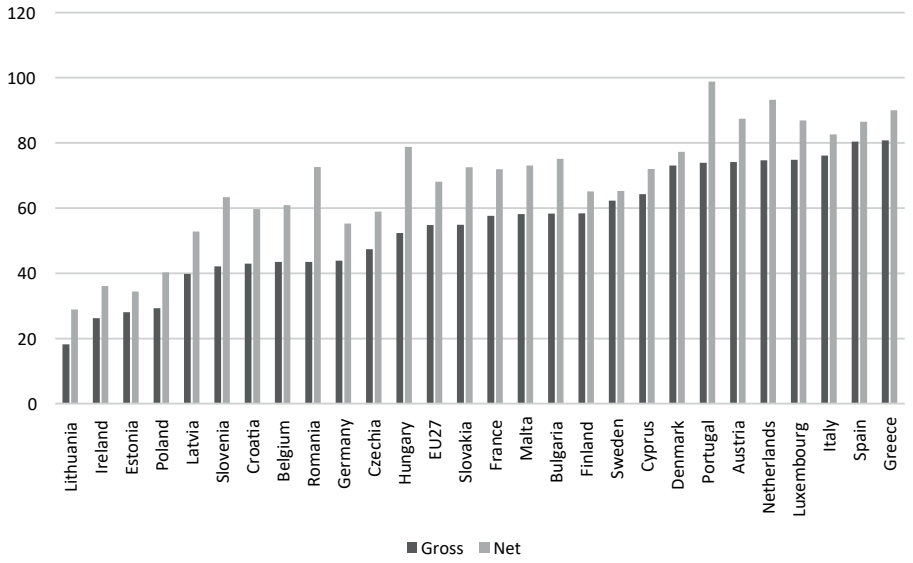

Figure 9. Gross pension replacement rate (*GPRR*) and net pension replacement rate (*NPRR*) in UE countries in 2022

Source: based on (OECD, 2023).

NPRR (after taxes and contributions on wages and pensions) is higher than the gross rate in all EU countries. The average replacement rate for the entire EU increases to 68.1%, while the spread of the rate between member states decreases (Figure 14). The lowest rate was again recorded in Lithuania (28.9%), but it was only three times lower than the highest rate, which was recorded in Portugal (98.8%).

Figure 10 shows the classification of EU countries in 2022 in terms of gross wealth pension (*GPW*) and net wealth pension (*NPW*) for men. The Figure shows that, in general, in most Western EU countries and the Mediterranean countries, the *GPW* is clearly higher than in most Central and Eastern European countries. However, after tax, the situation changes in many countries. Such countries as Romania, Hungary, Slovenia and Bulgaria significantly improve

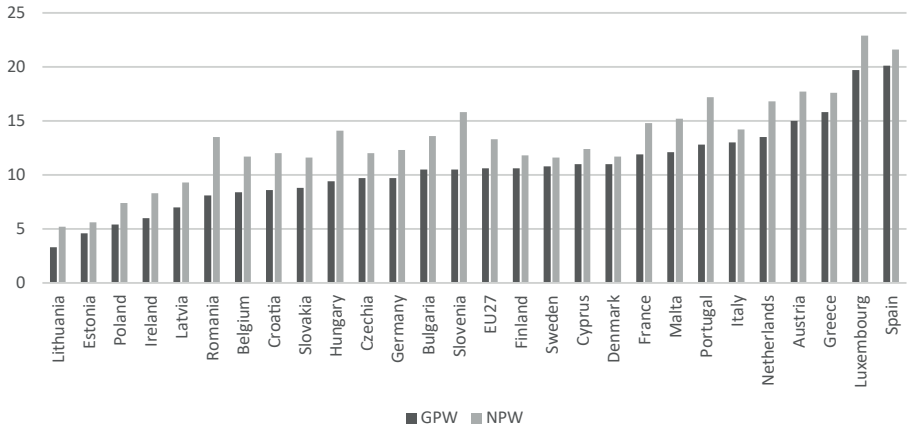

Figure 10. Classification of EU countries in 2022 according to the *GPW* and *NPW* indicators for men

Source: based on (OECD, 2023).

their positions, while Finland, Sweden, Cyprus, Denmark and Italy fall in the classification. In the case of women, the results of the analysis across countries are very similar, but the average *GPW* and *NPW* indicators for the entire EU are higher than the indicators for men. This is mainly due to the longer life expectancy of women (OECD, 2023).

6.4. Results of multidimensional comparative analysis

Table 4 presents classifications of countries in terms of two dimensions: the conditions for taxation of remuneration (Rank 1) and the conditions for taxation of retirement benefits (Rank 2). Mean is averaged data after unitarization of individual variables.

Table 4 generally shows that better remuneration taxation conditions exist in Western European countries than in Central and Eastern European countries. This is mainly due to much higher expenditure on gross and net salaries in the former compared to the latter (Figures 2 and 6). In this case, the total tax burden imposed on remuneration is less important, because the taxation rules vary greatly across individual countries (Figure 4). The highest *ITR* is recorded by Denmark, Romania and Lithuania, and the lowest by Cyprus, Bulgaria and Slovakia.

The taxation conditions for pension benefits are also generally better in Western European countries than in Central and Eastern European countries, despite the fact that in the former the tax rate on pension benefits (*PETR*) is

Table 4. Ranking of EU countries in terms of conditions of taxation of remuneration (Rank 1) and public pension (Rank 2)

Country	RANK 1: MEAN	RANK 2: MEAN	RANK 1: SCORE	RANK 2: SCORE
Belgium	0.702941	0.532712	1	13
Luxembourg	0.698623	0.880144	2	1
Austria	0.676405	0.761461	3	3
Sweden	0.634892	0.550966	4	11
Spain	0.608214	0.839146	5	2
France	0.601081	0.597740	6	9
Cyprus	0.574375	0.548604	7	12
Italy	0.566580	0.632108	8	7
Finland	0.560158	0.528033	9	14
Denmark	0.534153	0.616946	10	8
Ireland	0.521643	0.448817	11	17
Germany	0.497056	0.458843	12	16
Estonia	0.481840	0.217795	13	26
Slovakia	0.472773	0.429921	14	19
Malta	0.450844	0.585385	15	10
Czechia	0.448581	0.419530	16	20
Greece	0.426956	0.716641	17	4
Portugal	0.426905	0.636640	18	6
Netherlands	0.426299	0.678776	19	5
Croatia	0.417008	0.385998	20	22
Slovenia	0.401654	0.445860	21	18
Hungary	0.347027	0.466463	22	15
Poland	0.335962	0.233709	23	25
Bulgaria	0.300068	0.412303	24	21
Latvia	0.260655	0.245448	25	24
Lithuania	0.179612	0.124439	26	27
Romania	0.125827	0.368056	27	23

Source: own analysis.

much higher than in the latter (Figure 7). In particular, this is due to significantly higher gross and net pension expenditure per person in Western countries than in Central and Eastern European countries (Figure 6 and 8). Most of the former also recorded a higher rate of replacing wages with pensions.

6.5. Agglomeration analysis

The ranking results were analysed using the agglomeration method. The following results were obtained: 6 groups (6 clusters) regarding the conditions of taxation of wages (Figure 11 and Table 5) and 2 groups (2 clusters) in the field of taxation of pension benefits (Figure 12 and Table 6).

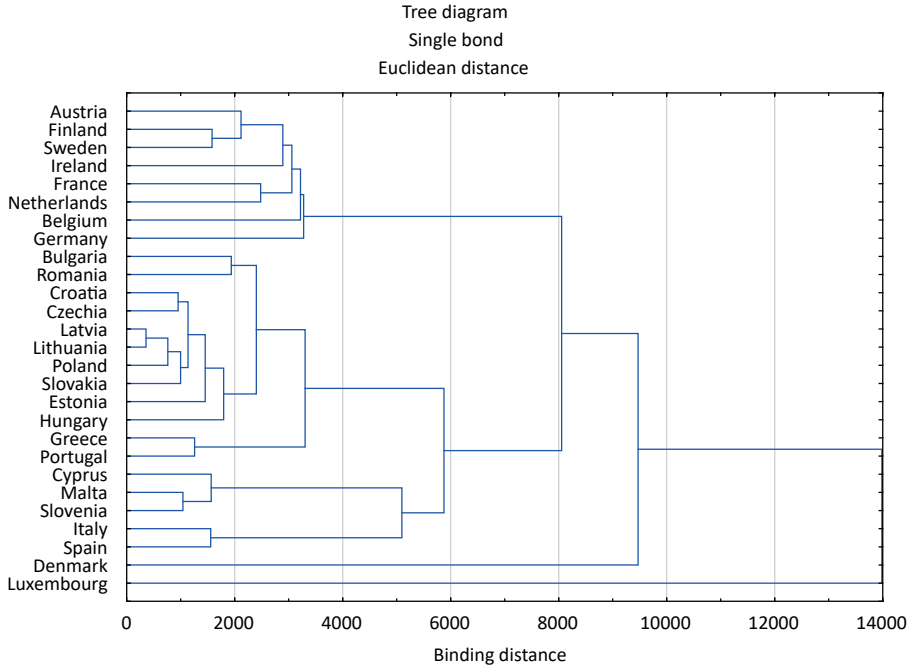

Figure 11. Grouping of EU countries using the agglomeration method in terms of taxation of salary expenses

Source: own calculations.

Taking into account the conditions for taxation of salary expenses, Denmark and Luxembourg constitute separate clusters. These countries clearly differ from other countries with higher average gross and net expenditure on wages. Moreover, these countries recorded a relatively low fiscal burden imposed

Table 5. Grouping of EU Member States using the agglomeration method in terms of remuneration taxation conditions

Country	Cluster	X_1	X_2	X_3	X_4
Denmark	1	49109.31	36463.67	20.90576	1.360517
Luxembourg	2	56857.55	48104.46	24.07332	9.219679
Italy	3	26020.30	24337.21	23.8007	23.18912
Spain	3	24695.24	23519.59	16.01317	20.66356
Cyprus	4	20816.43	20213.56	13.08473	18.41868
Malta	4	21816.73	19009.72	10.64343	5.576068
Slovenia	4	22283.86	18078.63	21.28347	11.34707
Austria	5	36316.25	28879.57	15.83743	18.62296
Belgium	5	39670.99	37244.51	16.91304	16.12316
Finland	5	37280.16	30765.73	24.59416	15.12468
France	5	34022.14	32447.03	28.78754	23.09641
Germany	5	33126.34	28128.63	25.08654	13.24185
Ireland	5	38980.48	34099.53	23.21672	8.022056
Netherlands	5	35920.94	34045.06	34.38331	11.32244
Sweden	5	38798.25	31212.8	23.88081	20.10379
Bulgaria	6	6589.063	6113.04	17.76654	12.27906
Croatia	6	14310.31	12253.07	12.62212	11.21446
Czechia	6	13362.05	12156.47	21.02964	22.16659
Estonia	6	13620.68	13587.65	17.97300	21.40355
Greece	6	16544.75	15127.29	18.24333	14.78820
Hungary	6	10241.25	8398.241	18.29423	14.09105
Latvia	6	11949.45	10752.89	28.10606	13.60554
Lithuania	6	11702.55	10493.57	22.21888	1.355045
Poland	6	11318.82	9835.81	20.23988	13.69864
Portugal	6	16370.14	16373.07	16.84640	12.92705
Romania	6	8354.274	6913.999	24.77601	2.325089
Slovakia	6	12333.32	11673.72	16.88872	21.16879

Source: own calculations.

on the employer for maintaining workplaces. Denmark constitutes a separate cluster, which is due to its characteristic pension system.

The third cluster consists of Italy and Spain. These countries seem very similar to each other, particularly in terms of the amount of expenditure on wages and the fiscal burden imposed on the employer.

The fourth cluster consists of Cyprus, Malta and Slovenia. These countries are characterised by lower expenditure on wages than the previous countries, but these expenditures are very similar. However, visible differences occur in the taxation of remuneration imposed on the employee and the employer.

The fifth cluster consists of the remaining Western European countries, except Greece and Portugal. These are countries with a similar level of socio-economic development and are characterized by much higher expenditure on wages than the countries in clusters 3 and 4. Taxation of wages, as in the case of the previous clusters, is quite diverse, although the tax burden imposed on the employee is, with the exception of Austria, higher than the burden imposed on the employer.

The last cluster consists of the countries of Central and Eastern Europe, together with Greece and Portugal. These countries are characterised by much lower expenditure on salaries than previous countries. Taxation of wages is

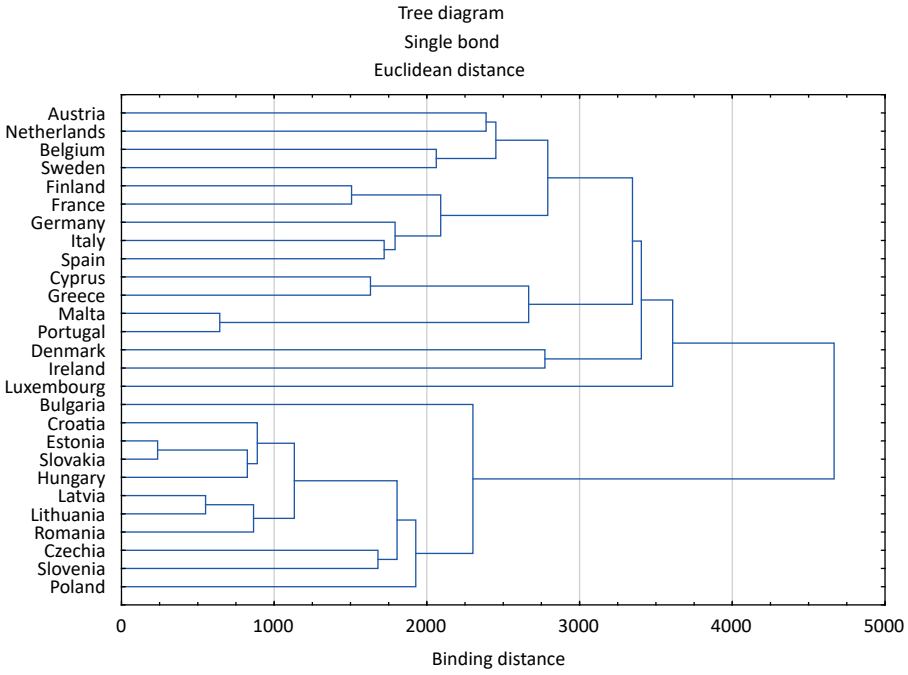

Figure 12. Grouping of EU countries using the agglomeration method in terms of pension taxation conditions

Source: own calculations.

Table 6. Grouping of EU Member States using the agglomeration method in terms of pension taxation conditions

Country	Cluster	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
Austria	1	20538.43	17059.1	0.169667	17880.41	74.1	87.4	15	17.7	16.6	19.5	0.280039
Belgium	1	18779.70	16405.78	0.126849	16302.01	43.5	60.9	8.4	11.7	9.1	12.8	0.357278
Cyprus	1	11185.93	10045.08	0.101089	11778.73	64.3	72	11	12.4	12	13.4	0.243500
Denmark	1	24480.02	18580.76	0.240714	17037.69	73.1	77.3	11	11.7	12.2	12.9	0.419390
Finland	1	17807.46	14789.33	0.168985	14048.06	58.4	65.1	10.6	11.8	12	13.3	0.514882
France	1	16706.92	15014.04	0.101213	15053.12	57.6	71.9	11.9	14.8	13.5	16.8	0.586232
Germany	1	14548.97	12288.38	0.154552	13704.54	43.9	55.3	9.7	12.3	10.8	13.6	0.571192
Greece	1	11716.19	10572.41	0.097217	13227.49	80.8	90	15.8	17.6	17.5	19.5	0.257619
Ireland	1	22057.55	19908.81	0.097451	16804.3	26.2	36.1	6	8.3	6.5	9.0	0.410290
Italy	1	16361.73	13045.38	0.202265	15665.66	76.1	82.6	13	14.2	14.8	16.0	0.378506
Luxembourg	1	25785.1	21780.18	0.154576	18081.42	74.8	86.9	19.7	22.9	21.8	25.3	0.530830
Malta	1	9066.93	8782.091	0.031826	10765.63	58.2	73.1	12.1	15.2	13.4	16.9	0.256101
Netherlands	1	20362.51	14730.12	0.27681	17381.99	74.7	93.2	13.5	16.8	14.5	18.1	0.370455
Portugal	1	9310.171	8194.577	0.118833	10864.33	73.9	98.8	12.8	17.2	14.4	19.3	0.296211
Spain	1	14758.20	13358.7	0.094283	15125.75	80.4	86.5	20.1	21.6	22.7	24.4	0.267952
Sweden	1	20147.73	16361.44	0.188151	14760.26	62.3	65.3	10.8	11.6	11.7	12.6	0.459870
Bulgaria	2	1825.067	1825.067	0	3897.101	58.3	75.1	10.5	13.6	11.9	15.3	0.641427
Croatia	2	4074.885	4019.549	0.013555	6117.869	43	59.7	8.6	12	9.6	13.3	0.443267
Czechia	2	5323.974	5323.974	0	8020.164	47.4	58.9	9.7	12	10.8	13.4	0.527800
Estonia	2	4777.291	4638.945	0.027608	6378.965	28.1	34.4	4.6	5.6	5.3	6.5	0.512434
Hungary	2	4150.538	4142.293	0.002129	7040.048	52.4	78.8	9.4	14.1	10.6	16.0	0.451401
Latvia	2	3637.928	3475.098	0.044165	5226.269	39.8	52.8	7	9.3	8	10.6	0.923196
Lithuania	2	3177.975	3177.975	0	5174.781	18.2	28.9	3.3	5.2	3.7	5.9	0.818186
Poland	2	5372.307	4471.863	0.168781	9748.034	29.3	40.3	5.4	7.4	5.4	7.4	0.426430
Romania	2	2931.378	2871.651	0.021265	5943.734	43.5	72.6	8.1	13.5	9.2	15.4	0.706103
Slovakia	2	4604.712	4603.598	0.000242	6529.703	54.9	72.5	8.8	11.6	9.9	13.1	0.452350
Slovenia	2	6486.458	6454.409	0.004777	7581.584	42.1	63.4	10.5	15.8	11.9	17.9	0.731182

Source: own calculations.

quite different between the employee and the employer, but in average categories the former is higher than the latter. At the same time, in most of the countries surveyed, the tax burden imposed on the employee is lower than the tax burden imposed in the countries belonging to cluster 5. In the countries surveyed, the tax burden between the employee and the employer is higher than the tax burden imposed in the countries belonging to cluster 4.

The agglomeration analysis in terms of taxation of pension expenditure divided the EU countries into two clusters. The first are the countries of Western Europe, and the second are the countries of Central and Eastern Europe (Table 6). This division shows that in Western countries the conditions for taxation of pension benefits are not as diverse as in the case of taxation of wages.

In general, Western EU countries offer much better conditions for the taxation of public pensions than the countries of Central and Eastern Europe. This first group of countries is characterized by significantly higher expenditure on pension benefits in gross, net and purchasing power terms, even though the effective tax burden imposed on pensions is clearly higher. Most Western countries also record higher replacement rates and retirement wealth indicators for both women and men. In Western countries, the relationship between the amount of an employee's fiscal payments and average annual gross pension expenditure per person is also clearly lower than in the countries of Central and Eastern Europe.

Conclusions

This paper has attempted to present and assess the conditions of taxation of public pensions in EU member states at the stage of work and at the stage of payment of pension benefits. Taxation systems were assessed separately for two life stages, using separate diagnostic variables. For this purpose, the multi-criteria comparative analysis method and the agglomeration analysis method were used.

The selection of variables and research methods resulted primarily from the approach to the analysis, namely, from the employee's point of view and the usefulness of the results achieved in making decisions about the choice of country of work and country of retirement.

The results of both the multi-criteria comparative analysis and the agglomeration analysis are very convergent. The analysis shows that the fundamental differences in the conditions of taxation of wages and pension benefits in EU countries should be divided into Western European countries and Central and Eastern European countries. In the former countries, there was a significantly higher tax burden imposed on the employee on the remuneration re-

Appendix A1. Diagnostic variables

Country	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
Austria	36 316.25	28 879.57	15.84	18.62	20 538.43	17 059.10	0.17	17880.41	74.10	87.40	15	17.7	16.6	19.5	0.28
Belgium	39 670.99	37 244.51	16.91	16.12	18 779.70	16 405.78	0.13	16302.01	43.50	60.90	8.4	11.7	9.1	12.8	0.36
Bulgaria	6 589.06	6 113.04	17.77	12.28	1 825.07	1 825.07	0.00	3897.101	58.30	75.10	10.5	13.6	11.9	15.3	0.64
Croatia	14 310.31	12 253.07	12.62	11.21	4 074.88	4 019.55	0.01	6117.869	43.00	59.70	8.6	12	9.6	13.3	0.44
Cyprus	20 816.43	20 213.56	13.08	18.42	11 185.93	10 045.08	0.10	11778.73	64.30	72.00	11	12.4	12	13.4	0.24
Czechia	13 362.05	12 156.47	21.03	22.17	5 323.97	5 323.97	0.00	8020.164	47.40	58.90	9.7	12	10.8	13.4	0.53
Denmark	49 109.31	36 463.67	20.91	1.36	24 480.02	18 580.76	0.24	17037.69	73.10	77.30	11	11.7	12.2	12.9	0.42
Estonia	13 620.68	13 587.65	17.97	21.40	4 777.29	4 638.94	0.03	6378.965	28.10	34.40	4.6	5.6	5.3	6.5	0.51
Finland	37 280.16	30 765.73	24.59	15.12	17 807.46	14 789.33	0.17	14048.06	58.40	65.10	10.6	11.8	12	13.3	0.51
France	34 022.14	32 447.03	28.79	23.10	16 706.92	15 014.04	0.10	15053.12	57.60	71.90	11.9	14.8	13.5	16.8	0.59
Germany	33 126.34	28 128.63	25.09	13.24	14 548.97	12 288.38	0.15	13704.54	43.90	55.30	9.7	12.3	10.8	13.6	0.57
Greece	16 544.75	15 127.29	18.24	14.79	11 716.19	10 572.41	0.10	13227.49	80.80	90.00	15.8	17.6	17.5	19.5	0.26
Hungary	10 241.25	8 398.24	18.29	14.09	4 150.54	4 142.29	0.00	7040.048	52.40	78.80	9.4	14.1	10.6	16	0.45
Ireland	38 980.48	34 099.53	23.22	8.02	22 057.55	19 908.81	0.10	16804.3	26.20	36.10	6	8.3	6.5	9	0.41
Italy	26 020.30	24 337.21	23.80	23.19	16 361.73	13 045.38	0.20	15665.66	76.10	82.60	13	14.2	14.8	16	0.38
Latvia	11 949.45	10 752.89	28.11	13.61	3 637.93	3 475.10	0.04	5226.269	39.80	52.80	7	9.3	8	10.6	0.92
Lithuania	11 702.55	10 493.57	22.22	1.36	3 177.97	3 177.97	0.00	5174.781	18.20	28.90	3.3	5.2	3.7	5.9	0.82
Luxembourg	56 857.55	48 104.46	24.07	9.22	25 785.10	21 780.18	0.15	18081.42	74.80	86.90	19.7	22.9	21.8	25.3	0.53
Malta	21 816.73	19 009.72	10.64	5.58	9 066.93	8 782.09	0.03	10765.63	58.20	73.10	12.1	15.2	13.4	16.9	0.26
Netherlands	35 920.94	34 045.06	34.38	11.32	20 362.51	14 730.12	0.28	17381.99	74.70	93.20	13.5	16.8	14.5	18.1	0.37
Poland	11 318.82	9 835.81	20.24	13.70	5 372.31	4 471.86	0.17	9748.034	29.30	40.30	5.4	7.4	5.4	7.4	0.43
Portugal	16 370.14	16 373.07	16.85	12.93	9 310.17	8 194.58	0.12	10864.33	73.90	98.80	12.8	17.2	14.4	19.3	0.30
Romania	8 354.27	6 914.00	24.78	2.33	2 931.38	2 871.65	0.02	5943.734	43.50	72.60	8.1	13.5	9.2	15.4	0.71
Slovakia	12 333.32	11 673.72	16.89	21.17	4 604.71	4 603.60	0.00	6529.703	54.90	72.50	8.8	11.6	9.9	13.1	0.45
Slovenia	22 283.86	18 078.63	21.28	11.35	6 486.46	6 454.41	0.00	7581.584	42.10	63.40	10.5	15.8	11.9	17.9	0.73
Spain	24 695.24	23 519.59	16.01	20.66	14 758.20	13 358.70	0.09	15125.75	80.40	86.50	20.1	21.6	22.7	24.4	0.27
Sweden	38 798.25	31 212.80	23.88	20.10	20 147.73	16 361.44	0.19	14760.26	62.30	65.30	10.8	11.6	11.7	12.6	0.46

Appendix A2. Unitarization

Country	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
Austria	0.5913682	0.5421711	0.781212	0.79087	0.781024	0.763415	0.387061	0.985829	0.892971	0.83691	0.696429	0.706215	0.678947	0.701031	0.946242
Belgium	0.6581046	0.7413769	0.735904	0.676379	0.7076212	0.7306756	0.541745	0.874551	0.404153	0.457797	0.303571	0.367232	0.284211	0.35567	0.832605
Bulgaria	0	0	0.699952	0.50032	0	0	1	0	0.640575	0.660944	0.428571	0.474576	0.431579	0.484536	0.41455
Croatia	0.1536001	0.1462212	0.916651	0.451561	0.0938988	0.1099709	0.951033	0.156565	0.396166	0.440629	0.315476	0.384181	0.310526	0.381443	0.706093
Cyprus	0.2830275	0.3357954	0.897165	0.781514	0.3906865	0.411925	0.634806	0.555658	0.736422	0.616595	0.458333	0.40678	0.436842	0.386598	1
Czechia	0.1347363	0.1439206	0.562499	0.953168	0.1460309	0.1753388	1	0.290677	0.466454	0.429185	0.380952	0.384181	0.373684	0.386598	0.581725
Denmark	0.845863	0.7227819	0.567718	0.000251	0.9455311	0.839669	0.1304	0.926416	0.876997	0.692418	0.458333	0.367232	0.447368	0.360825	0.741222
Estonia	0.1398813	0.1780032	0.691255	0.918221	0.1232145	0.1410103	0.900262	0.174972	0.158147	0.078684	0.077381	0.022599	0.084211	0.030928	0.604332
Finland	0.6105435	0.5870889	0.41235	0.630649	0.6670439	0.649671	0.389525	0.715646	0.642173	0.517883	0.434524	0.372881	0.436842	0.381443	0.60073
France	0.545731	0.6271278	0.235712	0.995754	0.6211116	0.660932	0.634358	0.786503	0.629393	0.615165	0.511905	0.542373	0.515789	0.561856	0.495756
Germany	0.5279108	0.524288	0.39161	0.544416	0.5310471	0.5243422	0.441667	0.691428	0.410543	0.377682	0.380952	0.40113	0.373684	0.396907	0.517885
Greece	0.1980504	0.2146688	0.679868	0.615238	0.4128175	0.4383511	0.648796	0.657796	1	0.874106	0.744048	0.700565	0.726316	0.701031	0.979227
Hungary	0.0726536	0.0544207	0.677724	0.583308	0.0970562	0.1161219	0.992308	0.221579	0.546326	0.713877	0.363095	0.502825	0.363158	0.520619	0.694126
Ireland	0.6443682	0.6664811	0.470373	0.305349	0.8444263	0.906221	0.647951	0.909962	0.127796	0.103004	0.160714	0.175141	0.147368	0.159794	0.754611
Italy	0.3865492	0.4339976	0.445773	1	0.6067044	0.5622776	0.269301	0.829688	0.92492	0.76824	0.577381	0.508475	0.584211	0.520619	0.801373
Latvia	0.1066351	0.1104953	0.264418	0.561072	0.0756618	0.0826871	0.840449	0.093707	0.345048	0.341917	0.220238	0.231638	0.226316	0.242268	0
Lithuania	0.1017236	0.1043197	0.512405	0	0.0564652	0.0677975	1	0.090077	0	0	0	0	0	0	0.154494
Luxembourg	1	1	0.43429	0.3602	1	1	0.441579	1	0.904153	0.829757	0.97619	1	0.952632	1	0.577267
Malta	0.3029268	0.3071265	1	0.193323	0.3022476	0.3486336	0.885026	0.484234	0.638978	0.632332	0.52381	0.564972	0.510526	0.56701	0.981461
Netherlands	0.5835042	0.665184	0	0.456506	0.7736819	0.6467042	0	0.95069	0.902556	0.919886	0.607143	0.655367	0.568421	0.628866	0.813218
Poland	0.0940899	0.0886555	0.595767	0.565336	0.1480482	0.1326374	0.390263	0.412493	0.177316	0.16309	0.125	0.124294	0.089474	0.07732	0.730864
Portugal	0.1945767	0.2443363	0.738711	0.529998	0.3123996	0.3191918	0.570707	0.491192	0.889776	1	0.565476	0.677966	0.563158	0.690722	0.922448
Romania	0.0351157	0.0190744	0.40469	0.044428	0.0461732	0.0524469	0.923177	0.144288	0.404153	0.625179	0.285714	0.468927	0.289474	0.489691	0.319396
Slovakia	0.1142716	0.1324242	0.736928	0.907469	0.1160117	0.139239	0.999127	0.185599	0.586262	0.623748	0.327381	0.361582	0.326316	0.371134	0.692731
Slovenia	0.3122194	0.2849532	0.551807	0.457635	0.1945486	0.2319878	0.982741	0.259757	0.381789	0.493562	0.428571	0.59887	0.431579	0.618557	0.2825
Spain	0.3601894	0.4145264	0.773809	0.884329	0.5397793	0.577979	0.659395	0.791624	0.99361	0.824034	1	0.926554	1	0.953608	0.964025
Sweden	0.6407431	0.5977354	0.442399	0.858692	0.7647179	0.7284536	0.320287	0.765857	0.704473	0.520744	0.446429	0.361582	0.421053	0.345361	0.681667

ceived and higher taxation of pension benefits. However, the predominance of these burdens does not translate into a deterioration of the overall conditions for taxation of salaries and pension benefits. In Western countries, the amount of expenditure on net wages and the amount of expenditure on net pension benefits is much higher than in the countries of Central and Eastern Europe, which also translates into better results in terms of other variables and, therefore, into an overall better assessment of the conditions for the taxation of pension benefits.

In the group of Western countries, the exceptions are Greece and Portugal, which recorded relatively low expenditure on salaries at the level of Central and Eastern European countries. It can be assumed that this was mainly a consequence of the financial crisis of 2008.

As a result, it can be concluded that in such conditions, effective rates of taxation of wages and pension benefits are not important factors in deciding on the choice of the country of work or the country of retirement. After retirement, choosing a country of residence other than for work is less important from a fiscal point of view, because the pension will be paid by the country of work, and double taxation agreements between EU Member States show that such pensions are exempt from income tax on the basis of reciprocity.

In the context of the analysis results, an employee in the common market should be guided mainly by the amount of net salary expenses or the amount of net pension benefits, which determine the remaining variables to the greatest extent. Then the best conditions for taxation of wages are offered by countries such as Belgium, Luxembourg and Austria, and the best conditions for taxation of pension benefits are offered by countries such as Luxembourg, Spain and Austria.

References

Armstrong, A., Davis, P., & Ebell, M. (2015). An economic analysis of pension tax proposals. *National Institute of Economic and Social Research*. https://eprints.lse.ac.uk/86276/1/CFMDP2015-33-Paper.pdf

Barrios, S., Mascarola, C., Figari, F., & Luca F. G. (2020). Size and distributional pattern of pension-related tax expenditures in European countries. *International Tax and Public Finance*, *27*, 1287–1320.

Cremer, H., & Pestieau, P. (2016). *Taxing pensions*. CESifo Working Paper Series, 5930. https://doi.org/10.2139/ssrn.2799616

Cremer, H., Lozachmeur, J. M., & Pestieau, P. (2008). Social security and retirement decision: A positive and normative approach. *Journal of Economic Surveys*, *22*, 213–233.

Diamond, P. (2009). Taxes and pensions. *Southern Economic Journal*, *76*, 2–15.

Ebbinghaus, B. (2021). Inequalities and poverty risks in old age across Europe: The double-edged income effect of pension systems. *Social Policy & Administration*, *55*, 440–455.

Euromod. (2011–2022). *Euromod country reports, models 2011–2022. Tax-benefit microsimulation model for the European Union*. https://euromod-web.jrc.ec.europa.eu/resources/country-reports

European Commission. (2001). The elimination of tax obstacles to the cross-border provision of occupational pensions, COM (2001) 214 final.

European Commission. (2022). *Taxation trends in the European Union, Taxation and Customs Union*. Luxembourg: Publications Office of the European Union.

European Commission. (n.d.). Taxes in Europe Database v4. https://ec.europa.eu/taxation_customs/tedb/#/home

European Commission Data. (n.d.). Data on Taxation Trends. https://taxation-customs.ec.europa.eu/taxation/economic-analysis-taxation/data-taxation-trends_en

Ferrarini, T., & Nelson, K. (2003). Taxation of social insurance and redistribution: A comparative analysis of ten welfare states. *Journal of European Social Policy*, *13*(1), 21–33.

Fratica-Dragomir A. M. (2023). The public pension system in Austria, Lichtenstein, Germany. *Univers Strategic—Revistă de Studii Strategice Interdisciplinare și de Securitate*, *4*(56), 77–99.

Genser, B., & Holtzmann, R. (2021). Frontloaded income taxation of old-age pensions: For efficiency and fairness in a world of international labor mobility. *CESifo Economic Studies*, *67*(1), 61–77.

Ivaškaitė-Tamošiūnė, V., & Thiemann, A. (2021). *The budgetary and redistributive impact of pension taxation in the EU: A microsimulation analysis*. European Commission, JRC Working Papers on Taxation and Structural Reforms, 08/2021. https://joint-research-centre.ec.europa.eu/document/download/1be7f6da-86e2-4994-a237-24fa7bbe14fe_en

Iwacewicz-Orłowska, A., & Sokołowska, D. (2018). Ranking of EU countries in terms of the value of environmental governance indicators in 2010 and 2015, using the Hellwig method. *Ekonomia i Środowisko, 23*(3), 50–62.

Jankowiak, M. (2021). The EU Member States' national healthcare systems compared using the single synthetic index. *Economics and Business Review*, *7*(2), 37–49.

Jun, B. W., Park, S. J., & Yoon, S. M. (2023). Effects of socio-economic factors on governmental tax expenditures for private pensions in selected OECD countries. *The Singapore Economic Review*, *68*(5), 1599–1611.

Keenay, G., & Whitehouse, E. (2003). The role of the personal tax system in old-age support: A survey of 15 countries. *Fiscal Studies*, *24*(1), 1–21.

Kluzek, M. (2022). Opodatkowanie emerytur w Polsce na tle rozwiązań w państwach Unii Europejskiej. *Studia BAS*, *4*(72), 53–71.

Leventi, C., & Picos, F. (2019). *The tax structure of an economy in crisis: Greece 2009–2017*. European Commission, JRC Working Papers on Taxation and Structural Reforms, 03/2019. https://joint-research-centre.ec.europa.eu/system/files/2020-01/jrc119508.pdf

Lozachmeur, J. M. (2006). Optimal age-specific income taxation, *Journal of Public Economic Theory*, *8*, 697–711.

Manor, M., & Ratajczak, J. (2020). Shift to private pension system: The case of Poland and Israel. *Economics and Business Review*, *6*(1), 82–102.

OECD. (2011). *Pensions at a glance 2011*. https://doi.org/10.1787/pension_glance-2011-en

OECD. (2021). *Pensions at a glance 2021*. https://doi.org/10.1787/ca401ebd-en

OECD. (2023). *Pensions at a glance 2023*. https://doi.org/10.1787/678055dd-en

Patterson, B. (2002). *Tax coordination in the EU: The latest position*. European Parliament Working Paper, Economic Affairs, 128. https://www.europarl.europa.eu/RegData/etudes/etudes/join/2002/312788/DG-4-ECON_ET(2002)312788_EN.pdf

Shi, J., & Kolk, M. (2022). How does mortality contribute to lifetime pension inequality? Evidence from five decades of Swedish taxation data. *Demography*, *59*(5), 1843–1871.

Swift, Z. L. (2006). *Managing the effects of tax expenditures on national budgets*. World Bank Policy Research Papers, WPS3927. https://documents.worldbank.org/curated/en/602361468341100836/Managing-the-effects-of-tax-expenditures-on-national-budgets

Vambol, V., Kowalczyk-Juśko, A., Vambol, S., Khan, N. A., Mazur, M., Goroneskul, M., & Kruzhilko, O. (2023). Multi criteria analysis of municipal solid waste management and resource recovery in Poland compared to other EU countries. *Scientific Reports*, *13*, *22053.*

Verbist, G. (2007). The distribution effects of taxes on pensions and unemployment benefits in the EU-15. micro-simulation in action: Policy analysis in Europe using EUROMOD. *Research in Labour Economics*, *25*, 73–99.

sciendo Economics and Business Review https://www.ebr.edu.pl Vol. 10 (2), 2024: 92–117 https://doi.org/10.18559/ebr.2024.2.1027

Labour productivity in Italian regions: A gravitational model approach

Katarzyna Filipowicz[1] Oleksij Kelebaj[2] Tomasz Tokarski[3]

Abstract

The aim of the paper is to assess the causes of spatial variations in labour productivity of Italian regions using the gravitational model of economic growth. The model is an extension of Robert Solow's economic growth model. The model parameters are calibrated using historical data and numerical simulations of the long-run equilibrium states of the model are carried out. The scenarios considered in the paper vary in forecast investment rates, employment growth rates and urbanisation rates. Based on the results of numerical simulations, it is claimed that to achieve the full convergence in labour productivity, it is necessary to maintain higher investment rates in the south of the country than in Lombardy (by about 4%–11%), and to keep investment rates in central and northern Italy at a similar level as in Lombardy. The fall in investment has affected the poorest regions, Southern Italy, the most, followed by central Italy and the richest regions of the north of the country the least.

Keywords

- model of economic growth
- labour productivity
- Italian regions

JEL codes: R11, O47, C2, E23

Article received 31 October 2023, accepted 18 April 2024.

Suggested citation: Filipowicz, K., Kelebaj, O., & Tokarski, T. (2024). Labour productivity in Italian regions: A gravitational model approach. *Economics and Business Review*, *10*(2), 92–117. https://doi.org/10.18559/ebr.2024.2.1027

[1] Department of Mathematical Economics, Jagiellonian University, 24 Gołębia St, 31-007 Cracow, Poland, k.filipowicz@uj.edu.pl, https://orcid.org/0000-0003-2131-0414.

[2] Faculty of Management and Social Communication, Jagiellonian University, 24 Gołębia St, 31-007 Cracow, Poland, corresponding author: oleksij.kelebaj@doctoral.uj.edu.pl, https://orcid.org/0000-0001-6037-1231.

[3] Department of Mathematical Economics, Jagiellonian University, 24 Gołębia St, 31-007 Cracow, Poland, t.tokarski@uj.edu.pl, https://orcid.org/0000-0002-9551-0892.

Introduction

Italy is both a highly developed country and a very spatially differentiated country in terms of economic development. At the end of the 1970s, analyses of the regional differentiation of the country's economic development began to use the division into so-called first, second and third Italy (Bianchini, 1991). Indeed, each of these areas is characterised by a different economic structure and significant differences in the level of economic development.

The paper's added value is to identify the causes of spatial differentiation of labour productivity in Italian regions using the gravitational model of economic growth. The gravitational growth model (which is a modification and extension of the Solow (1956) neoclassical model of growth) takes into account interactions between the economies being analysed (in this case, Italian regions). In general, it is assumed that an increase in the capita-labour ratio in one economy influences the increase in total factor productivity in the remaining economies.

The first version of gravitational growth model has been proposed in Mroczek et al. (2014), extension in Mroczek et al. (2015). A similar model has been used to analyse the spatial differentiation of economic development in the Polish economy (Filipowicz, 2019; Mroczek et al., 2014; Wisła & Nowosad, 2020), the Ukrainian economy (Wisła & Nowosad, 2020), the economies of EU countries (Wisła et al., 2018), as well as the economies of all European countries (Nowosad & Wisła, 2016) or Balkan countries (Filipowicz et al., 2015).

The structure of the paper is as follows: Section 1 contains a literature review, Second 2 presents the gravitational model of economic growth, and Section 3 describes the calibration of model parameters and numerical simulations. The study ends with the summary of the most important conclusions from the analysis.

1. Literature review

The search for the causes of heterogeneity in Italy's regional development is of great interest to researchers and has been the subject of an ongoing debate for several decades. Possible sources of spatial differentiation of economic development are explained in various ways. Those that stand out here include migration, differences in total factor productivity, economic complexity and diversity at the regional level and related trade links, fluctuations in unemployment, sectoral reallocation of resources, urban development, corruption and education.

Fratesi and Percoco (2014) point out that skill-selective migration can lead to an even greater polarisation of regional development in Italy. They also note that between 1980 and 2001 the migration of people with tertiary education from the southern regions of Italy to the north reduced the level of human capital in the south of the country, which had a negative impact on the economic development of this part of Italy.

The impact of interregional migration in Italy on disparities in total factor productivity between 1995 and 2015 was also analysed in the work by Calcagnini et al. (2021). Their empirical analysis indicated a non-linear (U-shaped) relationship between the employment of temporary workers and the increase in total factor productivity (i.e. the increase in total factor productivity in some southern Italian regions was favoured by an increase in the proportion of temporary workers, while the central and northern regions experienced a decrease). In addition, migration flows of skilled personnel had a positive impact on the increase in total factor productivity in the regions of migration destination.

The importance of the impact of the level of regional total factor productivity on differences in labour productivity was highlighted by Maffezzoli (2004), who studied convergence in Italian regions in relation to technological convergence. He pointed out that differences in relative total factor productivity between Italian regions were important and were the main source of convergence between 1980 and 2000.

Di Giacinto and Nuzzo (2004), on the other hand, attribute differences in labour productivity at the regional level in Italy to three factors. These factors are as follows: the structure of the regional economy in the south of Italy (a significant part of the labour force is employed in less productive sectors of the economy), the accumulated stock of physical and human capital in the region, and differences in the level of total factor productivity (which, broadly speaking, is in line with the economic growth model in Mankiw et al. (1992). The study also assessed the role of the aforementioned determinants of total factor productivity and empirically evaluated the factors influencing this variable. The factors influencing the variation in total factor productivity were: public and social capital, R&D investment, international openness, development of financial markets, development of agglomeration and diversification of economies, and geographical factors.

Changes in: labour productivity, technology efficiency and physical and human capital that occurred in Italian regions between 1980 and 2006 were also analysed by Gitto and Mancuso (2015). Their results show that the importance of labour productivity, technology efficiency and capital accumulation in terms of economic growth differs significantly between the southern regions of Italy and the rest of the country.

Basile and Cicerone (2022) studied the role of economic complexity as a driver of regional variation in labour productivity in Italy. Here, the economic

complexity was measured by their Economic Complexity Index (ECI). This index specifically seeks to explain the accumulated knowledge in the population, which is expressed in economic activity in a city, a country or a region. Basile and Cicerone argue that economic complexity plays a key role in the observed trend towards polarisation of labour productivity in Italian regions.

Regional diversity and trade links may be further determinants of differences in the economic growth of Italian regions. In the articles by Boschma and Iammarino (2007, 2009), the empirical part is based on export and import data for the period 1995–2003. Their results show that Italian regions with complementary sectors in terms of competences grow better economically. The study also assesses the impact of relatedness of international trade links on economic growth at the regional level. The authors conclude that related extra-regional knowledge stimulates cross-sectoral learning in regions and becomes a catalyst for regional economic development.

On the other hand, Busetta and Corso (2012) analyse Italian regions in terms of the impact of unemployment fluctuations on economic growth. The authors base their study on Okun's law, which can also be observed in Italy at the regional level. The issue of unemployment in the context of variations in Italy's level of regional development was also addressed by Carmeci and Mauro (2002). More specifically, these researchers analysed the relationship between stopping the convergence process of Italian regions in the early 1970s and the increase in regional differences in unemployment levels. In their analysis, they use a neoclassical growth model with an imperfect labour market. On this basis, they argue that labour market imperfections have a negative effect on the growth rate of output. In addition, they conclude that setting the national minimum wage too high (in relation to labour productivity) can negatively affect economic growth mainly in the less developed regions of Italy, as there a high minimum wage has a much stronger, negative impact on labour demand than in regions with high labour productivity.

The analysis of the factors differentiating Italy's regional economic development also raises the issue of sectoral reallocation of resources. Paci and Pigliaru (1997, 1998) conclude that aggregate convergence is to a large extent the result of structural change and that the shift from agriculture to non-agricultural production is particularly important for aggregate convergence. In doing so, it is important that the outflow of labour from low-productivity agriculture (in poorer regions) is a source of expansion of non-agricultural sectors.

Urban development, corruption and education are further determinants of regional labour productivity differences in Italy. Di Liddo (2015) draws attention to urban sprawl and its impact on economic development at the regional level. He also points out the negative impact of urban sprawl in Italian cities and recommends stimulating urban development in the main cities rather than in the provinces. On the other hand, Fiorino et al. (2012) analysed the impact of corruption on Italian economic growth at the regional level. They

found a negative correlation between corruption and economic growth. They argue, in addition, that in Italy corruption undermines the positive impact of public spending in mitigating regional economic growth differences.

Education is another determinant of differences in labour productivity at the regional level. Research by Di Liberto (2008) shows that an increase in educational attainment appears to have a statistically significant effect on labour productivity growth in the southern regions of Italy.

The problem of regional convergence in Italy is also addressed in the study by Terrasi (1999), which analyses the regional convergence of GDP per capita between 1953 and 1993. This research shows that 1960–1975 was a period of strong regional convergence in Italy, while after 1975 there was a tendency towards regional divergence. These latter processes are caused by both national development and spatial factors (Paci & Pigliaru, 1997, 1998; Terrasi, 1999). Moreover, a study of per capita income convergence in Italy over the period 1951–2000 confirmed the occurrence of convergence clubs (i.e. Italian regions with similar structural characteristics become more similar to each other, bridging the gap in per capita income (Arbia & Basile, 2005).

2. Gravitational model of economic growth

In analysing the determinants of spatial variation in labour productivity in Italian regions, we make the following assumptions:

I. The level of labour productivity in region is described by the labour productivity function (derived from the Cobb-Douglas production function) given by the formula:[4]

$$\forall i \; y_i(t) = a_i e^{\gamma t} g_i^{\beta}(t) k_i^{\alpha}(t) \tag{1}$$

where y_i is labour productivity (in region i), a_i – total factor productivity[5], g_i – total gravitational effects, k_i – capital per worker, $\alpha, \beta \in (0, 1)$ – elas-

[4] We assume that all macroeconomic variables analysed in this section are continuous and differentiable functions of time $t \in [0, +\infty)$. Moreover $\dot{x}_i(t) = \frac{dx_i}{dt}$, $\forall_i$ means $\forall_i = 1, 2, ..., N$. Records $\Pi_i x_i$ and $\Sigma_i x_i$. Records and are defined by formulas: $\Pi_i x_i = \prod_{i=1}^{n} x_i$ and $\Sigma_i x_i = \sum_{i=1}^{n} x_i$.

[5] Precisely speaking, a_i is a part of the total factor productivity $TFP = \frac{y}{k^{\alpha}} = \left(\frac{Y}{K}\right)^{\alpha} \cdot \left(\frac{Y}{L}\right)^{1-\alpha}$, which does not result from gravitational effects and is not the result of technical progress in the Hicks sense (for more on this, see Allen, 1975; or Dykas et al., 2023). Moreover, in the further empirical analyses, we assume that a_i is at a higher level the higher the urbanization rate urb_i in region i, which is described by the function $a_i = be^{\psi \cdot urb_i}$, where $b, \psi > 0$.

ticities of labour productivity with respect to capital per worker and to gravitational effects, $\gamma > 0$ – rate of technical progress in the Hicks sense. In addition, we assume (to obtain the asymptotic stability of the steady state of the model under consideration) that: $\beta < \frac{1-\alpha}{2}$.

II. The total gravitational effects g_i affecting region i are the geometric mean of the gravitational effects connecting region i to the other regions, i.e. g_{ij}. We therefore have:

$$\forall (i, j \wedge j \neq i)\, g_i(t) = \sqrt[N-1]{\prod_{j \neq i} g_{ij}(t)} = \prod_{j \neq i} g_{ij}^{1/(N-1)}(t) \tag{2}$$

III. Individual gravitational effects are given by the formula:

$$\vee (i, j \wedge i \neq i)\, g_{ij}(t) = \frac{k_i(t)\, k_j(t)}{d_{ij}^2} \tag{3}$$

where $d_{ij} > 0$ denotes the distance from the capital of region i to the capital of region j.

IV. As in the Solow model, the capital accumulation process is described by the differential equation:

$$\forall i\, \dot{k}_i(t) = s_i y_i(t) - \mu_i k_i(t) \tag{4}$$

where $s_i \in (0, 1)$ is an investment rate (in region i), μ_i – capital depreciation rate per working person, i.e. the sum of capital depreciation rates $\delta \in (0, 1)$ and growth rate of the number of workers $n_i > 0$.

Equations (2–3) give:

$$\forall i\, g_i(t) = k_i(t) \prod_{j \neq i} \frac{k_j^{1/(N-1)}(t)}{d_{ij}^{2/(N-1)}} = \frac{k_i(t)}{\overline{d}_i^2} \prod_{j \neq i} k_j^{\frac{1}{(N-1)}}(t) \tag{5}$$

where $\overline{d}_i = \prod_{j \neq i} d_{ij}^{1/(N-1)}$ is the geometric mean distance of the capital of the i-th region to the capitals of the other regions. Equations (1) and (5) give us:

$$\forall i\, y_i(t) = a_i e^{\gamma t} \left(\frac{k_i(t)}{\overline{d}_i^2} \prod_{j \neq i} k_j^{\frac{1}{(N-1)}}(t) \right)^{\beta} \cdot k_i^{\alpha}(t) = \frac{a_i e^{\gamma t}}{\overline{d}_i^{2\beta}} \left(\prod_{j \neq i} k_j^{\frac{\beta}{(N-1)}}(t) \right) k_i^{\alpha+\beta}(t) \tag{6}$$

We make substitutions:

$$\forall i\, y_{Ei}(t) = \exp\left(-\frac{\gamma}{1-2\beta-\alpha} t \right) y_i(t) \Rightarrow y_i(t) = \exp\left(\frac{\gamma}{1-2\beta-\alpha} t \right) y_{Ei}(t) \tag{7}$$

and:

$$\forall i\, k_{Ei}(t) = \exp\left(-\frac{\gamma}{1-2\beta-\alpha}t\right)k_i(t) \Rightarrow k_i(t) = \exp\left(\frac{\gamma}{1-2\beta-\alpha}t\right)k_{Ei}(t) \quad (8)$$

From equation (8) it follows that:

$$\forall i \frac{\dot{k}_{Ei}(t)}{k_{Ei}(t)} = -\frac{\gamma}{1-2\beta-\alpha} + \frac{\dot{k}_i(t)}{k_i(t)}$$

and hence and from the capital accumulation equation (4) we get:

$$\forall i \frac{\dot{k}_{Ei}(t)}{k_{Ei}(t)} = s_i \frac{y_i(t)}{k_i(t)} - \left(\mu_i + \frac{\gamma}{1-2\beta-\alpha}\right) = s_i \frac{y_{Ei}(t)}{k_{Ei}(t)} - \left(\mu_i + \frac{\gamma}{1-2\beta-\alpha}\right)$$

and finally:

$$\forall i \dot{k}_{Ei}(t) = s_i y_{Ei}(t) - \eta_i k_{Ei}(t) \quad (9)$$

where $\forall i \eta_i = \mu_i + \frac{\gamma}{1-2\beta-\alpha} > 0$. We include the substitutions (7–8) in the labour productivity function (6). Therefore, we obtain:

$$\forall i\, y_{Ei}(t) = \frac{a_i}{\overline{d}_i^{2\beta}}\left(\prod\nolimits_{j\neq i} k_{Ej}^{\beta/(N-1)}(t)\right)k_{Ei}^{\alpha+\beta}(t) \quad (10)$$

Relationship (10) is inserted into relationship (9) to obtain the following system of differential equations:

$$\forall i \dot{k}_{Ei}(t) = \frac{a_i s_i}{\overline{d}_i^{2\beta}}\left(\prod\nolimits_{j\neq i} k_{Ej}^{\beta/(N-1)}(t)\right)k_{Ei}^{\alpha+\beta}(t) - \eta_i k_{Ei}(t) \quad (11)$$

The system of equations (11) is a simple generalisation of the system of differential equations from the gravitational model of economic growth. Thus, using the Grobman-Hartman theorem (Ombach, 1999) it can be shown (Mroczek et al., 2014) or (Dykas et al., 2023) that at $\beta < \frac{1-\alpha}{2}$ the system has an asymptotically stable steady state $(k_{E1}^{\star}, \ldots, k_{EN}^{\star}) \in (0, +\infty)^N$, where:

$$\forall i \frac{k_i^{\star}}{k_B^{\star}} = \frac{k_{Ei}^{\star}}{k_{EB}^{\star}} = \left(\frac{a_i s_i \eta_B \overline{d}_B^{2\beta}}{a_B s_B \eta_i \overline{d}_i^{\beta}}\right)^{\frac{1}{1-\alpha-\frac{N-2}{N-1}\beta}} \quad \wedge \quad \frac{y_i^{\star}}{y_B^{\star}} = \frac{y_{Ei}^{\star}}{y_{EB}^{\star}} = \frac{\left(\frac{a_i s_i \eta_B}{a_B s_B \eta_i}\right)^{\frac{\alpha+\frac{N-2}{N-1}\beta}{1-\alpha-\frac{N-2}{N-1}\beta}}}{\left(\frac{\overline{d}_i^2}{\overline{d}_B^2}\right)^{\frac{\beta}{1-\alpha-\frac{N-2}{N-1}\beta}}} \quad (12)$$

where *B* subscripts refer to a base region[6] (the base region in the empirical analyses carried out hereafter is Lombardy, as the region with the highest economic potential in Italy). The quotients $\frac{y_i^*}{y_B^*}$ described by compounds (12), after taking into account the substitutions $\forall i\, \eta_i = \mu_i + \frac{\gamma}{1-2\beta-\alpha} > 0 \wedge \mu_i = \delta + n_i > 0$, reduce to the relation:

$$\forall i \frac{y_i^*}{y_B^*} = \frac{\left(\frac{a_i s_i \left(\delta + \frac{\gamma}{1-2\beta-\alpha} + n_B \right)}{a_B s_B \left(\delta + \frac{\gamma}{1-2\beta-\alpha} + n_i \right)} \right)^{\frac{\alpha + \frac{N-2}{N-1}\beta}{1-\alpha-\frac{N-2}{N-1}\beta}}}{\left(\frac{\overline{d}_i^2}{\overline{d}_B^2} \right)^{\frac{\beta}{1-\alpha-\frac{N-2}{N-1}\beta}}} \tag{13}$$

Equation (13) describes the relationship between the level of labour productivity in region and the value of this macroeconomic variable in the base region *B* in the long-run equilibrium of the gravitational growth model. From equation (13) we can also determine the combination of investment rates $(s_1, ..., s_N) \in (0, 1)^N$ at which full labour productivity convergence will occur in the long run (and thus $\forall i \frac{y_i^*}{y_B^*} = 1$). Thus, for all $i\, y_i^* = y_B^*$ if and only if:

$$\forall i \frac{s_i}{s_B} = \left(\frac{\overline{d}_i^2}{\overline{d}_B^2} \right)^{\frac{\beta}{\alpha + \frac{N-2}{N-1}\beta}} \cdot \frac{a_B \left(\delta + \frac{\gamma}{1-2\beta-\alpha} + n_i \right)}{a_i \left(\delta + \frac{\gamma}{1-2\beta-\alpha} + n_B \right)} \tag{14}$$

3. Calibration of model parameters and numerical simulations

For numerical simulations of the long-run equilibrium states of the gravitational economic growth model (equation (13), it is necessary to estimate

[6] Throughout the study, we refer regional macroeconomic variables to the base region, Lombardy. The choice of Lombardy is due to the fact that it is the region with the greatest economic potential in Italy (despite the fact that the highest GDP per capita is recorded in the Autonomous Province of Bolzano, which, however, is characterized by low demographic potential, so it does not play such an important role in the economic development of Italy as Lombardy).

the parameters of the labour productivity function (1). The parameters of this function were estimated in two ways. Firstly, they were estimated without taking external gravitational effects into account (i.e. without the impact of foreign countries on Italian regions, as is the case of the original gravitational model), and secondly, with these effects taken into account.

The labour productivity function without external gravitational effects (taking its logarithm) is given by the formula:

$$\ln y_{it} = a + b \cdot covid + \gamma t + \psi \cdot urb_{it} + \beta \ln g_{it} + \alpha \ln k_{it} \tag{15}$$

where y_{it} is labour productivity (in region i in year t), urb_{it} – percentage of people living in cities with more than 100 000 inhabitants,[7] g_{it} – total gravitational effects, k_{it} – capital per worker, t – time trend equal to 2010, 2011, …, *covid* – dummy variable equal to 1 in 2020, 0 in others, a, b, γ, ψ, β, α – parameters of the estimated equation (where γ is the rate of technical progress in the Hicks sense). Equation (16) shows that the total factor productivity in region i in year t, which is unrelated to the effect of gravity (denoted as TFP_{it}), satisfies the relationship:

$$\ln TFP_{it} = a + b \cdot covid + \gamma t + \psi \cdot urb_{it} \tag{16}$$

According to (16), the dummy variable $covid$ modifies TFP_{it} for the peak period of the pandemic. This is because, at that time, the volume of output (with given factor inputs) was falling for two reasons. Firstly, it was smaller because the volume of aggregate demand in the economy was decreasing due to Covid restrictions (catering, tourism, passenger transport, etc.). Secondly, TFP_{it} was also declining because part of the labour pool periodically did not take up work (illness, quarantine, etc.) (cf. Bärwolff, 2020; Dykas et al., 2023; Dykas & Wisła, 2022; Gori et al., 2022). In turn, the link between TFP_{it} and the previously defined urbanisation rate can be justified by the fact that it is easier to do business in large cities and (usually) the level of human capital is higher there (the best tertiary schools in Italy are traditionally located in a few of the largest cities, and graduates of these schools coming from the provinces often stay in these cities, thus increasing the human capital stock there at the expense of the provinces).

The parameters of equation (15) were estimated using the ordinary least squares method (OLS) and the generalised method of moments (GMM). Fixed

[7] Since the authors did not have access to data on the total number of city inhabitants or the rate of urbanization, they replaced this rate with the index: $urb_{it} = \frac{C_{it}}{POP_{it}}$, where c_{it} denotes the number of inhabitants in cities with more than 100,000 inhabitants in the region i and in the year t, POP_{it} – the total population in region i in year t. This indicator will be referred to (not very precisely) as the urbanisation rate.

effects were not included in these estimates due to the fact that the variables urb_{it} and g_{it} are strongly differentiated in geographical space and little differentiated in time (parameter estimates of equations in which the dependent variable was urb_{it} or $\ln g_{it}$, and the independent variables were a matrix of fixed effects variables yielded adjusted coefficient of determination of 0.968 and 0.867, respectively). Therefore, these variables are strongly collinear with the fixed effects.

Table 1. Estimated parameters of equation (15)

Independent variables	Method of estimation			
	OLS		GMM	
Constant	–27.972 (–7.117)	–33.300 (–7.631)	–35.401 (–7.926)	–42.700 (–8.580)
Covid	–	–0.0568 (–2.664)	–	–0.0722 (–3.413)
t	0.0146 (7.623)	0.0173 (8.079)	0.0183 (8.356)	0.0219 (8.968)
urb	0.371 (8.345)	0.370 (8.446)	0.369 (8.041)	0.369 (8.244)
$\ln g$	0.0378 (2.977)	0.0381 (3.041)	0.0394 (3.009)	0.0382 (2.994)
$\ln k$	0.572 (16.083)	0.574 (16.345)	0.601 (16.230)	0.593 (16.467)
R^2 Adjusted R^2	0.659 0.653	0.669 0.662	0.671 0.665	0.689 0.682
Sample	2010–2020		2011–2020	
Number of observations	231		210	

Note: t-Student statistic is given in brackets under the parameter estimates, R^2 is the coefficient of determination, adjusted R^2 is the adjusted coefficient of determination, in the GMM estimates, the instruments are lagged by one year for the dependent variable and the independent variables.

Source: own calculations.

Estimates of the parameters of equation (15) are presented in Table 1. Thus, all independent variables had a statistically significant effect on the regional variation of labour productivity in Italy. From the estimates presented, the rate of Hicksian technical progress was around 1.5%–2.2% between 2010 and 2020 or 2011–2020. In 2020, the COVID-19 pandemic led to an average decrease in production volumes in the Italian regions of 5.7%–7.2%. A 1-percentage-point increase in the urbanisation rate translated into a 0.37% in-

crease in labour productivity. Elasticity of labour productivity with respect to total gravitational effects is smaller then 0.04, elasticity of $\frac{Y}{L}$ with respect to $\frac{K}{L}$ is equal to 0.57–0.60. The OLS and GMM estimations of parameters of equation (15) are close to each other. It can therefore be hypothesised that these equations provide a good description of the regional variation of labour productivity in Italy.

In addition, equation (15) was extended to include the impact on the subsequent Italian regions of gravitational effects flowing from abroad (so-called external gravitational effects). The parameters of the relationship were then estimated:

$$\ln y_{it} = a + b \cdot covid - \alpha_G \cdot Genewa_i + \gamma t + \psi \cdot urb_{it} + \beta \ln g_{it} + \alpha \ln k_{it} \tag{17}$$

or:

$$\ln y_{it} = a + b \cdot covid + \gamma t + \psi \cdot urb_{it} + \beta \ln g_{it} + (\alpha - \alpha_G Genewa_i) \ln k_{it} \tag{18}$$

where $Genewa_i$ is (expressed in mingeo) distance of the capital of the i-th region from Geneva. The choice of Geneva as the main foreign centre of gravity affecting the Italian regions is due to the fact that Geneva and Zurich are among the most important financial centres in Europe. Furthermore, Geneva is a highly internationalised city, home to the largest number of international organisations, i.e. the UN (European Headquarters), the WHO, the ILO or the WTO. It should be stressed that Italy's level of economic development (particularly in the northern regions) is influenced not only by cooperation with Switzerland, but also (for historical reasons) with France, Austria or (to a lesser extent) Germany. Geneva is a reflection of the external gravitational effects arising from the cooperation of Italian regions with the economically highly developed countries of Western Europe.

Equation (17) shows that each successive geographical minute of distance of the capital of the i-th region from Geneva translated into a decrease in labour productivity of α_G%. In equation (18), on the other hand, distance from Geneva interactively modifies the elasticity α of labour productivity with respect to capital per worker. Specifically, it follows from this equation that if the distance of the capital of the i-th region from Geneva were G_i the elasticity of labour productivity with respect to capital per worker would be equal to $\alpha - G_i \alpha_G$.

The estimated OLS and GMM parameters of equations (17–18) can be found in Table 2. The following conclusions can be drawn from the estimates presented that as with the parameter estimates of the relationship (15), the parameter estimates of equations (17–18) also proved to be statistically significant. The estimated rate of Hicksian technical progress was around 1.3% (with OLS) or 1.7% (with GMM). So the estimates were slightly lower than the

estimates without foreign gravitational effects. The parameter estimates of equations (17–18) show that the COVID-19 pandemic (ceteris paribus) translated into an average decrease in labour productivity of about 5.4% (with OLS) or about 6.8% (with GMM). This means that the parameter estimates with the covid dummy variable after accounting for external gravitational effects were lower than the estimates with this variable without these effects. The same is true of the parameter estimates for the urbanisation rate. Estimates of the parameters of equation (15) indicated that each additional percentage point in the urbanisation rate translated into an increase in labour productivity of about 3.7%, while estimates of the parameters of equations (17–18) concluded that this increase amounted to about 3%.

Table 2. Parameter estimates for equations (17–18)

Independent variable	Method of estimation			
	OLS		GMM	
Constant	–23.026 (–5.219)	–22.811 (–5.214)	–31.980 (–6.286)	–31.694 (–6.284)
covid	–0.0541 (–2.725)	–0.0542 (–2.744)	–0.0677 (–3.415)	–0.0677 (–3.433)
t	0.0126 (5.886)	0.0125 (5.863)	0.0170 (6.865)	0.0168 (6.848)
Genewa	–0.000220 (–5.969)	–	–0.000202 (–5.258)	–
urb	0.297 (6.977)	0.297 (7.024)	0.303 (6.902)	0.302 (6.942)
$\ln g$	0.0274 (2.322)	0.0288 (2.468)	0.0283 (2.337)	0.0296 (2.467)
Geneva $\ln g$	–	$-4.67 \cdot 10^{-5}$ (–6.226)	–	$-4.31 \cdot 10^{-5}$ (–5.523)
$\ln k$	0.421 (10.116)	0.434 (10.962)	0.448 (10.307)	0.460 (11.130)
R^2 Adjusted R^2	0.715 0.707	0.718 0.711	0.729 0.721	0.732 0.724
Sample	2010–2020		2011–2020	
Number of observations	231		210	

Note: *t*-Student statistic is given in brackets under the parameter estimates, R^2 is the coefficient of determination, adjusted R^2 is the adjusted coefficient of determination, in the GMM estimates, the instruments are lagged by one year for the dependent variable and the independent variables.

Source: own calculations.

Elasticity of labour productivity with respect to gravitational effects is equal to about 0.03. The parameter estimates of equation (17) also show that the elasticity of labour productivity with respect to capital per worker was around 0.421–0.460. Analysing the parameter estimates of the expression $\alpha - \alpha_G \, Genewa_i$, we conclude that each successive geographic minute of distance of the capital of the i-th region from Geneva reduced the elasticity of labour productivity with respect to capital per worker by $4.31 \cdot 10^{-5} - 4.67 \cdot 10^{-5}$. Thus, e.g., the elasticity of y with respect to k in Lombardy was about 0.425–0.452, in Lazio 0.413–0.440, in Sicily 0.404–0.430, in Sardinia 0.413–0.440.

The resulting GMM estimates of parameters of equation (17) allowed labour productivity relations to be determined in the long-run equilibrium of the gravitational growth model, according to relations (13). Note, however, that (according to relations (17) the expressions a_i in equation (13) were replaced by:

$$a_i = \exp(a - \alpha_G \cdot Genewa_i + \psi \cdot urb_i^{\star})$$

where $urb_i^{\star}$ denotes the assumed long-run urbanisation rate. Since the quotients $\frac{y_i^{\star}}{y_B^{\star}}$ are higher, the higher the relations $\frac{a_i^{\star}}{a_B^{\star}}$ are, these expressions are equal to:

$$\exp\left(\alpha_G \cdot (Genewa_B - Genewa_i)\right)$$

Thus, the long-run labour productivity relationship $\frac{y_i^{\star}}{y_B^{\star}}$ presented hereafter is (*ceteris paribus*) the lower, the further the capital of the i-th region is located from Geneva.

Numerical simulations of the long-run labour productivity relationship in Italian regions in relation to the value of this variable in the base region (Lombardy) were carried out in two variants: a so-called baseline variant and a variant with development based on the six largest cities. In each of these variants, several scenarios are distinguished.

The baseline variant was simulated with the following several assumptions (scenarios):

1. Investment rates (s_i), employment growth rates (n_i) and urbanisation rates (urb_i) will evolve, as they did on average between 2000 and 2019 (these rates will be referred to hereafter as historical investment rates, historical employment growth rates and historical urbanisation rates).
2. n_i and urb_i rates: as they were on average in 2000–2019, while s_i – at a certain, same level in all regions $\overline{s} \in (0, 1)$.
3. s_i, n_i as they were on average in 2000–2019, $urb_i = \overline{urb} \in (0, 1)$.

4. Urbanisation rates: as before, while s_i and n_i rates – at levels $\overline{s} \in (0, 1)$ and $\overline{n} > 0$.
5. The growth rate of the number of employees: as on average in the period 2000–2019. While the investment rates at $\overline{s} \in (0, 1)$, the urbanisation rates: $urb \in (0, 1)$.
6. Investment rates: as before n_i and $\overline{n} > 0$ equal to and $urb_i = \overline{urb} \in (0, 1)$.
7. Investment rates equal to $\overline{s} \in (0, 1)$, growth rates of employment: $\overline{n} > 0$ and urbanisation rates: $urb \in (0, 1)$.

The simulation results of the baseline variant are summarised in Table 3. From the estimates of presented there, we notice that in Scenario 1 (i.e. with historical investment rates, labour growth rates and urbanisation rates), in the long run, the regions of northern Italy should have labour productivity 8.5% higher than in Lombardy, Central Italy 3.5% lower than the base region, while the south of the country should be 4% lower (for the Italian economy as a whole, output per worker should be 2.4% higher than in Lombardy).[8] In this scenario, the highest relative labour productivity $\frac{y_i^*}{y_B^*}$ will be recorded in the Aosta Valley (1.228), Piedmont (1.217), Molise (1.185), Emilia-Romagna (1.178) and the Autonomous Province of Bolzano and Umbria (1.119 each), and the lowest in Puglia (0.862), Marche (0.904), Sicily (0.925) and Lazio (0.940).

If all Italian regions had the same investment rates (Scenario 2), labour productivity would be 2.3% higher in the northern Italian regions than in Lombardy, 2.1% higher in Central Italy and 3.9% lower in the south than in the base region. Moreover, with the same investment rate, the highest relative labour productivity would be in Liguria (1.118), Piedmont (1.074), Emilia-Romagna (1.051), Umbria (1.044) and Lazio (1.002), while the lowest would be in Basilicata (0.900), Calabria (0.911) and Sardinia (0.920).

In Scenario 3 (and therefore with the same labour growth rates and historically shaped investment and urbanisation rates), labour productivity in northern Italy would be 6.6% higher than in Lombardy, 3.2% lower in Central Italy, while in the south the value of this variable would be just over 90% of labour productivity in the base region. For the Italian economy as a whole, labour productivity would be 0.1% lower than in Lombardy. The regions with the highest value of this variable should then be Emilia-Romagna (1.177), the Autonomous Province of Bolzano (1.154), the Aosta Valley (1.142), the Autonomous Province of Trento (1.106), Umbria (1.099), Molise (1.060) and Veneto (1.055), while the lowest relative labour productivity should re-

[8] The fact that labour productivity in Lombardy will be at a relatively low level (especially in relation to the regions of northern Italy) in the variants assuming historical investment rates is due to the fact that these rates in Lombardy were low in the period 2000–2019 (19.0%), while, for example, in the Aosta Valley they were 23.8%, in Molise and the Autonomous Province of Bolzano they were 23.1%, and in the Italian economy as a whole they were 19.5%.

Table 3. Labour productivity in long-run equilibrium—baseline variant (Lombardy = 100)

Region or group of regions	Year 2020	Scenario							
		1	2	3	4	5	6	7	8
Piedmont	85.8	121.7	107.4	114.2	120.2	100.8	106.0	112.8	99.5
Aosta Valley	100.6	122.8	101.1	114.5	129.1	94.3	106.2	120.4	99.1
Liguria	93.2	107.1	111.8	100.9	101.5	105.4	105.9	95.6	99.8
Autonomous Province of Bolzano	113.1	111.9	93.8	115.4	113.6	96.7	95.2	117.2	98.2
Autonomous Province of Trento	101.9	107.9	97.8	110.6	106.8	100.3	96.8	109.5	99.3
Veneto	86.8	107.6	100.2	105.5	107.7	98.3	100.4	105.7	98.4
Friuli-Venezia Giulia	86.8	106.3	97.6	103.1	106.6	94.7	98.0	103.4	95.0
Emilia-Romagna	91.1	117.8	105.1	117.7	112.4	105.1	100.3	112.4	100.3
Northern Italy	**93.5**	**108.5**	**102.3**	**106.6**	**107.2**	**100.6**	**101.1**	**105.4**	**99.4**
Tuscany	85.9	99.6	101.9	98.2	99.3	100.4	101.5	97.8	100.0
Umbria	73.3	111.9	104.4	109.9	107.9	102.5	100.7	106.0	98.9
Marche	76.9	90.4	96.6	88.0	93.5	94.0	99.9	91.0	97.2
Lazio	96.3	94.0	103.6	96.5	85.8	106.4	94.6	88.1	97.2
Central Italy	**88.7**	**96.5**	**102.1**	**96.8**	**92.6**	**102.4**	**98.0**	**92.8**	**98.2**
Abruzzo	76.4	100.9	97.0	98.4	103.4	94.7	99.5	100.9	97.1
Molise	70.2	118.5	100.0	106.0	124.5	89.5	105.2	111.4	94.1
Campania	77.1	100.4	98.0	95.0	99.6	94.5	99.0	94.3	93.8
Apulia	70.6	86.2	93.8	81.0	86.6	88.2	94.3	81.4	88.6
Basilicata	75.3	103.7	90.0	99.5	109.1	86.3	94.6	104.6	90.7
Calabria	71.1	96.5	91.1	89.1	98.9	84.2	93.4	91.4	86.3
Sicily	74.5	92.5	97.9	85.7	90.4	90.7	95.6	83.8	88.6
Sardinia	69.6	104.7	92.0	103.3	105.0	90.8	92.3	103.6	91.0
Southern Italy	**73.7**	**96.0**	**96.1**	**90.7**	**96.1**	**90.7**	**96.1**	**90.7**	**90.8**
ITALY	**87.2**	**102.4**	**100.5**	**99.9**	**100.9**	**98.1**	**99.1**	**98.5**	**96.7**

Source: own calculations based on equation (13).

cord: Puglia (0.810), Sicily (0.857), Marche (0.88) and Calabria (0.891). In the case where all regions were characterised by the same urbanisation rates (Scenario 4) then labour productivity in northern Italy would be 7.2% higher than in Lombardy, in Central Italy by 2.4%, while in the south it would be

3.9% lower than in the base region (in the whole economy relative labour productivity should be 1.009). The Aosta Valley (1.291), Molise (1.245) and Piedmont (1.202) would then have the highest value for this characteristic, while Lazio (0.858), Puglia (0.866), Sicily (0.904) and Marche (0.935) would have the lowest.

If, in all regions, investment rates and growth rates in the number of workers had been at the same level while urbanisation rates had been the same as in the first two decades of the 21st century (Scenario 5), then in the regions of northern and Central Italy labour productivity would have been similar to that recorded in Lombardy, while in the south the value of this variable would have been almost 10% lower than in the base region. Then, in the Autonomous Province of Trento, Tuscany, Piedmont, Umbria, Emilia-Romagna, Liguria and Lazio, the value of this variable would be a few percent higher than in Lombardy, while in Calabria, Basilicata, Apulia and Molise, it would be more than 10% lower. In Scenario 6 (and therefore with the same s_i and urb_i rates and historically shaped n_i rates), labour productivity in northern Italy would be 1.1% higher than in Lombardy, while in the central and southern Italian regions the value of this macroeconomic variable would be 2% and 3.9% lower, respectively. The highest values of this characteristic would be recorded in the Aosta Valley (1.062), Piedmont (1.06), Liguria (1.059), Molise (1.052), and the lowest in Sardinia (0.923), Calabria (0.934), Apulia (0.943), Basilicata (0.946) and Lazio (0.946).

If, on the other hand, investment rates had evolved as they did in the first two decades of the 21st century, and employment growth and urbanisation rates had been the same in all regions (scenario 7), labour productivity in the north of Italy would have been 5.4% higher than in Lombardy, 7.2% lower in Central Italy and 9.3% lower in the south of Italy than in the base region. For the Italian economy as a whole, relative labour productivity would be 0.985. The highest values for this characteristic should then be found in the Aosta Valley (1.204), the Autonomous Province of Bolzano (1.172) and the regions of Piedmont (1.128), Emilia-Romagna (1.124) and Molise (1.114). In the last scenario (and therefore with the s_i, n_i and urb_i rates the same for all regions), labour productivity levels in the north and centre of Italy would be 1%–2% lower than in Lombardy, while in the south they would be more than 9% lower than the base region (Italy's relative labour productivity would then be 0.967). In this scenario, labour productivity levels in Tuscany would be similar to Lombardy, in Emilia-Romagna they would be 0.3% higher than in the base region. In Calabria, Puglia and Sicily, on the other hand, the value of this variable should be more than 10% lower than in Lombardy. It is worth noting that in this scenario, the variation in labour productivity is the result of the geographical location of the regions alone.

In the second variant considered, based on the development of a few large cities, it was arbitrarily assumed that in the long run (i.e. at $t \to \infty$) the pop-

ulation of the six largest Italian cities (Rome, Milan, Naples, Turin, Palermo and Genoa) would increase by 50%, that of the remaining cities of 100,000 inhabitants by 25%, while the urb_i rate in regions without cities of 100,000 inhabitants would increase to 5%. Then (under the additional assumption that

Table 4. Urbanisation rates in the Italian regions in 2020 and at $t \to \infty$ (in %)

Region or group of regions	Year		Change in $t + \infty$ in relation to 2020 (percentage points)
	2020	$t \to \infty$	
Piedmont	22.3	27.9	5.6
Aosta Valley	0.0	5.0	5.0
Liguria	37.1	46.4	9.3
Lombardy	18.3	22.8	4.6
Autonomous Province of Bolzano	20.2	25.3	5.1
Autonomous Province of Trento	22.0	27.4	5.5
Veneto	17.1	21.4	4.3
Friuli-Venezia Giulia	16.7	20.8	4.2
Emilia-Romagna	36.1	47.3	11.2
Northern Italy	**22.6**	**28.6**	**6.0**
Tuscany	19.5	24.4	4.9
Umbria	31.5	39.3	7.9
Marche	0.0	5.0	5.0
Lazio	50.7	63.3	12.7
Central Italy	**33.1**	**42.0**	**8.9**
Abruzzo	9.2	13.8	4.6
Molise	0.0	5.0	5.0
Campania	20.8	31.2	10.4
Apulia	16.6	20.7	4.1
Basilicata	0.0	5.0	5.0
Calabria	9.2	13.8	4.6
Sicily	26.3	32.9	6.6
Sardinia	17.0	21.3	4.3
Southern Italy	**18.3**	**24.9**	**6.6**
ITALY	**23.2**	**30.0**	**6.8**

Source: own calculations based on www.istat.it.

the population of the regions does not change[9]) these rates should evolve as in Table 4. Based on the urbanization rates summarised in this Table, the observation can be made (in 2020 and at $t \to \infty$) that in 2020, in 4 regions the percentage of people living in cities with a population over 100,000 exceeded 30% (Lazio 50.7%, Liguria 37.1%, Emilia-Romagna 36.1% and Umbria 31.5%. On the other hand, in four regions (Aosta Valley, Marche, Molise and Basilicata) there were no cities with a population over 100,000. Based on the assumptions made here about long-term urbanization rates, six provinces (as in 2020 Lazio 63.3%, Emilia-Romagna 47.3%, Liguria 46.6%, Umbria 39.3%, and, additionally, Sicily 32.9% and Campania 31.2%) should be characterized by an index value of more than 30% of urb_i.

Comparing urbanisation rates at $t \to \infty$ with those recorded in 2020, it emerges that based on the assumptions made here, the highest (over 10 percentage points) increases in the value of this indicator should occur in Lazio (by 12.7 percentage points), Emilia-Romagna (11.2 points) and Campania (10.4 points). The lowest (less than 5 points) are to be found in Apulia (4.1 points), Friuli-Venezia Giulia (4.2 points), Sardinia (4.3 points), Veneto (4.3 points), Abruzzo, Calabria and Lombardy (4.6 points each) and Tuscany (4.9 points). For the country as a whole, this percentage should increase by 6.8 percentage points, but with Central Italy growing much faster (mainly as a result of population growth in Rome) than in the north or south of the country.

In the variant based on the development of six large cities, simulations of $\frac{y_i^*}{y_B^*}$ the relationship are examined in the following scenarios:

1. investment rates and growth rates in the number of employees are following the same pattern as on average between 2000 and 2019,
2. the growth rates of the number of employees are as they have been on average over the last two decades, while the investment rates in each region are $\overline{s} \in (0,1)$,

[9] However, this very strong assumption can be weakened relatively easily to come up with urbanisation rates, as in Table 4. Specifically, it can be assumed that the ratio of the population of the 6 largest cities to the population of the regions in which they are located will increase by 50%, while the remaining cities with populations between 100,000 and 500,000 will increase by 25%. Thus, denoting the current population of Rome (or any other city with a current population of more than 0.5 million people) by R, the long-run value of this characteristic by R_∞, the current population of Lazio (or the region in which any other city with a population of more than 500,000 people is located) as L, the long-run as L_∞ and assuming that $\frac{L_\infty}{L} = \lambda > 1$, we conclude that, by virtue of the assumption made, the following occurs: $R_\infty = 1{,}5\ \lambda R$. Therefore, assuming, for example, that with a 50% increase in population in the Italian regions the population of Rome should increase from 2.8 million people in 2020 to 6.3 million people in long run, Milan from 1.4 million to 3.1 million, Naples from 935,500 people to 2.1 million, Turin from 858,100 people to 1.9 million people, Palermo from 642,700 people to 1.4 million, and Genoa from 566,100 people to 1.3 million.

3. employment growth rates equal $\overline{n} > 0$, and investment rates equal $\overline{n} > 0$, as on average over the period 2000–2019,
4. these rates are equal to (respectively) $\overline{s} \in (0, 1)$ and $\overline{n} > 0$.

In the results of the simulations based on the development of the six cities presented in detail in Table 5, if investment and labour growth rates averaged as in 2000–2019 (Scenario I), in the long run labour productivity in the northern regions of Italy would be higher than in Lombardy (by 8.9%), while it would be lower in the central and southern regions of Italy (by 2.7% and 3.5%, respectively). Labour productivity in the Italian economy would then be 2.9% higher than in the base region. In the scenario considered, the highest relative labour productivity would be registered for Aosta Valley (1.229), Piedmont (1.218), Emilia-Romagna (1.202), Molise (1.186), the Autonomous Province of Bolzano (1.144) and Umbria (1.129). In contrast, Puglia (0.861) and Marche (0.890) would have the lowest relative labour productivity values. Changing the assumption on investment rates and assuming that they are the same in all regions changes the relationship in long-run labour productivity as follows. Northern and Central Italy would have higher labour productivity than Lombardy (by 2.7% and 3%, respectively), while Southern Italy would still have lower labour productivity than in the base region (by 3.4%). Total labour productivity for Italy would be 1% higher than in Lombardy. Liguria (1.130), Piedmont (1.075), Emilia-Romagna (1.073), Lazio (1.057) and Umbria (1.053) would have the highest relative labour productivity, while the Autonomous Province of Bolzano (0.959), Marche (0.951), Puglia (0.937), Sardinia (0.919), Calabria (0.911) and Basilicata (0.901) would have the lowest value of the analysed variable.

Changing (with respect to Scenario 1) the assumption on labour growth rates and assuming that they are the same in all regions results in simulated labour productivity in Northern Italy being 7.1% higher than in the base region. In the variant analysed, Central and Northern Italy have labour productivity lower than in Lombardy by 2.4% and 8.9%, respectively. In contrast, labour productivity in the Italian economy as a whole would be 0.5% higher than the value of this factor in Lombardy. Among the regions with the highest relative labour productivity were: Emilia-Romagna (1.202), Autonomous Province of Bolzano (1.180), Aosta Valley (1.147), Piedmont (1.144), Autonomous Province of Trento (1.109) and Umbria (1.109), while the lowest relative labour productivity would be characterised by: Calabria (0.891), Marche (0.866), Sicily (0.861) and Puglia (0.809).

In Scenario 4, it is assumed that both investment and labour growth rates are the same in all regions, in which case the variation in labour productivity is only affected by the geographical location of the regions. In this scenario, Central and Northern Italy would have 3.3% and 1% higher labour productivity than Lombardy, respectively, and Southern Italy would have 8.8% lower

labour productivity than in the base region. In this scenario, labour productivity in the Italian economy as a whole would be lower than in Lombardy in the long term (by 1.4%). In the scenario analysed, the highest relative labour productivity would be in: Lazio (1.086), Emilia Romagna (1.072), Liguria (1.065), Umbria (1.035), Piedmont (1.009), the Autonomous Province of Trento (1.005) and Tuscany (1.004), while the lowest would be Molise (0.896), Apulia (0.881), Basilicata (0.864) and Calabria (0.842).

Table 5. Labour productivity in long-run equilibrium—variant based on the six largest cities (Lombardy = 100)

Region or group of regions	Scenario			
	1	2	3	4
Piedmont	121.8	107.5	114.4	100.9
Aosta Valley	122.9	101.2	114.7	94.4
Liguria	108.2	113.0	102.0	106.5
Autonomous Province of Bolzano	114.4	95.9	118.0	98.9
Autonomous Province of Trento	108.1	98.0	110.9	100.5
Veneto	107.3	100.0	105.2	98.1
Friuli-Venezia Giulia	106.0	97.4	102.9	94.5
Emilia-Romagna	120.2	107.3	120.2	107.2
Northern Italy	**108.9**	**102.7**	**107.1**	**101.0**
Tuscany	99.7	101.9	98.2	100.4
Umbria	112.9	105.3	110.9	103.5
Marche	89.0	95.1	86.6	92.6
Lazio	95.9	105.7	98.4	108.6
Central Italy	**97.3**	**103.0**	**97.6**	**103.3**
Abruzzo	100.9	97.0	98.4	94.7
Molise	118.6	100.2	106.1	89.6
Campania	101.9	101.3	96.5	96.0
Apulia	86.1	93.7	80.9	88.1
Basilicata	103.9	90.1	99.6	86.4
Calabria	96.5	91.1	89.1	84.2
Sicily	92.9	98.3	86.1	91.1
Sardinia	104.6	91.9	103.2	90.7
Southern Italy	**96.5**	**96.6**	**91.1**	**91.2**
ITALY	**102.9**	**101.0**	**100.5**	**98.6**

Source: own calculations based on equation (14).

In addition to the scenarios discussed for the development of labour productivity, using equation (14), the relationship between investment rates in the subsequent regions (relative to Lombardy's investment rate) at which full labour productivity convergence would occur was determined (please see Table 6). These simulations were performed under the following scenarios:

Table 6. Relationship of investment rates ensuring full convergence (Lombardy = 100)

Region or group of regions	Scenario			
	1	2	3	4
Piedmont	92.4	93.7	100.5	92.3
Aosta Valley	98.9	93.5	101.0	98.7
Liguria	88.4	93.8	100.2	87.3
Autonomous Province of Bolzano	107.3	105.5	102.0	104.7
Autonomous Province of Trento	102.5	103.7	100.8	102.2
Veneto	99.8	99.6	101.7	100.0
Friuli-Venezia Giulia	102.7	102.3	105.8	102.9
Emilia-Romagna	94.7	99.7	99.7	92.5
Northern Italy	**97.5**	**98.8**	**100.7**	**97.1**
Tuscany	98.0	98.4	100.0	97.9
Umbria	95.4	99.2	101.2	94.4
Marche	103.9	100.1	103.2	105.7
Lazio	96.1	106.3	103.2	94.0
Central Italy	**97.7**	**102.3**	**102.0**	**96.8**
Abruzzo	103.4	100.6	103.4	103.4
Molise	100.0	94.6	107.0	99.8
Campania	100.3	101.1	107.4	98.6
Apulia	107.3	106.8	114.3	107.4
Basilicata	112.4	106.3	111.4	112.2
Calabria	110.8	107.8	117.7	110.8
Sicily	102.4	105.1	114.3	101.9
Sardinia	109.6	109.3	110.9	109.8
Southern Italy	**104.5**	**104.4**	**111.3**	**103.9**
ITALY	**99.4**	**101.0**	**103.8**	**98.9**

Source: own calculations based on equation (14).

1. historically shaped growth rates of employment and urbanisation rates,
2. historically shaped growth rates of employment and the same rates of urbanization,
3. the same rates of growth in employment and the same rates of urbanisation,
4. historical growth rates of employment and urbanisation rates as in Table 4,
5. same growth rates of employment and urbanisation rates as in Table 4.

The results of simulations of investment rate relations guaranteeing full labour productivity convergence are presented in Table 6. These simulations show that in all the scenarios analysed, relative investment rates[10] (understood as the quotient $\frac{s_i}{s_B}$ defined by equation (14) in Southern Italy should exceed 1. In scenarios 1, 2 and 4, investment rates in the southern Italian regions should be about 4% higher than in Lombardy, in scenarios 3 and 5, more than 10% higher. In scenarios 1, 4 and 5, investment rates in the central Italian regions should be lower than in Lombardy (by around 3%–4%), while in scenarios 2 and 3 they should be higher by around 2%–2.5%. In the regions of northern Italy, outside scenario 3, investment rates in these regions should be 2%–3% lower than in the base region. Generally, it can be hypothesised that for full labour productivity convergence to occur in Italy, investment rates in the south of the country are needed to be higher than in Lombardy, while those in the centre and north are needed to be similar to the region.

Conclusions

The analyses presented in the paper can be summarised as follows: The fall in investment in the Italian economy following the global financial crisis has translated into a reduction in the rate of capital accumulation, a fall in the value of domestic gravitational effects, leading to a fall in production, labour productivity, employment and an increase in unemployment. In addition, a significant drop in production was recorded there in 2020 as a result of restrictions related to the COVID-19 pandemic. As a consequence of these processes, the volume of GDP in Italy in 2021 is close to the value of this macroeconomic variable at the beginning of the 21st century, while capital

[10] Since we are examining the quotients $\frac{s_i}{s_B}$, a value of this quotient of, say, 1.02 means that the investment rate in region i should be 2% higher than in the base region (i.e. if the base region had an investment rate of 20%, region i should have an investment rate of 20.4%).

per worker, gravitational effects and labour productivity are lower (which is unique among EU countries).

The fall in investment has affected the poorest regions, Southern Italy, the most, followed by Central Italy, and the richest regions of the north of the country the least. In turn, declines in investment have led to a further polarisation of capital per worker, gravitational effects and labour productivity in Italian regions. The coefficient of variation (defined as the quotient of the standard deviation and the unweighted mean) for capital per worker increased by more than 25% between 2010 and 2020, for gravitational effects by more than 3%, and for labour productivity by more than 5%. Generally, moving from north to south Italy, the level of capital per worker and labour productivity decreases, while the strongest gravitational effects are recorded in Central Italy in the quadrilateral connecting Rome, Florence, Bologna and Perugia.

Analysing the spatial variation in labour productivity on the basis of the gravitational growth model, we find that the level of labour productivity in the country's regions is most strongly influenced by the variation in capital per worker, followed by urbanisation rates and domestic and foreign gravitational effects. The study calibrated the parameters of the gravitational model of economic growth in two ways: not taking into account and taking into account external gravitational effects, i.e. effects coming from abroad. In both cases, estimates were made using the method of least squares (OLS) and the generalized method of moments (GMM). A better explanation of the model was obtained in the model with foreign effects. The GMM-estimated parameters of the model including external gravitational effects were used in numerical simulations.

Numerical simulations of the long-run relationship of labour productivity in Italian regions in relation to the value of this variable in the base region (Lombardy) were carried out in two variants: a baseline variant and a variant with development based on the six largest cities. The baseline variant assumed eight different scenarios for the development of investment rates, labour force growth rates and urbanization rates. In the variant with development based on the 6 largest cities, four different scenarios were analysed. In the baseline variant, regardless of the scenario adopted, Northern, Central and Southern Italy would have higher relative labour productivity in long-run equilibrium relative to Lombardy than in 2020. Northern Italy's relative long-run labour productivity would be highest under a scenario in which investment rates, labour force growth rates and urbanization rates would be at historical levels (that is, as they were on average from 2000 to 2019). The lowest relative labour productivity growth in the area would occur under the assumption that investment rates, urbanization rates and labour force growth rates equalize across all regions. In contrast, for the central Italian regions, the most favourable scenario was the one in which urbanization rates are assumed to be at historical levels and other rates are assumed to be at some level that is the

same for all regions (in this scenario, Central Italy achieves a higher level of labour productivity than Northern Italy in long-term equilibrium). The least favourable scenario for the central regions is the one that assumes investment and labour growth rates at historical levels and an urbanisation rate that is the same for all regions. The southern Italian regions, regardless of the scenario adopted, are characterized in long-run equilibrium (as they are now) by the lowest relative labour productivity.

In the variant with development based on the six largest cities, the most favourable scenario for Northern Italy would be one with investment rates and labour growth rates maintained at historical levels, the least favourable for the area would be an equalization of investment rates and labour growth rates between regions (in this scenario, in long-term equilibrium, they would have lower relative labour productivity than central Italy). Equally favourable for Central Italy would be a scenario in which labour growth rates are at historical levels and only investment rates equalize across all regions (this is also the most favourable scenario for southern Italian regions). At the same time, as in the baseline scenario, Southern Italy would have the lowest relative level of labour productivity in the long term regardless of the assumptions used. The study also determined the relationship between investment rates in the regions relative to Lombardy, at which full labour productivity convergence would occur. From the analysis, a general conclusion can be drawn that for full convergence in labour productivity to occur in Italy, it is necessary to keep investment rates in the south of the country higher than in Lombardy (by about 4%–11%, depending on the scenario adopted for labour and urbanization growth rates), and investment rates in central and northern Italy at similar levels to those in Lombardy.

One can expect that the following three actions may be helpful in striving for the convergence of the economic development of Italian regions. Firstly, striving to increase broadly understood social capital in the southern regions of the country (that are more susceptible to corruption and crime actions at least since the unification of Italy in the second half of the 19th century). Secondly, differentiating the minimum wage depending on the situation on the labour market (which may result in an increase in demand for labour and an increase in employment in regions with the highest unemployment, without losses to the labour markets in regions with low unemployment). Thirdly, differential taxation of capital depending on regional investment rates.

Further analyses may concern extending the model with variables characterizing the diversity of the labour market (in order to capture the interaction between variables in the labour market and the product market). Moreover, the impact of the economic policy instruments mentioned in the comments on the processes of regional convergence (both on the product and labour markets) can also be analysed.

References

Allen, R. G. D. (1975). *Teoria makroekonomiczna. Ujęcie matematyczne*. Wydawnictwo Naukowe PWN.

Arbia, G., & Basile, R. (2005). Spatial dependence and non-linearities in regional growth behaviour in Italy. *Statistica*, *65*(2), 145–167. https://doi.org/10.6092/issn.1973-2201/83

Bärwolff, G. (2020). Mathematical modeling and simulation of the COVID-19 pandemic. *Systems*, *8*(3), 24. https://doi.org/10.3390/systems8030024

Basile, R., & Cicerone, G. (2022). Economic complexity and productivity polarization: Evidence from Italian provinces. *German Economic Review*, *23*(4), 567–594. https://doi.org/10.1515/ger-2021-0070

Bianchini, F. (1991). The Third Italy: Model or myth? *Ekistics*, *58*(350/351), 336–345. https://www.jstor.org/stable/43646779

Boschma, R., & Iammarino, S. (2007). *Related variety and regional growth in Italy*. SPRU Electronic Working Paper Series.

Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography*, *85*(3), 289–311. https://doi.org/10.1111/j.1944-8287.2009.01034.x

Busetta, G., & Corso, D. (2012). Labor productivity and Okun's law: An empirical application to Italian regional panel data. Rivista italiana degli economisti, *The Journal of the Italian Economic Association*, *2*, 279–298. https://doi.org/10.1427/37497

Calcagnini, G., Marin, G., & Perugini, F. (2021). Labour flexibility, internal migration and productivity in Italian regions. *Structural Change and Economic Dynamics*, *57*(C), 308–320.

Carmeci, G., & Mauro, L. (2002). The convergence of Italian regions and unemployment: Theory and evidence. *Journal of Regional Science*, *42*, 509–532.

Di Giacinto, V., & Nuzzo, G. (2004). *Explaining labor productivity differentials on Italian regions*. ERSA Conference Papers. European Regional Science Association.

Di Liberto, A. (2008). Education and Italian regional development. *Economics of Education Review*, *27*(1), 94–107. https://doi.org/10.1016/j.econedurev.2006.08.004

Di Liddo, G. D. (2015). Urban sprawl and regional growth: Empirical evidence from Italian regions. *Economics Bulletin*, *35*, 2141–2160.

Dykas, P., Tokarski, T., & Wisła, R. (2023). *The Solow model of economic growth. Application to contemporary macroeconomic issues*. Routledge.

Dykas, P., & Wisła, R. (Eds.). (2022). *The socioeconomic impact of COVID-19 on Eastern European countries*. Routledge.

Filipowicz, K. (2019). *Wpływ interakcji przestrzennych na zróżnicowanie rozwoju gospodarczego Polski*. Wydawnictwo Uniwersytetu Jagiellońskiego.

Filipowicz, K., Nowosad, A., & Tokarski, T. (2015). Oddziaływanie efektu grawitacyjnego na zróżnicowanie wydajności pracy w krajach bałkańskich. *Gospodarka Narodowa*, *2*, 15–53.

Fiorino, N., Galli, E., & Petrarca, I. (2012). Corruption and growth: Evidence from the Italian regions. *European Journal of Government and Economics*, *1*, 126–144.

Fratesi, U., & Percoco, M. (2014). Selective migration, regional growth and convergence: Evidence from Italy. *Regional Studies*, *48*(10), 1650–1668. https://doi.org/10.1080/00343404.2013.843162

Gitto, S., & Mancuso, P. (2015). The contribution of physical and human capital accumulation to Italian regional growth: A nonparametric perspective. *Journal of Productivity Analysis*, *43*, 1–12.

Gori, L., Manfredi, P., Marsiglio, S., & Sodini, M. (2022). COVID-19 epidemic and mitigation policies: Positive and normative analyses in a neoclassical growth model. *Journal of Public Economic Theory*, *24*, 968–992.

Maffezzoli, M. (2004). *Convergence across Italian regions and the role of technological catch-up*. IGIER Working Paper, 274. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=628481

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, *107*(2), 407–437. https://doi.org/10.2307/2118477

Mroczek, K., Tokarski, T., & Trojak, M. (2014). Grawitacyjny model zróżnicowania rozwoju ekonomicznego województw. *Gospodarka Narodowa*, *3*, 5–34.

Mroczek, K., Tokarski, T., & Trojak M. (2015). Złote reguły akumulacji w grawitacyjnym modelu wzrostu gospodarczego. *Gospodarka Narodowa*, *3*, 27–47.

Nowosad, A., & Wisła, R. (red.). (2016). *Zróżnicowanie rozwoju współczesnej Europy*. Wydawnictwo Uniwersytetu Jagiellońskiego.

Ombach, J. (1999). *Wykłady z równań różniczkowych wspomagane komputerowa – Maple*. Wydawnictwo Uniwersytetu Jagiellońskiego.

Paci, R., & Pigliaru, F. (1997). Structural change and convergence: an Italian regional perspective. *Structural Change and Economic Dynamics*, *8*(3), 297–318. https://doi.org/10.1016/S0954-349X(96)00077-X

Paci, R., & Pigliaru F. (1998). *Growth and sectoral dynamics in the Italian regions*. Working Paper CRENoS 199803. Centre for North South Economic Research, University of Cagliari and Sassari.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65–94. https://doi.org/10.2307/1884513

Terrasi, M. (1999). Convergence and divergence across Italian regions. *The Annals of Regional Science*, *33*, 491–510. https://doi.org/10.1007/s001680050117

Wisła R., Filipowicz, K., & Tokarski, T. (2018). Zróżnicowanie rozwoju ekonomicznego krajów UE na podstawie grawitacyjnego modelu wzrostu. *Wiadomości Statystyczne*, *7*, 37–55.

Wisła, R., & Nowosad, A. (Eds.). (2020). *Economic transformation in Poland and Ukraine. National and Regional Perspective*. Routledge.

sciendo Economics and Business Review https://www.ebr.edu.pl Vol. 10 (2), 2024: 118–142 https://doi.org/10.18559/ebr.2024.2.1149

Personal bankruptcy prediction using machine learning techniques

Magdalena Brygała[1] Tomasz Korol[2]

Abstract

It has become crucial to have an early prediction model that provides accurate assurance for users about the financial situation of consumers. Recent studies have focused on predicting corporate bankruptcies and credit defaults, not personal bankruptcies. Due to this situation, the present study fills the literature gap by comparing different machine learning algorithms to predict personal bankruptcy. The main objective of the study is to examine the usefulness of machine learning models such as SVM, random forest, AdaBoost, XGBoost, LightGBM, and CatBoost in forecasting personal bankruptcy. The study relies on two samples of households (learning and testing) from the Survey of Consumer Finances, which was conducted in the United States. Among the models estimated, LightGBM, CatBoost, and XGBoost showed the highest effectiveness. The most important variables used in the models are income, refusal to grant credit, delays in the repayment of liabilities, the revolving debt ratio, and the housing debt ratio.

Keywords

- personal bankruptcy
- SVM
- random forest
- AdaBoost
- XGBoost
- LightGBM
- CatBoost
- SHAP

JEL codes: G17, G51

Article received 23 January 2024, accepted 7 April 2024.

Funding: This research was prepared within grant project No. 2017/25/B/HS4/00592, "Forecasting the risk of consumer bankruptcy in Poland." This research was funded by the National Science Center in Poland (Narodowe Centrum Nauki).

Suggested citation: Brygała, M., & Korol, T. (2024). Personal bankruptcy prediction using machine learning techniques. *Economics and Business Review*, *10*(2), 118–142. https://doi.org/10.18559/ebr.2024.2.1149

[1] Faculty of Management and Economics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland, corresponding author: magdalena.brygala@pg.edu.pl, https://orcid.org/0000-0003-3222-1046.

[2] Faculty of Management and Economics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland, tomasz.korol@zie.pg.gda.pl, https://orcid.org/0000-0002-7623-3404.

Introduction

The economies of countries that have not managed to recover from the COVID-19 pandemic have to face another challenge for countries around the world, which is the war in Ukraine. Consequently, the risk of consumer bankruptcy has increased dramatically. Therefore, searching for more precise methods and testing modern solutions for consumer bankruptcy prediction is essential. Before the COVID-19 pandemic, it was hard to imagine the impact it could have on economies worldwide. Some countries very quickly implemented support for companies to minimise the effects of the financial crisis. However, soon, before the COVID-19 pandemic was forgotten, the war in Ukraine began, which also affected countries on different continents. Additionally, forecasting the timing of economic recessions is very difficult (Altman & Kuehne, 2016). It is important to anticipate bankruptcy as soon as possible in order to avoid it. In addition to declaring bankruptcy, consumers have various options to deal with problems associated with paying off liabilities. The later these problems are noticed, the more difficult it is to avoid bankruptcy. This study aims to create a good classification model for predicting bankruptcy. However, developing such models to predict bankruptcy risk with high accuracy is challenging because bankruptcy rates are low, and there are few datapoints on which to base predictions (Garcia, 2022).

Many empirical studies have been developed on predicting the risk of corporate bankruptcy and non-performing loans (Barboza, Basso et al., 2021; Barboza, Kimura et al., 2017; Garcia, 2022; Kovacova et al., 2019; Kovacova & Kliestikova, 2017; Letza et al., 2003; Wang et al., 2022), and few studies concern personal bankruptcy (Brygała, 2022; Korol, 2021; Korol & Fotiadis, 2022; Sahiq et al., 2022; Syed Nor et al., 2019). The small body of empirical research on consumer bankruptcy stems from such factors as limited access to data related to consumer bankruptcy. Due to the fact that very few publications focus on forecasting consumer bankruptcy and that there is a research gap in this area, the main goal of this study is to develop predictive machine learning models of consumer bankruptcy based on data from the United States (Survey of Consumer Finances). To fill this gap in the literature, this study is one of the first literary attempts to develop machine learning models in personal bankruptcy prediction.

The contribution of this study to the literature on forecasting the risk of personal bankruptcies is four-fold. First, our research analysed the performance of six machine learning methods: support vector machine (SVM), random forest (RF), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and categorical boosting (CatBoost), which were applied to the problem of personal bankruptcy prediction. Second, it identifies the most important predictors of filing

for personal bankruptcy. Third, it compares the machine learning models to the results obtained by other methods in the literature of corporate bankruptcy and default prediction. Fourth, it examines SHapley Additive exPlanations (SHAP) to help interpret machine learning model predictions and explore the importance of various features that affect bankruptcy. Moreover, the authors of this research formulated the following research questions:

1. Which model can obtain the highest total effectiveness and the lowest type I and II errors?
2. What are the main microeconomic predictors of filing for personal bankruptcy?

The paper is organised into five sections. In the introduction, the authors justify the topic, the research objectives, and the study's contribution to the literature. Section 1 provides a review of bankruptcy and default predictive models. Section 2 describes the data used in the analysis and the forecasting methods implemented. Section 3 presents six machine learning models. Section 4 discusses the results obtained from the testing sample. Finally, the conclusion section summarises the research.

1. Literature review

Researchers and practitioners have conducted intensive research on models for predicting company bankruptcy and default on loans, both among enterprises and consumers. Among the algorithms used for prediction purposes are traditional statistical techniques (e.g., discriminant analysis and logistic regression), deep learning (e.g., artificial neural networks), and machine learning models (e.g., support vector machine, bagging, boosting, and random forest) (Shi et al., 2022). Machine learning techniques identify characteristics that differentiate the observations of different groups (Barboza, Kimura et al., 2017). They are used in many fields, such as economics, medicine and engineering.

Machine learning and deep learning models have been very successful in financial applications, with many studies looking at their use in predicting bankruptcy. Both models have advantages over traditional statistical methods when there are a large number of variables, the relationships between the variables are complex, the values of each variable change over time, and when it is more important to understand the correlations between variables than to look for causality (Shi et al., 2022; Syam & Sharma, 2018). The advantages of using machine learning and artificial intelligence include their dynamism, which allows for running background processes and making decisions in real time (Syam & Sharma, 2018). To overcome the limitations of statisti-

cal models, research has been developed that actively uses pattern recognition methods in machine learning (Son et al., 2019). In the latest research, the most commonly used algorithms are neural networks, boosting and bagging methods, and logistic regression (Al Daoud, 2019). Al Daoud (2019) noted how research has shown that gradient-boosting algorithms are used successfully and represent a very important strategy. Prior research showed that machine learning models are more suitable for predicting the risk of bankruptcy than statistical models (Garcia, 2022; Machado & Karray, 2022; Son et al*.,* 2019). Carmona et al. (2022) also pointed out how recent research shows that gradient boosting can reduce the weaknesses of traditional models and provide an effective model for predicting business failures. The term "black box" is applied to models where we know the inputs and outputs, but we can say little about what is going on inside (Gramegna & Giudici, 2021). However, machine learning models are often considered a black box due to their complexity and hidden internals (Carmona et al., 2022). Brotcke (2022) stated that the less transparency and explainability of machine learning models compared to traditional regression models may lead to discussions about the compliance of models with fair lending regulations. Black boxes that are more complex are more accurate for the highest predictive performance but are often more challenging to interpret. However, in recent years, researchers have proposed improvements to increase the interpretability of machine learning models. One common approach to explaining machine learning models is the SHAP method (Bussmann et al*.,* 2020), which is often performed to interpret complex models (Bussmann et al*.,* 2020; Jabeur, Mefteh-Wali et al*.,* 2021). Brotcke (2022) also pointed out that machine learning can reduce potential discrimination by limiting discretionary and judgmental decisions. This can be crucial, for example, in the case of using indicators or variables containing discriminatory factors such as age, marital status and gender. Due to the importance of the topic of discrimination, in the United States it is illegal for lenders to discriminate against consumers on the basis of: race, colour, religion, national origin, sex, marital status, age, attendance in a public assistance programme (CFPB, 2022).

In the latest research, Papík and Papíková (2023) analysed studies focusing on gradient-boosting algorithms. They noticed that most studies achieved higher performance with gradient boosting, especially XGBoost. In the research analysed, only one study applied Catboost, which proved to be the most effective algorithm. In two cases, the application of a neural network outperformed gradient boosting. Jabeur, Gharib et al. (2021) developed neural network and machine learning models to overcome the limitations of such initial models like discriminant analysis and logistic regression. Sahiq et al. (2022) examined the usefulness of logistic regression in forecasting consumer bankruptcy. They showed that the key determinants of personal bankruptcy include race, education, employment sector, personal loan, study loan, microfinance,

and total outstanding balance. Korol and Fotiadis (2022) proposed artificial intelligence techniques: fuzzy sets, artificial neural networks, and genetic algorithms in forecasting the risk of personal bankruptcy. The fuzzy sets outperformed the other techniques in total effectiveness and with the lowest type I and II errors both for Taiwanese and Polish households. The research also proved that artificial intelligence models outperformed the statistical models estimated in previous research (Korol, 2021) based on the same samples. Research conducted by Shi et al. (2022) showed that most deep learning models outperform classical machine learning and statistical algorithms in estimating credit risk. Moreover, team methods provide greater accuracy compared to single models. Alam et al. (2021) compared deep learning with discrete hazard models. Deep learning performed better than discrete hazard models in predicting corporate failure. Halim et al. (2021) developed deep learning models such as: recurrent neural network, long short-term memory, gated recurrent unit, as well as logistic regression, support vector machine, neural network and decision tree. Their research showed that all deep learning models outperform other widely used methods. Bragoli et al. (2022) found that XGBoost performed better in correctly classifying bankrupt firms. Other methods, such as random forest and neural network, were better at classifying non-bankrupt firms. Machado and Karray (2022) proposed hybrid machine learning algorithms for predicting commercial customer credit scores. They compared the effectiveness of hybrid and individual algorithms (AdaBoost, decision tree, random forest, support vector machine, artificial neural network, and gradient boosting). For hybrid models, data is first grouped using a classifier method (k-Means and DBSCAN), then different machine learning models are applied to each of the obtained clusters to predict a given event. Hybrid models outperformed individual ones.

2. Data and methodology

2.1. Data

The study used microdata from 37,900 surveys conducted between 2001 and 2019 in the Survey of Consumer Finances (SCF). SCF is a survey in the United States, which includes household characteristics such as: demographic, behavioural and financial. Unanswered questions in the survey were covered by the multiple imputation technique. The data include a dependent variable of 1 for households that have filed for bankruptcy in the last five years, and 0 otherwise. The independent variables selected during model prepara-

tion include demographic and financial characteristics (Table 1). The data include only consumers who have debt. In the models, the inverse hyperbolic sine transformation (IHS) of income is used, which allows the use of samples with zero and negative values (Berlemann & Salland, 2016; Georgarakos et al., 2014). The formula of IHS applied to income is:

$$\log\left(x+(x^2+1)^{\frac{1}{2}}\right)$$

Independent variables such as sex, age, marital status and race were not used. The application of these variables may be discriminatory to the consumer due to the Equal Credit Opportunity Act in the United States (Brotcke, 2022). The above-mentioned federal civil rights apply to credit cards, car loans, home loans, student loans and business loans (CFPB, 2022). The regulations are designed to protect consumers by prohibiting unfair and discriminatory approaches (Brotcke, 2022).

Table 1. The list of variables used in evaluating models

Variable	Description
children	The number of children.
saving_account	The dummy variable is 1 if the respondent has a saving account.
turndown	The dummy variable is 1 if the respondent applied for a loan and was turned down.
late	The dummy variable is 1 if the household had any past payments due in the last year.
income	The inverse hyperbolic sine transformation of income.
homeownership	The dummy variable is 1 if the respondent owns, e.g., ranch/farm/mobile home/house/condo, 0: otherwise.
education	The variable education is described by four values: no high school, high school, college associate degree, bachelor's degree or higher.
mortgage_asset	It represents the proportion of housing debt to the value of total assets.
consumer_debt	It represents the total non-mortgage and non-revolving consumer debt proportion to the total monthly payments.
revolving_debt	It represents the proportion of revolving debt to the total monthly payments.
house_debt	It represents the proportion of housing debt to the total monthly payments.
income_debt	It represents the proportion of income to the total monthly payments.

Source: own research.

Data were divided to train and test samples to avoid overfitting and bias. Models were assessed for different proportions between the training sample and the testing sample to maximise the training sample. To evaluate the models, 80% of the dataset is used for the learning sample (1531 consumers), while 20% is set aside for the testing sample (383 consumers). The dataset is highly unbalanced and skewed towards consumers who did not decide to go bankrupt (negative class). The proportion of bankruptcies to non-bankruptcies stands at the level of 4.38%. Predicting rare events like bankruptcies is often challenging in view of possible bias in estimating probabilities. Without using methods dealing with an unbalanced dataset, the minority class may be ignored in the prediction. The researchers proposed several methods to deal with this challenge, both at the algorithm level and data level (Yen & Lee, 2009). Among the methods used for such a challenge are: undersampling, oversampling, a combination of undersampling and oversampling methods, choosing a cut-off point, and using class weight. Therefore, in the research, the undersampling method was used to balance consumers who decided to file for bankruptcy and those who did not file for bankruptcy. Moreover, the data were preprocessed using StandardScaler. This is a normalisation technique which normalises the features to create standardised features by removing the mean and scaling to unit variance (Le et al*.,* 2018). Stata and Python software was used in the preprocessing step. Next, the models were implemented using Python software packages. The authors used six methods to forecast personal bankruptcy: SVM, RF, AdaBoost, XGBoost, LightGBM, and CatBoost. Such models were calculated for 1914 consumers: 957 bankrupts and 957 non-bankrupts.

2.2. Machine learning models

2.2.1. Support Vector Machine

SVM (Support Vector Machine) is a machine learning algorithm used both for regression and classification problems. The objective of SVM is to find a hyperplane which can segregate the n-dimensional space into classes. The hyperplane is the boundary of classification between two classes with the highest margin. Support vectors are the datapoints which are closest to the hyperplane and create the hyperplane. The strength of SVM is that despite the significant overlap between different data classes, it finds the decision boundary. The SVM model is then ready to classify the new datapoints on the side of the hyperplane to which they should be mapped. However, SVM is more time-consuming, has high algorithm complexity, and requires large memory capacity (Jabeur, Gharib et al*.,* 2021).

2.2.2. Random Forest

The RF (Random Forest) method was proposed by Breiman (2001). The forest consists of several subsets that generate the same number of classification trees, and responses are combined (Barboza, Kimura et al., 2017; Schonlau & Zou, 2020 ; Wu et al., 2016). In RF, the sample features and the number of samples are selected randomly (Wang et al., 2022). Creating multiple trees instead of one and combining the results gives a more stable prediction than a single tree. Each tree is built on a different bootstrap sample that was created by randomising and returning N objects from all N training samples. The prediction result of RF in classification problems is the largest class among all the prediction results of decision trees (Wang et al., 2022).

2.2.3. Adaptive Boosting

AdaBoost (Adaptive Boosting) is one of the machine learning algorithms proposed by Freund and Schapire (1997). The method involves fitting a sequence of weak classifiers, which are models that are only slightly better than random guessing, to multiple modified versions of the data (Barboza, Basso et al., 2021). By incorporating weak classifiers, AdaBoost constructs a more powerful learning algorithm, enhancing the strength of the classifiers (Heo & Yang, 2014). In AdaBoost, the approach is sequential, and the successive classifiers are closely related. If the resulting classifier achieves higher accuracy compared to the default rule, it means that the classification method has identified certain patterns or structures in the data that allow it to perform better (Alfaro et al., 2008).

2.2.4. Extreme Gradient Boosting

XGBoost (Extreme Gradient Boosting) is a methodology for regression as well as classification. It constitutes the implementation of a gradient boosting framework developed by Chen and Guestrin (2016). XGBoost is an ensemble model based on gradient boosted trees (Mo et al., 2019). XGBoost starts with creating a first weak tree with poor performance, then it builds another tree based on the previous tree in the next stage, trying to predict what the first tree could not have predicted. The algorithm continues to build trees, each of which corrects the previous one, until a stop condition is reached, such as the number of trees to be built. In the objective function, normalisation is used to prevent overfitting, estimate the model more efficiently and minimalise the complexity of the model (Jabeur, Mefteh-Wali et al., 2021). Al Daoud (2019) pointed out that the technique used in XGBoost makes the model faster and more stable during model fitting. In addition, there are several hyperparameters, which can be modified to maximise the power of the model and to prevent the overfitting of the model.

2.2.5. Light Gradient Boosting Machine

LightGBM (Light Gradient Boosting Machine) is the implementation of a gradient-boosted framework proposed by Ke et al. (2017). Research has shown that in the case of the used dataset, LightGBM is faster and more accurate than CatBoost and XGBoost (Al Daoud, 2019). Decision trees in the LightGBM algorithm are grown leaf-wise instead of checking all previous leaves for each new leaf, as with XGBoost (Al Daoud, 2019). LightGBM uses a histogram algorithm to combine exclusive features (Wang et al., 2022). The advantage of the LightGBM algorithm is its high accuracy and model training speed, low memory consumption, and that it is adapted to the use of large datasets (Al Daoud, 2019; Ke et al., 2017). Ke et al. (2017) pointed out that LightGBM, in addition to reducing the training time by more than 20 times compared to the gradient-boosting decision tree, achieved almost the same accuracy. However, having a large dataset affects the model training time. Therefore, the choice between a shorter training time and the model's accuracy is not so obvious, especially when the accuracy is not much higher with a shorter model training time.

2.2.6. Categorical Boosting

CatBoost (Categorical Boosting) also belongs to the gradient-boosted binary trees. This is a new gradient algorithm proposed by Prokhorenkova et al. (2018). CatBoost, like other gradient-boosting implementations, constructs each new tree to approximate the gradients of the current model (Dorogush et al., 2018). The objective of CatBoost is to minimise the loss function of the model by adding weak learners with a gradient-descent-like procedure (Papík et al., 2023). One of the advantages of CatBoost is that this algorithm has the ability to work with categorical variables. Dorogush et al. (2018) noted that CatBoost followed by LightGBM are rivals for the fastest method, while XGBoost is much slower than both methods. This is important for large datasets. Hancock and Khoshgoftaar (2020) pointed out that CatBoost exhibits sensitivity to hyperparameters and emphasized the significance of hyperparameter tuning.

2.3. Evaluation metrics

Some of our models, such as RF, AdaBoost, XGBoost, LightGBM, and CatBoost provide a measurement of the importance of features. Feature importance was also used in selecting variables. The importance is the average for each single decision tree in the model, and it is computed as the amount by which the feature split point improves accuracy, weighted by the number

of samples on each node (Son et al., 2019). A higher score for feature importance means that the specific feature will have a greater effect on that model. It determines which features contribute most to the predictive power of the model. The same technique can be used for both feature selection and feature importance. However, feature selection is most commonly used before or during model training to select features, while feature importance measures are used during or after training to explain the trained model (Saarela & Jauhiainen, 2021).

This study will use total effectiveness (S), type I error (E_1), type II error (E_2), and AUC as measurements of performance. A type I error shows false prediction of bankrupts (D_1) among all bankruptcies (BR), while a type II error indicates false prediction of non-bankrupts (D_2) among all non-bankruptcies (NBR). The measurement was calculated using a confusion matrix, which is intended to compare the actual classification with the predicted classification. Total effectiveness shows the probability of an accurate prediction of bankrupts and non-bankrupts. The total effectiveness is calculated as (Korol, 2021):

$$S = \left(1 - \frac{D_1 + D_2}{BR + NBR}\right) \cdot 100\% \tag{1}$$

a type I error is computed as:

$$E_1 = \frac{D_1}{BR} \cdot 100\% \tag{2}$$

and a type II error is calculated as:

$$E_2 = \frac{D_2}{NBR} \cdot 100\% \tag{3}$$

The area under the ROC curve (AUC) is suitable for evaluating a method's performance in imbalanced datasets, as it is insensitive to misclassification costs and imbalanced distributions, with a higher AUC value indicating better classifier performance (Zelenkov & Volodarskiy, 2021). AUC measures the probability that a model will rank a randomly chosen positive instance higher than a randomly chosen negative one (Liang et al., 2016). The combination of these four indicators allows for a thorough analysis of the predictive results, taking into account effectiveness among both bankrupts and non-bankrupts.

2.4. Shapley additive explanation

The SHAP method is an approach based on game theory to explain the output of any machine learning model. It was proposed by Lundberg and Lee (2017).

The SHAP value is calculated to provide interpretable prediction results. It also shows the key factors influencing the predictive results, providing more valuable information to identify potential bankruptcies. To make the model interpretable, SHAP uses an additive feature attribution method, and the output model is defined as a linear addition of the input variables (Mangalathu et al., 2020). In view of the fact that machine learning models are considered black boxes, the SHAP summary plot helps to explain the predictions. SHAP is used to interpret each parameter on a global and individual scale (see Section 3). Each point on the graph represents a person, and the set of points constructs the SHAP value of the attribute. The horizontal axis shows the positive and negative correlation between the characteristic variables and the output scores, while the vertical axis is the absolute value ranking of the attribute values (Zhang et al., 2023). Another important aspect is the colour of a given observation. Blue represents a lower value and red represents a higher value. A higher SHAP value means a higher probability of bankruptcy.

3. Results

Among the five most important variables in the prepared models, where feature importance is possible, the most common features were: income, refusal to grant a loan (turndown), having any past payments due (late), the proportion of housing debt to the total monthly payments (house_debt), and the proportion of revolving debt to the total monthly payments (revolving_debt). Figure 1 shows the ranking of the features using RF, AdaBoost, XGBoost, LightGBM, and CatBoost. The most significant variable for RF, XGBoost and CatBoost was the prior refusal of credit. Income was the most significant variable for LightGBM and AdaBoost and the second most significant variable for CatBoost and RF. In the case of selected variables, having a savings account (saving_account) and owning a house (homeownership) turned out to be the least significant in most of the proposed models.

After developing six prediction models using the learning sample, we performed effectiveness analyses of these models on the testing sample. The classification results are provided in Table 2. From the results obtained, out of the six models, LightGBM, CatBoost, XGBoost, and RF perform significantly better than AdaBoost and SVM. LightGBM achieved a higher total accuracy of 0.78 percentage points, a lower type I error of 1.06 percentage points, and a lower type II error of 0.51 percentage points than Catboost. A type I error is considered more costly than a type II error because it can lead to granting a loan to a person who will encounter problems with repayment. The costs of misclassification should minimise the risk of insolvency but also focus on

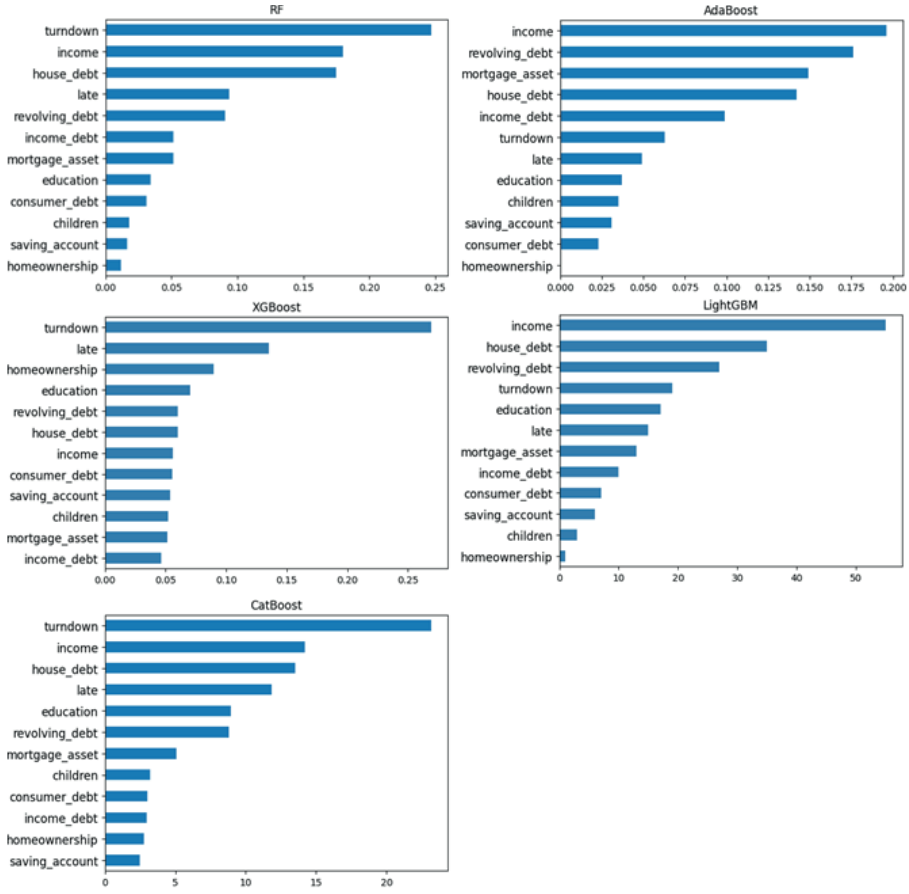

Figure 1. Feature importance

Source: own research.

maximising the number of loans granted, depending on the strategy adopted in this area, because it is profit for banks and financial institutions. The lowest type I error among the proposed models was achieved by LightGBM (21.93%), followed by RF, with a result of 22.46%, and CatBoost (22.99%). The lowest type II error was achieved by XGBoost (27.04%), followed by LightGBM with 28.06%, and then CatBoost (28.57%). The lowest total effectiveness was achieved by AdaBoost (70.76%) and SVM (70.76%). RF is an important alternative to boosting methods, worth verifying in the case of consumer bankruptcy prediction. In terms of AUC, RF and LightGBM showed the best performance, followed by CatBoost and XGBoost.

The study used a small dataset, thus the learning time of the models was not too long. Therefore, it is not required to use this criterion when choos-

Table 2. The results of the effectiveness of models (training and testing sample)

Sample	Model	Type I error (%)	Type II error (%)	Total effectiveness (%)	AUC (%)
Training	SVM	24.55	29.57	72.96	81.73
	Random Forest	23.51	29.30	73.61	81.62
	AdaBoost	27.14	31.14	70.87	78.22
	XGBoost	22.73	28.78	74.27	82.69
	LightGBM	21.30	29.30	74.72	82.61
	CatBoost	24.42	27.22	74.13	82.40
Testing	SVM	25.67	32.65	70.76	77.68
	Random Forest	22.46	31.12	73.11	79.77
	AdaBoost	27.81	30.61	70.76	78.02
	XGBoost	24.60	27.04	73.63	78.63
	LightGBM	21.93	28.06	74.93	79.62
	CatBoost	22.99	28.57	74.15	79.42

Source: own research.

ing an effective model. However, if the dataset was larger, both the indicators showing the effectiveness of the models and the time needed in the model learning process should be considered.

The SHAP summary plots for LightGBM, Catboost, and XGBoost are illustrated in Figures 2, 3, and 4. For the LightGBM model, the income feature provides the highest contribution to prediction, as shown in Figure 2. Consumers who have been turned down in the past (turndown) are more likely to file for bankruptcy. Moreover, a lower proportion of housing debt to total monthly payments (house_debt) leads to a lower risk of bankruptcy.

In Figure 3, for the Catboost model, variables that contribute most to bankruptcy prediction are the refusal to grant a loan (turndown) and the proportion of housing debt to the total monthly payments (house_debt). Additionally, having any past payments due (late) is also one of the most significant factors for the prediction results.

In Figure 4, for the XGBoost model, income, the proportion of housing debt to the total monthly payments (house_debt), and the proportion of housing debt to the value of total assets (mortgage_asset) provide the highest contribution to prediction.

SHAP values can also be used to create an explanation for every observation in the dataset, not only for the global effect presented in the SHAP summary plot. Figures 5, 6, and 7 present explanations of individual predictions

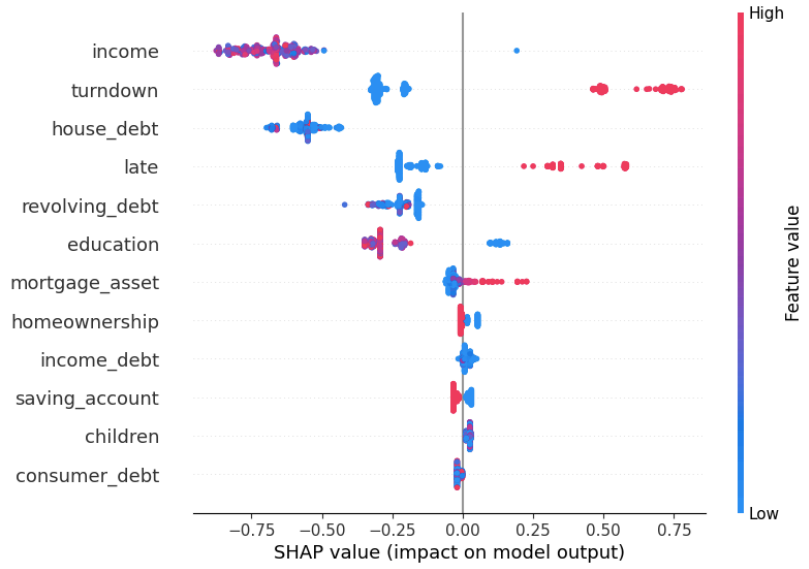

Figure 2. The SHAP summary plot for the LightGBM model illustrates the range and distribution of the impacts of input features

Source: own research.

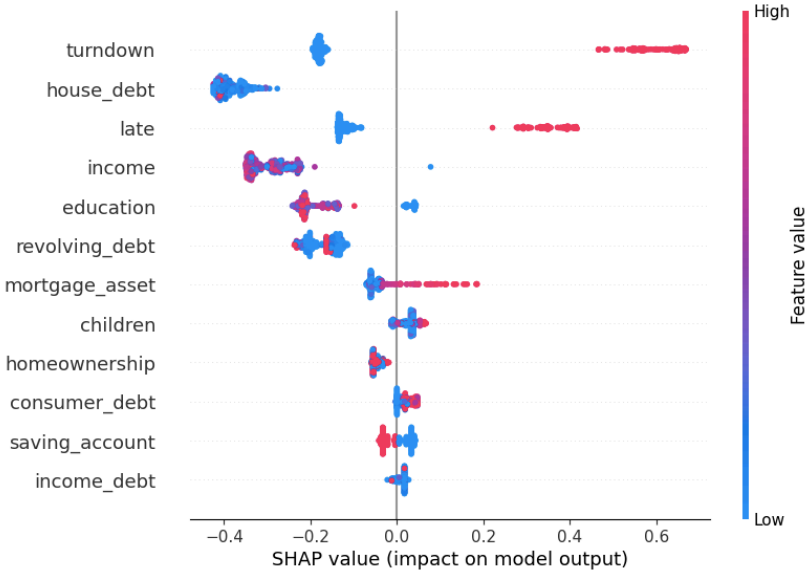

Figure 3. The SHAP summary plot for the Catboost model illustrates the range and distribution of the impacts of input features

Source: own research.

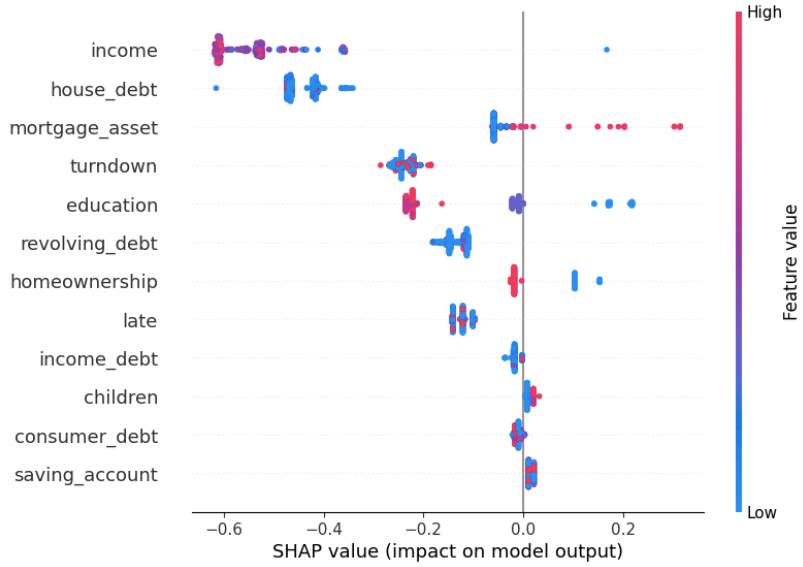

Figure 4. The SHAP summary plot for the XGBoost model illustrates the range and distribution of the impacts of input features

Source: own research.

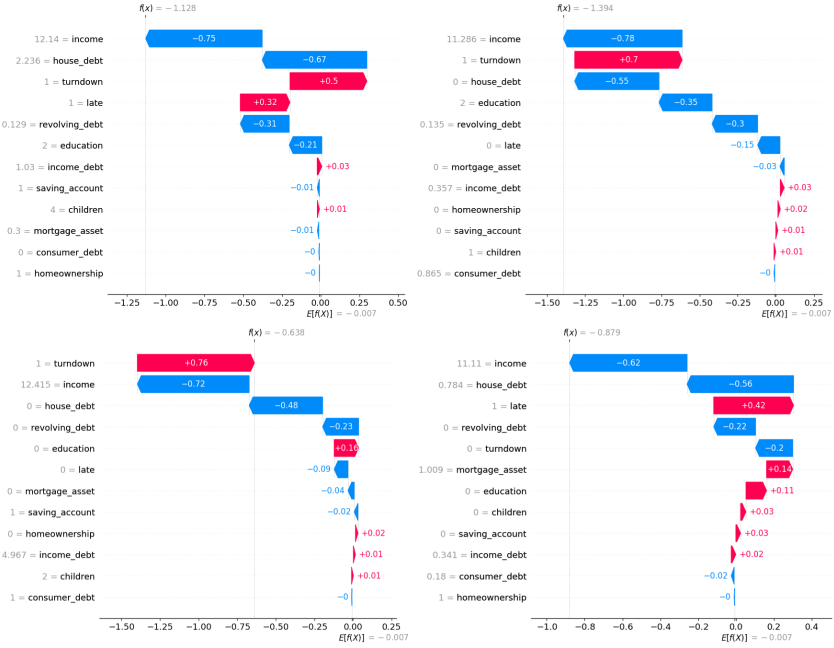

Figure 5. Explanation of individual prediction for the LightGBM model

Source: own research.

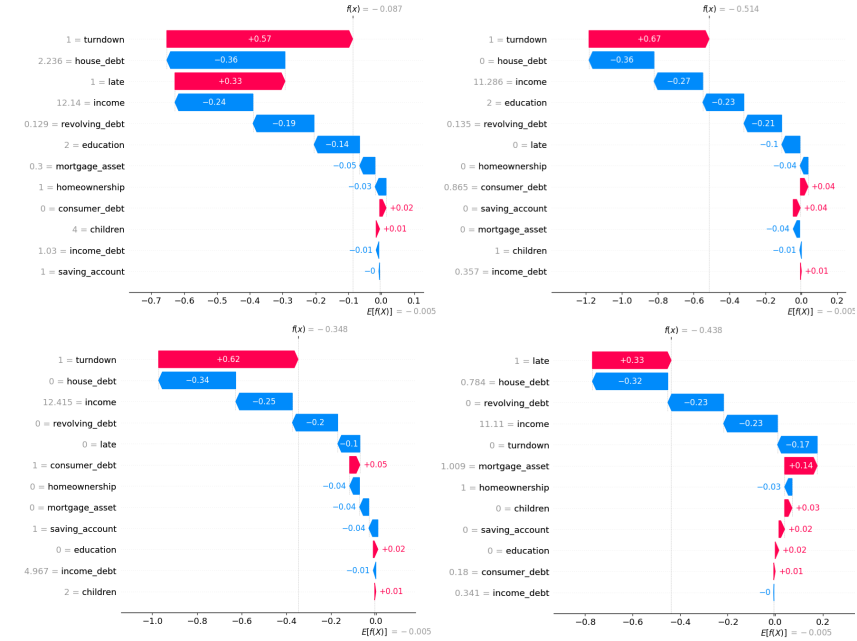

Figure 6. Explanation of individual prediction for the Catboost model

Source: own research.

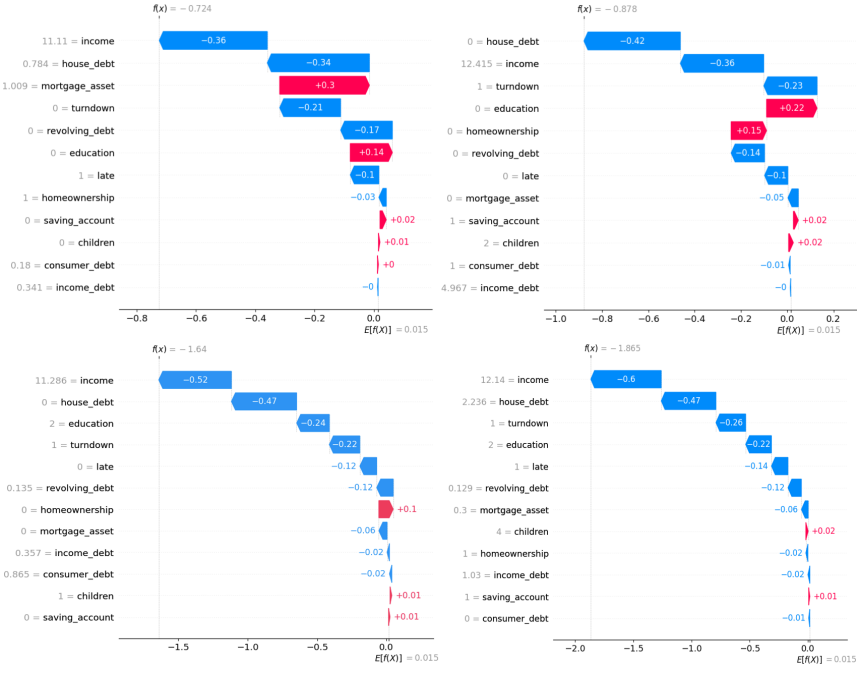

Figure 7. Explanation of individual prediction for the XGBoost model

Source: own research.

for LightGBM, Catboost, and XGBoost for four predictions. The grey values in front of the variables are the values of particular features. The baseline value (E[f(X)]) is displayed below the x-axis and shows the expected value of the model. The value (f(x)) is the model output for each individual, calculated as a sum of the SHAP values for all variables.

4. Discussion

The application of machine learning to financial forecasting is still a relatively new area, but is one worth exploring. The advantage of the effectiveness of machine learning over statistical methods has been confirmed in many studies on both the bankruptcy of enterprises and loan defaults (Garcia, 2022; Machado & Karray, 2022; Son et al., 2019). In comparison to corporate bankruptcy prediction models, it is difficult to compare the effectiveness of consumer bankruptcy prediction models because the literature on this subject contains little research. Existing research mainly focuses on factors affecting consumer bankruptcy rather than predictive models.

Syed Nor et al. (2019) analysed the effectiveness of a decision tree in predicting personal bankruptcy in Malaysia for consumers with terminated or defaulted loans on both an unbalanced and a balanced dataset. The dataset was balanced by the undersampling method. In the case of unbalanced data, despite the higher accuracy (83.29%), by balancing the dataset, accuracy decreased to 70,90%, but specificity (for the minority class) increased from 6.62% to 81.23%, and sensitivity (for the majority class) decreased from 99% to 60.57%. Sensitivity is the probability of the model properly predicting bankrupts (Syed Nor et al., 2019). Despite the higher accuracy, the model for unbalanced data is not effective, due to the large prediction error of the minority class. A more efficient model was presented for a balanced sample. The dataset was obtained from an authorised debt management agency in Malaysia. In the study by Brygała (2022), the results also show that the predictive performance of the logistic regression model based on a balanced dataset is more effective compared to one based on an imbalanced dataset. Two methods of dealing with unbalanced data were used: the undersampling method and the optimal threshold. The research relies on a dataset from the Survey of Consumer Finances from the United States. The total effectiveness of the prediction model on an imbalanced dataset was 95.98%, with a type I error of 99.71% and a type II error of 0%. The total effectiveness of the prediction model on a balanced dataset (undersampling technique) was 69.85%, with a type I error of 29.41% and a type II error of 30.88%. After adjusting the cut-off point to an imbalanced dataset, as one method of dealing with unbal-

anced data (Mihalovič, 2016), the total effectiveness of the model reached 68.99%, with a type I error of 31.18% and a type II error of 31%. Korol (2021) deployed a decision tree, logistic regression, and discriminant analysis to predict personal bankruptcy on a balanced dataset.

The results show that the highest total effectiveness for European households was achieved by logistic regression (92.70%), followed by discriminant analysis (89.60%) and the decision tree (85.60%). For Far-East Asian households, the highest total effectiveness was also achieved by logistic regression (90.10%), followed by discriminant analysis (87.70%) and the decision tree (83.70%). For the same sample, Korol and Fotiadis (2022) compared fuzzy sets, artificial neural networks, and genetic algorithms in forecasting the risk of personal bankruptcy on a balanced dataset. The dataset was also balanced by the undersampling method. The fuzzy sets outperformed artificial neural networks and genetic algorithms. For Taiwanese households, the fuzzy sets are characterised by 90.60% correct classifications, while for European consumers, it amounts to 93.90%. Artificial neural networks and genetic algorithms obtained a total effectiveness of 89.30% for Taiwanese households and 92.90% for Polish households. The research is based on datasets from Poland and Taiwan. Sahiq et al. (2022) also examined the usefulness of logistic regression in forecasting consumer bankruptcy and compared balanced and imbalanced datasets. The dataset was balanced by the SMOTE technique. The total effectiveness of the prediction model on the imbalanced dataset was 84.82%, with sensitivity (for the majority class) of 100%, and specificity (for the minority class) of 0%. The total effectiveness of the prediction model on a balanced dataset was 73.43%, with sensitivity (for the majority class) of 69.50%, and specificity (for the minority class) of 77.35%. The research relies on a dataset from the Debt Management Programme conducted in Malaysia. The comparison of the personal bankruptcy forecasting models is presented in Table 3.

Comparing the effectiveness of the developed models to the effectiveness of models from the literature related to company bankruptcy and defaults, Bragoli et al. (2022) noted that XGBoost performed better in correctly classifying bankrupt firms, but RF and neural networks were better in classifying non-bankrupt firms. In our study, the lowest type I error was achieved by LightGBM and RF, but the lowest type II error by XGBoost and LightGBM. Al Daoud (2019) compared three algorithms: XGBoost, CatBoost, and LightGBM, in two areas: accuracy and CPU runtime. LightGBM proved to be both faster than other methods used and more accurate. Due to the small dataset in our research, time was not a determinant when choosing a model. However, LightGBM proved to be highly effective in our research, demonstrating higher total effectiveness than other models tested. De Castro, Vieira et al. (2019) compared SVM, bagging, AdaBoost, decision trees, logistic regression, and discriminant analysis in predicting default in a residential mortgage programme. The boosting, bagging, and RF algorithms outperformed other meth-

Table 3. The comparison of the personal bankruptcy forecasting models

Authors	Dataset	Method	Total effectiveness (%)	Type I error (%)	Type II error (%)
Syed Nor et al. (2019)	Imbalanced dataset	Logistic regression	83.29	–	–
	Balanced dataset: undersampling		70.90	–	–
Korol (2021)	Balanced dataset: undersampling (Poland)	Logistic regression	92.70	6.20	8.40
		Discriminant analysis	89.60	8.20	12.60
		Decision tree	85.60	15.80	13
	Balanced dataset: undersampling (Taiwan)	Logistic regression	90.10	10.60	9.20
		Discriminant analysis	87.70	13.80	10.80
		Decision tree	83.70	17.40	15.20
Korol and Fotiadis (2022)	Balanced dataset: undersampling (Poland)	Fuzzy logic	93.90	4.80	7.40
		Artificial neural networks	92.90	5.80	8.40
		Genetic algorithms	92.30	5.80	9.60
	Balanced dataset: undersampling (Taiwan)	Fuzzy logic	90.60	7.80	11
		Artificial neural networks	89.30	8.80	12.60
		Genetic algorithms	89.30	8.80	12.60
Brygała (2022)	Imbalanced dataset	Logistic regression	95.98	99.71	0
	Imbalanced dataset: adjusting cut-off point		68.99	31.18	31.00
	Balanced dataset: undersampling		69.85	29.41	30.88
Sahiq et al. (2022)	Imbalanced dataset	Logistic regression	84.82	–	–
	Balanced dataset: SMOTE		73.43	–	–

Source: own research.

ods. Support vector machines were one of the weaker methods compared to RF, bagging, AdaBoost, and decision trees. In our research, SVM achieved the lowest efficiency with AdaBoost among the methods used, also achieving some of the highest type I and II errors. Furthermore, Coşer et al. (2019) compared RF, logistic regression, LightGBM, and XGBoost to predict loan default. The highest results were obtained for random forest. In the case of our

study, RF registered the second lowest type I error but also one of the highest type II errors, which meant that the total effectiveness was not the highest among the models tested.

The dataset was finally divided into 80% of the training sample and 20% of the testing sample, due to the highest efficiency of this division. Previous research (Khare & Sait, 2018; Schonlau & Zou, 2020) has shown that through using a larger training sample, it is possible to obtain higher model efficiency and optimise the size of the training sample. This is especially important for small datasets, a point which was confirmed in our research. It is important to adjust the division of the sample, taking into account the research problem, the proportions between the minority and the majority class, and the size of the dataset.

Conclusions

The application of machine learning models to financial forecasting is still a relatively new area. Moreover, there is a research gap in predicting consumer bankruptcy due to insufficient research in this field. This study is one of the first literary attempts to develop machine learning models in personal bankruptcy prediction. Because machine learning models are considered a black box, the research used SHAP to help interpret and explain model predictions. The use of SHAP offers a meaningful and insightful measure of the importance of each variable in predicting bankruptcy. Increasing the interpretability of models gives the opportunity to use more complex models that may show higher efficiency, but so far, due to less transparency and explainability, they could not be used.

The main objective of this study was to predict personal bankruptcy through machine learning classification algorithms. Six machine learning models (SVM, RF, AdaBoost, XGBoost, LightGBM, and CatBoost) were utilised to predict personal bankruptcy. In summary, the highest total effectiveness was obtained by LightGBM (74.93%), CatBoost (74.15%), followed by XGBoost (73.63%), and RF (73.11%). The lowest type I error was achieved by LightGBM (21.93%), followed by RF, with a result of 22.46%, and CatBoost (22.99%). The lowest type II error was achieved by XGBoost (27.04%), followed by LightGBM, with 28.06%, and then CatBoost (28.57%). The lowest total effectiveness was achieved by AdaBoost (70.76%) and SVM (70.76%). Due to the small dataset in our research, time was not a determinant when choosing a model. However, it is worth noting that CatBoost, followed by LightGBM, are rivals for the fastest method, and XGBoost is slower than these two methods (Dorogush et al., 2018). In the case of financial institutions and banks, where the dataset is

large, this can be one of the more important factors when choosing a method. The models were evaluated for varying ratios between the training and testing samples, but an 80% to 20% split was more effective for this dataset. This is in agreement with Schonlau and Zou (2020), who noted that in a small dataset, a 50% to 50% split might reduce the size of the training sample, while a large dataset will not be affected by such a split. Therefore, it is possible to optimise effectiveness on the same dataset by increasing the training sample.

The authors are aware of the limitations of their research. First of all, only data from the United States were taken into account. Having more bankrupts would also make it possible to predict bankruptcy based on data from a shorter period of time.

In the future, the authors will continue to explore the use of other methods, such as deep learning, to predict personal bankruptcy. Moreover, future studies should also explore different feature selection methods, which can be compared with traditional techniques, various common techniques for unbalanced data, such as undersampling, oversampling, a combination of undersampling and oversampling, class weight, threshold tuning, and different techniques increasing the interpretability of machine learning models.

References

Al Daoud, E. (2019). Comparison between XGBoost, LightGBM and CatBoost using a home credit dataset. *International Journal of Computer and Information Engineering*, *13*(1), 6–10.

Alam, N., Gao, J., & Jones, S. (2021). Corporate failure prediction: An evaluation of deep learning vs discrete hazard models. *Journal of International Financial Markets, Institutions and Money*, *75*, 101455. https://doi.org/10.1016/j.intfin.2021.101455

Alfaro, E., García, N., Gámez, M., & Elizondo, D. (2008). Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks. *Decision Support Systems*, *45*(1), 110–122. https://doi.org/10.1016/j.dss.2007.12.002

Altman, E. I., & Kuehne, B. J. (2016). Credit markets and bubbles: Is the benign credit cycle over? *Economics and Business Review*, *2*(3), 20–31. https://doi.org/10.18559/ebr.2016.3.3

Barboza, F., Basso, L. F. C., & Kimura, H. (2021). New metrics and approaches for predicting bankruptcy. *Communications in Statistics-Simulation and Computation*, *52*(6), 2615–2632. https://doi.org/10.1080/03610918.2021.1910837

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, *83*, 405–417. https://doi.org/10.1016/j.eswa.2017.04.006

Berlemann, M., & Salland, J. (2016). The Joneses' income and debt market participation: Empirical evidence from bank account data. *Economics Letters*, *142*, 6–9. https://doi.org/10.1016/j.econlet.2016.02.030

Bragoli, D., Ferretti, C., Ganugi, P., Marseguerra, G., Mezzogori, D., & Zammori, F. (2022). Machine-learning models for bankruptcy prediction: do industrial variables matter? *Spatial Economic Analysis*, *17*(2), 156–177. https://doi.org/10.1080/17421772.2021.1977377

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Brotcke, L. (2022). Time to assess bias in machine learning models for credit decisions. *Journal of Risk and Financial Management*, *15*(4), 165. https://doi.org/10.3390/jrfm15040165

Brygała, M. (2022). Consumer bankruptcy prediction using balanced and imbalanced data. *Risks*, *10*(2), 24. https://doi.org/10.3390/risks10020024

Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2020). Explainable AI in fintech risk management. *Frontiers in Artificial Intelligence*, *3*, 26. https://doi.org/10.3389/frai.2020.00026

Carmona, P., Dwekat, A., & Mardawi, Z. (2022). No more black boxes! Explaining the predictions of a machine learning XGBoost classifier algorithm in business failure. *Research in International Business and Finance*, *61*, 101649. https://doi.org/10.1016/j.ribaf.2022.101649

Chen, T., & Guestrin, C. (2016, August). *Xgboost: A scalable tree boosting system*. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 785–794. https://doi.org/10.1145/2939672.2939785

CFPB (Consumer Financial Protection Bureau). (2022). *Is a lender allowed to consider my age or where my income comes from when deciding whether to give me a loan?* https://www.consumerfinance.gov/ask-cfpb/is-a-lender-allowed-to-consider-my-age-or-where-my-income-comes-from-when-deciding-whether-to-give-me-a-loan-en-1181/

Coşer, A., Maer-matei, M. M., & Albu, C. (2019). Predictive models for loan default risk assessment. *Economic Computation & Economic Cybernetics Studies & Research*, *53*(2). https://doi.org/10.24818/18423264/53.2.19.09

de Castro Vieira, J. R., Barboza, F., Sobreiro, V. A., & Kimura, H. (2019). Machine learning models for credit analysis improvements: Predicting low-income families' default. *Applied Soft Computing*, *83*, 105640. https://doi.org/10.1016/j.asoc.2019.105640

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). *CatBoost: Gradient boosting with categorical features support*. https://doi.org/10.48550/arXiv.1810.11363

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119–139. https://doi.org/10.1006/jcss.1997.1504

Garcia, J. (2022). Bankruptcy prediction using synthetic sampling. *Machine Learning with Applications*, *9*, 100343. https://doi.org/10.1016/j.mlwa.2022.100343

Georgarakos, D., Haliassos, M., & Pasini, G. (2014). Household debt and social interactions. *The Review of Financial Studies*, *27*(5), 1404–1433. https://doi.org/10.1093/rfs/hhu014

Gramegna, A., & Giudici, P. (2021). SHAP and LIME: An evaluation of discriminative power in credit risk. *Frontiers in Artificial Intelligence*, *4*, 752558. https://doi.org/10.3389/frai.2021.752558

Halim, Z., Shuhidan, S. M., & Sanusi, Z. M. (2021). Corporation financial distress prediction with deep learning: Analysis of public listed companies in Malaysia. *Business Process Management Journal*, *27*4), 1163–1178. https://doi.org/10.1108/bpmj-06-2020-0273

Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: An interdisciplinary review. *Journal of Big Data*, *7*(1), 94. https://doi.org/10.1186/s40537-020-00369-8

Heo, J., & Yang, J. Y. (2014). AdaBoost based bankruptcy forecasting of Korean construction companies. *Applied Soft Computing*, *24*, 494–499. https://doi.org/10.1016/j.asoc.2014.08.009

Jabeur, S. B., Gharib, C., Mefteh-Wali, S., & Arfi, W. B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, *166*, 120658. https://doi.org/10.1016/j.techfore.2021.120658

Jabeur, S. B., Mefteh-Wali, S., & Viviani, J. L. (2021). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Annals of Operations Research*, *334*, 679–699. https://doi.org/10.1007/s10479-021-04187-w

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, *30*.

Khare, N., & Sait, S. Y. (2018). Credit card fraud detection using machine learning models and collating machine learning models. *International Journal of Pure and Applied Mathematics*, *118*(20), 825–838.

Korol, T. (2021). Examining statistical methods in forecasting financial energy of households in Poland and Taiwan. *Energies*, *14*(7), 1821. https://doi.org/10.3390/en14071821

Korol, T., & Fotiadis, A. K. (2022). Implementing artificial intelligence in forecasting the risk of personal bankruptcies in Poland and Taiwan. *Oeconomia Copernicana*, *13*(2), 407. https://doi.org/10.24136/oc.2022.013

Kovacova, M., Kliestik, T., Valaskova, K., Durana, P., & Juhaszova, Z. (2019). Systematic review of variables applied in bankruptcy prediction models of Visegrad group countries. *Oeconomia Copernicana*, *10*(4), 743–772. https://doi.org/10.24136/oc.2019.034

Kovacova, M., & Kliestikova, J. (2017). *Modelling bankruptcy prediction models in Slovak companies*. SHS Web of Conferences, vol. 39, p. 01013. EDP Sciences. https://doi.org/10.1051/shsconf/20173901013

Le, T., Lee, M. Y., Park, J. R., & Baik, S. W. (2018). Oversampling techniques for bankruptcy prediction: Novel features from a transaction dataset. *Symmetry*, *10*(4), 79. https://doi.org/10.3390/sym10040079

Letza, S. R., Kalupa, Ł., & Kowalski, T. (2003). Predicting corporate failure: How useful are multi-discriminant analysis models? *Economics and Business Review*, *3*(2), 5–11. https://doi.org/10.18559/ebr.2003.2.494

Liang, D., Lu, C. C., Tsai, C. F., & Shih, G. A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European*

Journal of Operational Research, *252*(2), 561–572. https://doi.org/10.1016/j.ejor.2016.01.012

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*.

Machado, M. R., & Karray, S. (2022). Assessing credit risk of commercial customers using hybrid machine learning algorithms. *Expert Systems with Applications*, *200*, 116889. https://doi.org/10.1016/j.eswa.2022.116889

Mangalathu, S., Hwang, S. H., & Jeon, J. S. (2020). Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Engineering Structures*, *219*, 110927. https://doi.org/10.1016/j.engstruct.2020.110927

Mihalovič, M. (2016). Performance comparison of multiple discriminant analysis and logit models in bankruptcy prediction. *Economics & Sociology*, *9*(4). https://doi.org/10.14254/2071-789x.2016/9-4/6

Mo, H., Sun, H., Liu, J., & Wei, S. (2019). Developing window behavior models for residential buildings using XGBoost algorithm. *Energy and Buildings*, *205*, 109564. https://doi.org/10.1016/j.enbuild.2019.109564

Papík, M., & Papíková, L. (2023). Impacts of crisis on SME bankruptcy prediction models' performance. *Expert Systems with Applications*, *214*, 119072. https://doi.org/10.1016/j.eswa.2022.119072

Papík, M., Papíková, L., Kajanová, J., & Bečka, M. (2023). *CatBoost: The case of bankruptcy prediction*. International Conference on Business and Technology, pp. 3–17. Springer.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, *31*.

Saarela, M., & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. *SN Applied Sciences*, *3*, 272. https://doi.org/10.1007/s42452-021-04148-9

Sahiq, A. N. M., Ismail, S., Nor, S. H. S., Ul-Saufie, A. Z., & Yaacob, W. F. W. (2022, September). *Application of logistic regression model on imbalanced data in personal bankruptcy prediction*. 2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS) (pp. 120–125). IEEE. https://doi.org/10.1109/aidas56890.2022.9918779

Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal*, *20*(1), 3–29. https://doi.org/10.1177/1536867x20909688

Shi, S., Tse, R., Luo, W., D'Addona, S., & Pau, G. (2022). Machine learning-driven credit risk: A systemic review. *Neural Computing and Applications*, *34*(17), 14327–14339. https://doi.org/10.1007/s00521-022-07472-2

Son, H., Hyun, C., Phan, D., & Hwang, H. J. (2019). Data analytic approach for bankruptcy prediction. *Expert Systems with Applications*, *138*, 112816. https://doi.org/10.1016/j.eswa.2019.07.033

Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, *69*, 135–146. https://doi.org/10.1016/j.indmarman.2017.12.019

Syed Nor, S. H., Ismail, S., & Yap, B. W. (2019). Personal bankruptcy prediction using decision tree model. *Journal of Economics, Finance and Administrative Science*, *24*(47), 157–170. https://doi.org/10.1108/jefas-08-2018-0076

Wang, D. N., Li, L., & Zhao, D. (2022). Corporate finance risk prediction based on LightGBM. *Information Sciences*, *602*, 259–268. https://doi.org/10.1016/j.ins.2022.04.058

Wu, D. J., Feng, T., Naehrig, M., & Lauter, K. E. (2016). Privately evaluating decision trees and random forests. *Proceedings on Privacy Enhancing Technologies*, (4), 335–355. https://doi.org/10.1515/popets-2016-0043

Yen, S. J., & Lee, Y. S. (2009). Cluster-based under-sampling approaches for imbalanced data distributions. *Expert Systems with Applications*, *36*(3), 5718–5727. https://doi.org/10.1016/j.eswa.2008.06.108

Zelenkov, Y., & Volodarskiy, N. (2021). Bankruptcy prediction on the base of the unbalanced data using multi-objective selection of classifiers. *Expert Systems with Applications*, *185*, 115559. https://doi.org/10.1016/j.eswa.2021.115559

Zhang, L., Wang, J., & Liu, Z. (2023). What should lenders be more concerned about? Developing a profit-driven loan default prediction model. *Expert Systems with Applications*, *213*, 118938. https://doi.org/10.1016/j.eswa.2022.118938

sciendo Economics and Business Review https://www.ebr.edu.pl Vol. 10 (2), 2024: 143–164 https://doi.org/10.18559/ebr.2024.2.1180

Enhancing garbage fee compliance: Insights from a Slovak municipality

Anetta Caplanova[1] Eva Sirakovova[2] Estera Szakadatova[3]

Abstract

Tax avoidance and tax evasion remain critical challenges for central or local governments and municipalities. This non-compliance also represents an ethical issue since individuals who benefit from publicly provided services do not contribute to their financing as they are legally required. The study aimed to test whether the use of behavioural interventions would reduce the number of non-payers of the garbage collection fee in the city of Hlohovec, Slovakia. The experiment was carried out by distributing leaflets to households with permanent residence in Hlohovec. The subjects of the experiment were randomly divided into three groups. Households in the control group (number of households is 1,718) did not receive any leaflets, households in the first intervention group (number of households is 1,721) received a leaflet containing a social norm, and households in the second intervention group (number of households is 1,625) received a leaflet containing a deterrent message. Ordinary least squares (OLS) regressions were used to analyse the data. The results showed that using a social norm did not significantly increase the number of garbage collection fee payers. At the same time, leaflets with deterrent messages led to an increase in outstanding debt. These results suggest that people reacted unfavorably to the perceived threat (deterrent message), and the opposite effect occurred, i.e. this type of intervention led to an increase in the number of non-payers of garbage collection fee.

Keywords

- behavioural interventions
- nudge
- randomised controlled trial
- garbage fee collection
- local government
- social norm
- deterrence

[1] Department of Economics, Faculty of Economics and Finance,University of Economics in Bratislava, Dolnozemska cesta 1, 85235 Bratislava, Slovakia, anetta.caplanova@euba.sk, https://orcid.org/0000-0003-3767-7042.

[2] Department of Economics, Faculty of Economics and Finance, University of Economics in Bratislava, Dolnozemska cesta 1, 85235 Bratislava, Slovakia, corresponding author: eva.sirakovova@euba.sk, https://orcid.org/0000-0003-1591-4166.

[3] Department of Finance, Faculty of Economics and Finance, University of Economics in Bratislava, Dolnozemska cesta 1, 85235 Bratislava, Slovakia, estera.szakadatova@euba.sk, https://orcid.org/0000-0003-4979-8773.

JEL codes: C93, D10, R28, R51

Article received 23 February 2024, accepted 10 June 2024.

This paper is prepared with the support of the project of the Slovak Research and Development Agency (APVV 18-0435) "Behavioural Interventions in Local Government: Increasing the Efficiency of Local Public Policies", the project VEGA of the Scientific Grant Agency of the Ministry of Education of the Slovak Republic No 1/0545/23 "Enhancing economic processes through insights from behavioural economics (Theoretical and empirical approach)."

Suggested citation: Caplanova, A., Sirakovova, E., & Szakadatova, E. (2024). Enhancing garbage fee compliance: Insights from a Slovak municipality. *Economics and Business Review*, *10*(2), 143–164. https://doi.org/10.18559/ebr.2024.2.1180

Introduction

In recent years, the use of behavioural interventions has gained popularity in the field of public policy, and governments have increasingly applied behavioural insights to enhance the effectiveness of public policies and their benefits to society (OECD, 2017). Behavioural interventions implemented in public policies rely primarily on nudges and choice architecture to guide people's actions while preserving their freedom of choice.

Tax avoidance and evasion remain critical challenges for central and local governments and municipalities, despite their being illegal. Additionally, lower budget revenues are collected at the central or local levels, leading to fewer resources to finance the provision of public goods and services.

Traditionally, efforts to address tax non-compliance have focused on enhanced enforcement and tax audits. These measures rely on fines and other forms of penalties as deterrents to tax evasion. However, taxpayers' behaviour is significantly influenced by behavioural factors, such as social norms, sense of justice, and the simplicity and comprehensibility of the tax system.

Incorporating behavioural interventions in tax/fee collection systems offers a promising avenue for improving the effectiveness of these processes and addressing the challenge of non-compliance. Policymakers can achieve better outcomes with minimal cost and effort by understanding the behavioural factors that influence people's decisions and designing interventions that nudge them toward desired behaviours.

This paper is based on the results of a randomized controlled trial (RCT) conducted in 2021 in the Western Slovakian city of Hlohovec, which was aimed at improving compliance with payment of the garbage collection fee. The payment of this fee was identified as a pressing problem by the municipality, as non-compliance rates were increasing and were the highest among locally imposed fees and taxes.

Households in Hlohovec were randomly allocated to one of two treatment groups or a control group based on the first letter of the household head's surname in alphabetical order. Households in the treatment groups received a modified version of the invoice along with either a leaflet emphasising social norms or a deterrence leaflet, while the control group received an unmodified invoice. The experiment was designed to evaluate the effectiveness of these two types of behavioural interventions for improving compliance with the garbage collection fee.

The study emphasises the importance of a well-designed and implemented methodology within the local governance, and the need to adhere to relevant legislative frameworks and the specific competencies of the corresponding level of government. However, our results indicate that behavioural interventions based on social norms and deterrence did not have a positive influence on households' behaviour. Thus, our study emphasises the need for a more in-depth exploration of the factors that shape households' compliance behaviour.

The structure of the paper follows a logical sequence. Section 1 provides an overview of the relevant literature on behavioural interventions in tax/fee collection, which provides the context for the subsequent sections. Section 2 describes the context of the experimental design and the implemented intervention. Section 3 provides details on the methodology used to evaluate the effectiveness of the intervention. Section 4 presents the results. Finally, the paper ends with concluding remarks summarising the study findings and their implications for future research and policy.

1. Literature review

Existing behavioural interventions aiming to increase tax compliance and payment of fees commonly involve sending taxpayers modified payment reminders or invoices. In this kind of experiment, subjects in the treatment groups receive modified documents, and those in the control group get unmodified documents or placebo treatments.

One of the most common forms of behavioural interventions in this area is the use of social-norm letters, which aim to influence taxpayers' percep-

tions of tax compliance by pointing to the compliant behaviour of other citizens (Castro & Scartascini, 2015; Dell'Anno, 2009; Grabowski & Korczak, 2023; Lyulyov et al., 2023). Individuals who receive such letters are expected to adjust their perception of tax compliance upwards and, consequently, increase their own compliance. However, Castro and Scartascini (2015) point out that social-norm interventions may not be effective for those individuals whose non-compliance with tax obligations is consistent with their attitudes and beliefs. In such cases, individuals may even change their behaviour in the opposite direction, which then may lead to a boomerang effect.

Several field experiments have been conducted to test the effectiveness of different behavioural interventions with regard to boosting tax compliance. In 1995 and 1996, in Minnesota, USA, tax payment reminders highlighting social norms were used to increase tax compliance. The results showed a statistically significant increase in tax collected from subjects in the intervention group compared to those in the control group (Coleman, 2007).

To enhance timely personal income tax payments, Hernandez et al. (2017) conducted a field experiment in Poland, finding that alterations in the tax form wording significantly boosted payment rates. Negatively framed letters proved to be more effective than positively framed ones. Hallsworth et al. (2017) found that framing reminders with social norms increased tax compliance in large-scale experiments. In contrast, reminders emphasizing budget shortfalls due to lower tax inflows did not yield better outcomes. Iyer et al. (2010) noted a significant increase in the reported tax base when firms were informed about penalties for non-compliance or were made aware of the risk of detection.

In contrast, John and Blume (2018) showed that a descriptive social norm message did not prompt households to pay local taxes on time. Schultz et al. (2007) documented mixed success in changing behaviour using messages, with descriptive normative messages strengthening unwanted behaviour in particular contexts. Del Carpio's (2013) field experiment in Peru revealed that disclosing compliance information significantly increased compliance, and payment reminders had also a positive impact. Similarly, Schächtele et al. (2023) demonstrated that a visually redesigned tax bill highlighting public services increased payment rates among tax non-payers by over 20%. Castro et al. (2022) tested three message types for rental income tax compliance, with detection-focused messages proving to be most successful for boosting tax revenues. However, messages appealing to social norms or altruism had limited direct effects on compliance.

A study by Bobek et al. (2012) looked at the use of social norms as a non-monetary stimulus. The aim of their research was to study the influence of social norms on tax compliance. The authors tested the direct and indirect effects of social norms by simulating a compliance scenario involving 174 experienced taxpayers. The results of the path analysis showed that personal and subjective

norms directly affect tax compliance decisions, while injunctive and descriptive norms exert an indirect influence. This suggests that social norms play an important role in shaping tax compliance behaviour, both directly and indirectly.

The significance of moral principles arising from compliance with social norms was also analysed by Dulleck et al. (2016). The authors studied the influence of psychological stress induced by the potential violation of social norms in the context of tax compliance. The results of their laboratory experiments revealed a positive relationship between psychological stress and tax compliance and highlighted the importance of moral sentiments for tax compliance.

Kleven et al. (2011) stressed the significance of deterrence, noting a large effect on compliance in their research in Denmark. In contrast, Ariel's (2012) findings in the context of corporate tax compliance did not support the deterrence theory, as messages containing deterrence did not have a statistically significant positive effect on tax collection.

Between 2012–2013, Bott et al. (2017) conducted experiments in Norway to increase tax collection from income earned abroad. Nudge letters, particularly those emphasising social norms and the tax office's knowledge of individual income from abroad, led to the highest increase in tax compliance.

Castro and Scartascini (2015) analysed efforts to increase compliance with property tax payments in Argentina. Letters sent to deter taxpayers from tax evasion had a positive, statistically significant effect on compliance. However, letters containing information on the degree of tax compliance of other people living in the municipality or how the tax revenue was to be used did not have a statistically significant impact. Kettle et al. (2016) focused on non-compliers in Guatemala. They found out that letters prompting income tax declaration, providing payment information, including deterrence messages about tax audits, and emphasizing social norms increased most compliance with personal and corporate income taxes.

A field experiment was conducted to investigate the impact of increasing the salience of public disclosure of penalties on the tax compliance behaviour of firms. The findings revealed that messages highlighting the potential for incarceration due to tax evasion were highly effective in increasing tax compliance (Holz et al., 2022). In comparing the standard enforcement regime, involving the threat of outstanding debt being handed over to an Enforcement Agency with nudges comprising letters reminding tax delinquents to settle their tax dues, Andersson et al. (2023) found that the nudge had an effect of approximately 7 percentage points. On the other hand, the standard enforcement regime led to an increase in outstanding debt payments by more than 9 percentage points. Consequently, the effect of the standard enforcement regime was only marginally higher than the effect of the nudge.

In a study conducted in the City of Philadelphia for the calendar year 2015, Chirico et al. (2019) ran an experiment targeting taxpayers in arrears with regard to property tax. Seven alternative formulations of reminder letters were

used, and those threatening economic sanctions for non-compliance were found to be the most effective. However, despite this effectiveness, no evidence indicated improved tax compliance behaviour in 2016 for those who received reminders.

Anderson (2017) investigated the impact of a postcard "nudge" on income tax filers in Nebraska, which aimed to encourage self-reporting of tax liabilities. The nudge more than doubled the likelihood of tax reporting and nearly doubled the revenue collected; however, the overall tax reporting rate remained low, suggesting that an information nudge alone may not substantially change tax reporting behaviour. Simplifying communication from the tax administration consistently improved tax compliance in Belgium, prompting timely payments and encouraging both late filers and payers to comply more swiftly, with positive relationships noted between the use of deterrence messages and tax payments (De Neve et al., 2021).

Latvia implemented a behavioural experiment (Jamison et al., 2021) demonstrating the effectiveness of behaviourally informed tax communication in improving compliance, particularly with the use of a stricter-toned message emphasising deliberate active choice when targeting partially or fully self-employed individuals who delayed or failed to submit returns.

In Belgium, Luts and van Roy (2019) sent reminders to non-compliers with income tax payments, and those containing information about possible sanctions for non-compliance proved most effective. Doshi's (2017) UK intervention motivated timely tax payments by using letters with behavioural nudges, with the letter emphasising individuals as belonging to a non-paying minority having the largest positive impact on tax collection.

Giarrizzo (2012) conducted a detailed comparative study of monetary and non-monetary incentives for fulfilling tax obligations. She found that paying taxes is often a resisted action, with only a few people predisposed to do so voluntarily, and this reluctance diminishes when people perceive inefficiencies in the system. In such cases, controls and penalties become insufficient and require the introduction of parallel incentives. This study highlights the effectiveness of positive incentives and points to the need to replace the traditional control scheme and penalties with a system that includes both punishments and rewards.

The literature review indicates that behavioural interventions are an inexpensive and relatively simple way to motivate people to comply with their tax obligations. However, the effectiveness of different types of nudges differs. As it has been shown, the effect of nudges based on normative and descriptive social norms was different, as was the effect of those based on social norms and deterrence messages. Also, other factors, such as attitudes toward tax avoidance or cultural factors, affect the impact of behavioural nudges.

The aim of this study is to test whether the use of behavioural intervention would lead to the desired behaviour, specifically, whether the use of be-

havioural intervention would result in a reduction in the number of non-payers of the garbage collection fee in the city of Hlohovec, Slovakia. The results presented in this article contribute to the literature with the experience of implementing behavioural intervention focused on increasing garbage fee compliance in a municipality in Western Slovakia.

2. The context of the experiment

The city of Hlohovec is located in Western Slovakia and has slightly more than 20,000 inhabitants. The municipality identified a problem of non-compliance with the payment of fees for garbage collection, which has been higher compared to compliance with the payment of other municipal taxes and fees. To reduce the share of non-compliers, a behavioural intervention was implemented in 2021. The municipality sends a reminder letter to those residents who do not pay the fee or local tax within the pre-determined period. If the entity does not fulfil its payment obligation even after the late payment deadline, the municipal office refers the case to its legal department, which starts the enforcement process.

The payment procedure for the garbage collection fee is clearly specified. If the fee is below 25€, it must be paid within 15 days of the date of the decision on the payment. If the fee is higher than 25€, the subject can pay the fee in four equal instalments (the first instalment within 15 days and then the remaining three instalments every three months throughout the year, with the last instalment to be paid by the end of November). Despite the possibility of paying the fee in instalments, most subjects decide to pay the fee in one instalment. This can be due to the relatively low size of the nominal fee.[4] However, the garbage collection fee differs among households living in apartment buildings and those living in individual-family houses, and households living in apartments pay lower fees.

According to the Act (2004), if a taxpayer fails to comply and does not demonstrate the payment of the fee within the specified period, the tax or fee administrator may proceed with enforcement proceedings. This involves filing a proposal for enforcement proceedings, which requests a court or executor to enforce the unpaid fee from the debtor. Upon approval of the enforcement proceedings, the executor may proceed to recover the outstanding amount. This can involve various measures, such as confiscating the property of the debtor or freezing their bank accounts. If the court approves the enforcement

[4] The average annual garbage collection fee among the households in the sample was approximately 79 euros.

proceedings, the executor may confiscate the debtor's assets to recover the outstanding amount, including any interest and costs associated with the enforcement. However, specific procedures may be adjusted depending on the type of the fee and local regulation.

3. Methods

Leaflets[5] containing behavioural nudges were sent to citizens by regular mail, together with the invoice for the annual fee for garbage collection. Thus, the subjects who use electronic communication with the municipal office were excluded from the sample. Thus, there may be some selection bias in the sample, given that younger or more educated people may use electronic communication more frequently. Business entities were not included in the intervention.

The subjects were assigned to one of the two intervention groups or a control group. Randomization was performed based on the alphabetic order of the subject's surname. The information system used by the municipality assigned a registration number to each taxpayer and this way to anonymise the subjects. The intervention group that received the social norm leaflet comprised 1,721 households, and the intervention group that received the deterrence leaflet consisted of 1,625 households.

Households in the control group received no leaflet, only a standard invoice for the annual garbage collection. Households in the first treatment group received a leaflet focusing on social norms. This leaflet emphasised that 9 out of 10 citizens of Hlohovec paid this fee on time and urged the recipient to join the paying majority and pay the fee within a pre-determined payment period. Even though the share of non-compliers increased during the COVID-19 period of 2020 and subsequently also in 2021 (see Table 1), it based this nudge on the rate of non-compliers for natural persons (households), which were the subjects of the intervention (i.e. pre-COVID) in 2019. It is believed that the increase in non-compliers during the period when the pandemic affected life and the economy was only temporary. Households in the second treatment group were sent a deterrence leaflet highlighting that if a citizen did not pay the garbage collection fee on time, enforcement proceedings could follow, corresponding to the actual practice applied in the municipality. In order to strengthen the effect of the leaflets, they also included pictures. The format of the invoice for the garbage collection was the same as in the previous years.

[5] The English version of leaflets used is provided in Appendix B.

In line with existing behavioural research (Athief & Ma'ruf, 2023; Castro & Scartascini, 2015; Dell'Anno, 2009 ; Govender & David, 2023; Hamdan et al., 2023; Nikolajenko-Skarbalė & Viederytė-Žilienė, 2023), in this experiment, a leaflet with a behavioural component aimed to influence the perception of compliance with the garbage collection fee payment by other residents of the city and the awareness of the repercussions in case of non-payment.

The experiment was carried out between spring and autumn 2021, which represents the payment period for the garbage collection fee. The data on compliance with the fee payment were collected at the end of 2021.

The data used for the estimation were obtained from a randomised controlled trial in 2021 in cooperation with the city of Hlohovec, Slovakia.

Since 2018, the number of non-compliers with the payment of the municipal waste fee has been increasing among natural persons and legal entities. While in 2018, the share of non-compliers represented approximately 7.3% for natural persons and approximately 5% for legal entities, by 2020, the share of non-compliers among natural persons had more than quadrupled to over 29%, and among legal entities, it had increased more than 8 times and reached almost 45%. Additional information on the number of taxpayers and non-compliers with paying other taxes and fees can be found in Appendix A, Table A1.

The data shows that in the pre-intervention period, the average amount of tax to be paid by Hlohovec citizens in our sample was about 71 euros, with the maximum amount of tax to be paid being about 280 euros (see Table 1).

Table 1. Summary statistics of the sample

	Mean	Standard deviation	Minimum value	Maximum value
Pre-intervention period				
Amount of tax to be paid	71.3239	41.0261	0	279.43
Amount of tax paid	65.8510	42.9821	0	279.43
Net outstanding debt	5.4728	21.7273	–23.7	216.3
Living in high-density area	0.8348	0.3714	0	1
Post-intervention period				
Amount of tax to be paid	79.2933	46.0452	0	311.56
Amount of tax paid	62.4242	49.3659	0	311.56
Net outstanding debt	16.7917	36.5535	–155.78	311.56
Living in high-density area	0.8352	0.3710	0	1

Note: The negative value of the net outstanding debt indicates that the payment made by a citizen to the city council was higher than the outstanding fee.

Source: own calculations based on the data provided by the municipality office of the city Hlohovec.

In the post-intervention period, the average amount of tax to be paid increased by about 8 euros, while the maximum amount of tax to be paid by citizens in our sample increased by about 32 euros. The summary statistics also show that compared to the pre-intervention period, the average amount of tax paid by citizens decreased in the post-intervention period, while the average net outstanding debt increased on average threefold in the post-intervention period. The data also shows that the maximum amount of net outstanding debt in the post-intervention period equalled the maximum amount of tax to be paid by citizens (however, in the pre-intervention period, the maximum amount of net outstanding debt was lower than the maximum amount of tax to be paid).

The data sample includes households living in Hlohovec in 2020 (i.e. the pre-intervention period) and in 2021 (i.e. the post-intervention period). The entire data sample consists of 18,403 observations (both for 2020 and 2021), which includes observations in both the pre- and post-intervention periods. However, the sample size used for the estimation is smaller than the entire sample of the households included in the experiment, since those households that moved to/from the city of Hlohovec between 2020–2021 were excluded from the sample. Moreover, some households have permanent residency in other locations and were excluded from the sample.

Just over 1,700 households were randomly assigned to treatment group 1, which received a social norm leaflet with the social norm intervention, and over 1,600 households were randomly assigned to treatment group 2, which received a deterrence leaflet. Approximately 1,700 households were randomly allocated to the control group (see Table 2).

The data shows that in the pre-intervention period, the average outstanding debt of a household in the treatment and control groups was between 5–6 euros. The average outstanding debt of a household in treatment group 1 was 5.91 euros, and for households in treatment group 2, it was 4.99 euros. Households in the control group had an average debt of 5.50 euros. However, in the post-intervention period, when compared with the pre-intervention period, the mean outstanding debt increased in all groups. The average debt increased most among households in treatment group 1 (i.e. about 1.65 percentage points) and the least among households in treatment group 2 (i.e. about 1.35 percentage points). It also increased by 1.58 percentage points among households in the control group during the same period.

Since the average treatment effect on the treated (ATT) and the average treatment effect (ATE) are equal in an RCT, the OLS estimator is consistent for both. In addition, given the substantial sample size of more than 3,000 observations, we used the ordinary least squares (OLS) estimation method, which also offers a straightforward interpretation of the coefficients and facilitates an intuitive understanding of the relationship between the independent and dependent variables:

$$Complier_i = \alpha_0 + \beta_1 \, Social\ norm_i + \beta_2 \, Deterrence_i + \varepsilon_i \quad (1)$$

where the dependent variable $Complier_i$ is a dummy variable that indicates whether a household i is a complier or not, $Social\ norm_i$ is a binary variable taking value 1 if household i was allocated in the treatment group 1 and received the leaflet with the social norm, $Deterrence_i$ is a binary variable taking value 1 is a binary variable taking value 1 if a household i was allocated to the treatment group 2 and received the deterrence leaflet and ε_i is the robust standard error[6].

The coefficients of interest are β_1 and β_2 represent the estimate of the impact of the social norm and deterrence leaflet interventions on compliance with the garbage collection fee payment.

Secondly, the study estimated the impact of the social norm and deterrence treatments on the amount of net outstanding debt on the garbage collection fee. To estimate the causal relationship between the treatments and the amount of net outstanding debt, an ordinary least squares estimation method with the following specification was employed:

$$Debt_i = \alpha_1 + \gamma_1 \, Social\ norm_i + \gamma_2 \, Deterrence_i + \theta_i \quad (2)$$

where the dependent variable $Debt_i$ is the natural logarithm of the net outstanding debt of household i. The remaining variables are defined in the same way as in the regression specification (1), θ_i is the robust standard error.

The coefficients of interest are γ_1 and γ_2 which represent the estimate of the impact of the social norm and deterrence leaflet interventions on the amount of net outstanding debt on the garbage collection fee.

This assumption is reasonable, as the treatment assignment was random, and households in the intervention and control groups had similar characteristics in terms of demographics, income, and education level.

4. Results and discussion

The data shows that the share of compliers with the garbage collection fee payment declined between experimental years 2020 and 2021, both in the treatment and control groups (see Appendix A, Table 2A). The proportion of compliers declined the greatest among households in treatment

[6] Robust standard errors adjust standard errors to accommodate potential misspecifications in the variance structure of the error term within the classical linear regression framework, e.g., to account for clustering of observations. Therefore, robust standard errors enhance the validity of regression analysis in empirical research.

group 1, which received the social norm leaflet (by 16.2 percentage points). Table 2 presents the results of the balance test. The results show that the estimated coefficients of the covariates, or background characteristics, and the outcome variables are statistically insignificant. This suggests that the background characteristics are balanced along the pre-intervention outcomes. This implies that the allocation to treatment and to control groups is as good as randomly assigned.

Table 2. The balance in covariates along the pre-intervention outcomes

	Treatment group 1	Treatment group 2
Compliance	–0.0195 (0.0730)	–0.0292 (0.0733)
Tax to be paid	0.1622 (1.6367)	0.6506 (1.6519)
Amount of tax paid	–0.8582 (1.7441)	1.9181 (1.7598)
Net outstanding debt	1.1455 (1.1841)	–1.2132 (1.1952)
High population density area	–0.0049 (0.0137)	0.0039 (0.0138)

Standard errors are in parentheses.
***, **, * statistically significant at 1%, 5% and 10% level of significance.

Source: own elaboration.

Table 3 presents the estimation results of the effect of the social norm and deterrence leaflets on compliance with garbage collection fee payment. The results show that the social norm leaflet had a positive but statistically insignificant effect on compliance, increasing it by approximately 0.67 percentage points. On the other hand, the deterrence leaflet had a significant negative effect, decreasing compliance with the payment of the garbage collection fee by 2.58 percentage points among households in the treatment group receiving the deterrence leaflet. This suggests that the deterrence letter did not help increase the garbage collection fee payment rate in the city of Hlohovec.

Table 4 presents the analysis of the effect of the social norm and deterrence leaflet on the amount of net outstanding debt on the garbage collection fee. The estimates show that the social norm intervention had a negative but statistically insignificant impact on the net outstanding debt, reducing it by approximately 0.93 percentage points. In contrast, the deterrence leaflet had a statistically significant positive effect on the net outstanding debt, increasing it by approximately 1.33 percentage points.

Table 3. The effect of individual treatments on compliance with the garbage collection fee payment

Variables	(1) Col 1
Treatment 1 (social norm leaflet)	0.0068
	(0.0084)
Treatment 2 (deterrence leaflet)	–0.0258**
	(0.0085)
Constant	0.7899***
	(0.0059)
Number of observations, *N*	10,128

Standard errors are in parentheses.
***, **, * statistically significant at 1%, 5% and 10% level of significance.

Source: own elaboration.

Table 4. The effect of individual treatments on outstanding debt of the garbage collection fee

Variables	(1) Col 1
Treatment 1 (social norm leaflet)	–0.0093
	(0.00669)
Treatment 2 (deterrence leaflet)	0.0133**
	(0.0067)
Constant	0.1260***
	(0.0047)
Number of observations, *N*	10,128

Standard errors are in parentheses.
***, **, * statistically significant at 1%, 5% and 10% level of significance.

Source: own elaboration.

The estimation results do not suggest a positive effect of the implemented behavioural interventions on compliance with the payment of the garbage collection fee among the sample of households. The social norm leaflet did not have a statistically significant impact on compliance. In contrast, the deterrence leaflet led to a decrease in compliance and an increase in net outstanding debt on the garbage collection fee payment. These results show that the behavioural interventions implemented in the city of Hlohovec did not contribute to increasing the payment rate of the garbage collection fee.

To check for the possible effects of spill-over, the study estimated the impact of the interventions on compliance and net outstanding debt among households living in less and more densely populated areas.[7] The results show significant negative effects of the deterrence leaflet on compliance and significant positive effects on net outstanding debt among households living in less populated areas. However, no statistically significant relationships were observed between individual treatments and compliance with the garbage collection fee payment or the value of net outstanding debt among households in high-density areas. Therefore, the observed impact of interventions on households' compliance behaviour is the result of the interventions and not due to possible mutual communication among households.

The results in Table 5, column 1, confirm the findings in Tables 3 and 4. Specifically, the deterrence leaflet significantly negatively affected the garbage collection fee payment rate among households in areas with low population density, leading to an 11.22 percentage point decrease. Additionally, the same intervention resulted in a statistically significant 8.03 percentage point increase in net outstanding debt among these households. In contrast, the results for households in high-density areas (column 2) did not reveal any significant relationship between the treatments and compliance behaviour or net outstanding debt.

The results of our research are different from the findings of most previous studies, which document a positive impact of social norms or deterrence leaflets on compliance with the payment of fees, or taxes. Specifically, in contrast to Coleman (2007), Hallsworth et al. (2017), and Kettle et al. (2016), as a consequence of the deterrence leaflet, we found an increase in non-compliance with the garbage collection fee. Our findings partially align with Ariel (2012), who found no discernible positive influence of the deterrence message, and Castro et al. (2022), who highlighted the lack of a direct positive impact of social norms. Castro and Scartascini (2015) previously emphasized the occurrence of a reverse or boomerang effect in behavioural interventions addressing tax and fee non-compliance, and this is also reflected in our findings regarding the adverse impact of the deterrence message.

This negative impact on compliance may be attributed to people not liking to be threatened, especially in countries with a history of authoritarian regimes and low compliance levels. Negative compliance may also be linked to trust in public institutions, which is generally low in Slovakia. Therefore, it

[7] A highly populated area was defined as one with more than 100 households living on the same street. Streets with more than 100 households are the streets with apartment blocks, therefore, people living in these areas are more likely to be exposed to their neighbours than people living on the streets with a lower number of households. In the sample, there are not any streets, with the number of households close to 100-household threshold.

Table 5. The effect of individual treatments on compliance with the garbage collection fee payment among households in low- and high population density areas

Dependent variable – compliance with garbage collection fee payment	(1) Low population density	(2) High population density
Treatment 1 (social norm leaflet)	0.0195	0.0054
	(0.0214)	(0.0091)
Treatment 2 (deterrence leaflet)	–0.1123***	–0.0071
	(0.0203)	(0.0093)
Constant	0.8087***	0.7860***
	(0.0144)	(0.0065)
Dependent variable – net outstanding debt		
Treatment 1 (social norm leaflet)	–0.0173	–0.0083
	(0.0182)	(0.0071)
Treatment 2 (deterrence leaflet)	0.0803***	–0.00131
	(0.0172)	(0.0072)
Constant	0.1179***	0.1277***
	(0.0122)	(0.0051)
Number of observations, *N*	1,679	8,449

Standard errors are in parentheses.
***, **, * statistically significant at 1%, 5% and 10% level of significance.

Source: own elaboration.

is essential to consider the potentially negative effects of interventions that emphasise the negative consequences of non-compliance.

Moreover, the negative impact of the deterrence leaflet on compliance with the payment of the garbage collection fee may suggest that people react adversely to perceived threats and that deterrence interventions may generate a backfire effect. However, the long-term effects of the intervention would need to be studied further, as the experiment was carried out during the COVID-19 pandemic, which could have affected households' compliance behaviour, due to decreased income. However, these results suggest that policymakers should formulate their interventions positively and motivate citizens to comply with policies using a non-deterrence formulation.

Furthermore, additional data on household characteristics such as income or size could provide insights into the results. Nevertheless, it is assumed that these characteristics remained constant over time, and thus, the estimation results are valid and capture the effect of interventions. In addition,

other factors may influence the results of behavioural interventions, and not all of them can be foreseen in advance or reflected in the intervention design. Moreover, the willingness of public authorities to cooperate in the implementation of behavioural interventions plays a crucial role in testing them.

Conclusions

The study aimed to test whether the use of behavioural interventions would lead to the desired behaviour, specifically, whether the use of behavioural interventions would reduce the number of non-payers of the garbage collection fee in the municipality of Hlohovec, Slovakia. The results of the experiment did not confirm the research assumption that the behavioural interventions based on social norm and deterrence would positively affect households' behaviour, increase the payment rate of the garbage collection fee, and decrease the value of related outstanding debt. On the contrary, the deterrence leaflet had a negative impact on compliance with the payment of the garbage collection fee in the post-intervention period. It led to an increase in the corresponding treatment group's outstanding debt. The social norm intervention did not significantly impact compliance with the garbage collection fee payment and related net outstanding debt.

The negative impact of the deterrence leaflet on compliance with the payment of the garbage collection fee indicates that people may react adversely to perceived threats and that deterrence interventions may have unintended consequences and lead to increased non-compliance. Moreover, the analysis highlights the importance of appropriately designed and implemented methodology and the need to adhere to relevant legislative frameworks and specific competences of the corresponding level of government.

This study provides valuable insights into the effectiveness of behavioural interventions in the context of local government, highlighting the importance of carefully designed and implemented interventions and the need for further exploration of the factors that influence households' compliance behaviour.

Appendix A. The share of non-compliers with the payment of taxes and fees in the city of Hlohovec, Slovakia

Table A1. The share of non-compliers with the payment of individual taxes and fees in the city of Hlohovec

Year	Indicator	Property tax	Dog upkeep tax	Hotel accommodation tax	Vending machine tax	Garbage collection fee (natural persons)	Garbage collection fee (legal entities)
2018	Number of taxpayers	8,581	1,468	10	3	9,778	699
	Number of non-compliers	58	23	1	1	687	35
	Share of non-compliers (%)	0.67	1.57	–	–	7.3	5.01
2019	Number of taxpayers	8,512	1,491	10	3	10,194	706
	Number of non-compliers	66	38	0	0	1172	132
	Share of non-compliers (%)	0.77	2.55	–	–	11.5	18.7
2020	Number of taxpayers	8,575	1,460	5	2	10,137	676
	Number of non-compliers	683	148	2	1	2,980	303
	Share of non-compliers (%)	8	10.14	–	–	29.4	44.82
2021	Number of taxpayers	8,540	1,477	4	2	10,247	683
	Number of non-compliers	386	36	1	1	2,831	351
	Share of non-compliers (%)	4.52	2.43	–	–	27.63	51.39

Source: Data based on information provided by the municipality office of the city of Hlohovec.

Table A2. Proportion of compliers in treatment and control groups across both periods

	(1) Pre-intervention period	(2) Post-intervention period	(3) Number of observations, *N*
Treatment group 1 (social norm leaflet)	0.9111	0.7635	1,721
Treatment group 2 (deterrence leaflet)	0.8929	0.7465	1,625
Control group	0.9191	0.7590	1,718

Source: own computations based on the data obtained from the experiment.

Appendix B. Leaflets sent to households in intervention groups

The Deterrence Leaflet

The Social Norm Leaflet

References

Act. (2004). Act No. 582/2004 Coll. on Local Taxes and Fees. (2004). https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2004/582/20230901.html

Anderson, J. E. (2017). Paying the state use tax: Is a "nudge" enough? *Public Finance Review*, *45*(2), 260–282. https://doi.org/10.1177/1091142115614390

Andersson, H., Engström, P., Nordblom, K., & Wanander, S. (2023). Nudges and threats: Soft vs hard incentives for tax compliance. *Economic Policy*, *38*(116), 771–819. https://doi.org/10.1093/epolic/eiad017

Ariel, B. (2012). Deterrence and moral persuasion effects on corporate tax compliance: Findings from a randomized controlled trial. *Criminology*, *50*(1), 27–69. https://doi.org/10.1111/j.1745-9125.2011.00256.x

Athief, F. H. N., & Ma'ruf, A. (2023). Tracing the asymmetry of religiosity-based loyalty of Islamic bank depositors. *Banks and Bank Systems*, *18*(1), 1–13. https://doi.org/10.21511/bbs.18(1).2023.01

Bobek, D. D., Hageman, A. M., & Kelliher, C. F. (2012). Analyzing the role of social norms in tax compliance behavior. *Journal of Business Ethics*, *115*(3), 451–468. https://doi.org/10.1007/s10551-012-1390-7

Bott, K. M., Cappelen, A. W., Sorensen, E., & Tungodden, B. (2017). *You've got mail: A randomised field experiment on tax evasion*. Working Paper, 2017-051). https://econpapers.repec.org/paper/hkawpaper/2017-051.htm

Castro, J. F., Velásquez, D., Beltrán, A., & Yamada, G. (2022). The direct and indirect effects of messages on tax compliance: Experimental evidence from Peru. *Journal of Economic Behavior & Organization*, *203*, 483–518. https://doi.org/10.1016/j.jebo.2022.09.022

Castro, L., & Scartascini, C. (2015). Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior & Organization*, *116*, 65–82. https://doi.org/10.1016/j.jebo.2015.04.002

Chirico, M., Inman, R. P., Loeffler, C., Macdonald, J. M., & Sieg, H. (2019). Deterring property tax delinquency in Philadelphia. An experimental evaluation of nudge strategies. *National Tax Journal*, *72*(3), 479–506. https://doi.org/10.17310/ntj.2019.3.01

Coleman, S. (2007, November 1). The Minnesota Income tax compliance experiment: Replication of the social norms experiment. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1393292

De Neve, J., Imbert, C., Spinnewijn, J., Tsankova, T., & Luts, M. (2021). How to improve tax compliance? Evidence from population-wide experiments in Belgium. *Journal of Political Economy*, *129*(5), 1425–1463. https://doi.org/10.1086/713096

Del Carpio, L. (2013). *Are the neighbors cheating? Evidence from a social norm experiment on property taxes in Peru (Job Market Paper)*. https://www.semanticscholar.org/paper/Are-the-Neighbors-Cheating-Evidence-from-a-Social-Carpio/77488ffd5e3ca1d44c85287a74bf7d50f5612e4d

Dell'Anno, R. (2009). Tax evasion, tax morale and policy maker's effectiveness. *The Journal of Socio-economics*, *38*(6), 988–997. https://doi.org/10.1016/j.socec.2009.06.005

Doshi, M. (2017, June 20). *How the British government got more citizens to pay their taxes on time*. https://www.bqprime.com/politics/uk-the-nudge-unit-uses-behavioural-science-to-influence-policy-outcomes-such-as-improved-tax-collections-and-pension-enrolment

Dulleck, U., Fooken, J., Newton, C., Ristl, A., Schaffner, M., & Torgler, B. (2016). Tax compliance and psychic costs: Behavioral experimental evidence using a physiological marker. *Journal of Public Economics*, *134*, 9–18. https://doi.org/10.1016/j.jpubeco.2015.12.007

Giarrizzo, V. (2012). Monetary And Non-Monetary Incentives to Boost Tax Payment: A controlled experiment. *Visión De Futuro*, *16*(2). https://revistacientifica.fce.unam.edu.ar/index.php/visiondefuturo/article/view/655/572

Govender, K., & David, E. (2023). Experience and satisfaction: Exploring students' perceptions of private and public higher education services. *Problems and Perspectives in Management*, *21*(2), 371–382. https://doi.org/10.21511/ppm.21(2).2023.36

Grabowski, W., & Korczak, K. (2023). Identifying the envelope wages phenomenon based on LFS and SES data sets—evidence from Poland. *Economics & Sociology*, *16*(4), 130–154. https://doi.org/10.14254/2071-789X.2023/16-4/6

Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, *148*, 14–31. https://doi.org/10.1016/j.jpubeco.2017.02.003

Hamdan, M., Calavia, P. F., & Aminu, N. (2023). Sin stocks in European countries: The influence of wealth and familiarity bias on investment choices. *Investment Management and Financial Innovations*, *20*(2), 256–266. https://doi.org/10.21511/imfi.20(2).2023.22

Hernandez, M., Jamison, J., Korczyc, E., Mazar, N., & Sormani, R. (2017). *Applying behavioral insights to improve tax collection*. World Bank Group. https://elibrary.worldbank.org/doi/abs/10.1596/27528

Holz, J., List, J. A., Zentner, A., Cardoza, M., & Zentner, J. (2022). The $100 million nudge: Increasing tax compliance of firms using a natural field experiment. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4010639

Iyer, G. S., Reckers, P. M., & Sanders, D. (2010). Increasing tax compliance in Washington State: A field experiment. *National Tax Journal*, *63*(1), 7–32. https://doi.org/10.17310/ntj.2010.1.01

Jamison, J. R., Mazar, N., & Sen, I. (2021). Applying behavioral insights to tax compliance: Experimental evidence from Latvia. *Journal of Tax Administration*, *6*(2). https://hdl.handle.net/10871/121984

John, P., & Blume, T. (2018). How best to nudge taxpayers? The impact of message simplification and descriptive social norms on payment rates in a central London local authority. *Journal of Behavioral Public Administration*, *1*(1). https://doi.org/10.30636/jbpa.11.10

Kettle, S., Hernandez, M., Ruda, S., & Sanderson, M. A. (2016). *Behavioral interventions in tax compliance: Evidence from Guatemala*. Policy Research Working Paper, 7690. https://ssrn.com/abstract=2811337

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, *79*(3), 651–692. https://doi.org/10.3982/ecta9113

Luts, M., & van Roy, M. (2019). *Nudging in the context of taxation: How the Belgian FPS finance uses behavioural insights to encourage taxpayers to pay faster*. Intra-European Organisation of Tax Administrations IOTA. https://www.iota-tax.org/sites/default/files/documents/iota_paper_belgium_nudging_final.pdf

Lyulyov, O., Pimonenko, T., Chen, Y., & Kwilinski, A. (2023). Macroeconomic stability of the country: The nexus of institutional and behavioural dimensions. *Economics & Sociology*, *16*(4), 264–288. https://doi.org/10.14254/2071-789X.2023/16-4/13

Nikolajenko-Skarbalė, J., & Viederytė-Žilienė, R. (2023). Transformations of customer loyalty attitude in marketing: Key components of modern loyalty. *Innovative Marketing*, *19*(4), 105–123. https://doi.org/10.21511/im.19(4).2023.09

OECD. (2017). *Behavioural insights and public policy: Lessons from around the world*. OECD Publishing. https://doi.org/10.1787/9789264270480-en

Schächtele, S., Eguino, H., & Roman, S. (2023). Fiscal exchange and tax compliance: Evidence from a field experiment. *Journal of Policy Analysis and Management*, *42*(3), 796–814. https://doi.org/10.1002/pam.22460

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x

sciendo
Economics and Business Review
https://www.ebr.edu.pl
Vol. 10 (2), 2024: 165–186
https://doi.org/10.18559/ebr.2024.2.1255

Growth prospects for the silver economy in the market segment of residential care services provided to dependent elderly people

Rafał Iwański[1]

Abstract

The aim of this study is to characterise the determinants of the development of the silver economy in the field of care services provided in a residential form for dependent elderly persons in Poland. The analysis was carried out on the basis of statistical and financial background data, including those from the Ministry of Family and Social Affairs, the Ministry of Health, OECD, etc. Although the demand for care services will continue to grow in the coming years, the following barriers to the development of this segment of the silver economy can be identified: lack of employees, unattractiveness of monetary gratification, inefficient financing mechanisms, lack of public investment in the development of care facilities, and increasing costs of providing services in all forms.

Keywords

- economics of the elderly
- silver economy
- demographic trends
- labour demand
- welfare programme

JEL codes: H53, I38, J11, J14, J23

Article received 23 November 2023, accepted 30 April 2024.

Suggested citation: Iwański, R. (2024). Growth prospects for the silver economy in the market segment of residential care services provided to dependent elderly people. *Economics and Business Review*, *10*(2), 165–186. https://doi.org/10.18559/ebr.2024.2.1255

[1] Applied Gerontology and Thanatology Group, Institute of Pedagogy, University of Szczecin, ul. Ogińskiego 17, 71-431 Szczecin, Poland, rafal.iwanski@usz.edu.pl, https://orcid.org/0000-0001-6345-9677.

Introduction

Demand for care services will grow in the coming years as a consequence of an ageing population and the diminishing care potential of families. According to estimates, approximately 2 million elderly people in Poland will require the assistance of others in 2022, rising to more than 3 million by 2060. Many societies across the world are facing the challenge of securing care for dependent older people. This article contributes to the discussion on the role of individual segments of the silver economy in constructing a care system for dependent persons in Poland. It fills a research gap, as there is a noticeable lack of studies in this area that adopt an interdisciplinary perspective on the development of the silver economy. The process of population ageing, combined with a decline in the care potential of families, results in demand for care services growing every year. It is not always possible to meet the care needs of an elderly person in the form of a health visitor service, in which case residential forms of care apply.

Care services could be a significant segment of the silver economy in the coming years. While demand for services will grow rapidly, the availability of publicly funded or co-funded services is limited, which may contribute significantly to the development of the private sector. The process of population ageing is global and affects countries on every continent. However, it is worth noting the variation in the progression of the population ageing process. In the case of European countries, including Poland, there has been a steady increase in the proportion of older people in the population for several decades, and demographic forecasts predict that this trend will continue (OECD, 2021b; WHO, 2022). In contrast, in the case of the African continent, for example, a significant number of societies are entering the demographic explosion stage, according to the theory of the first demographic transition. The effect of demographic processes at this stage is dynamic population growth and, for the time being, a slight increase in the percentage of elderly people. However, a change in the reproductive pattern and a gradual shift away from dispersive reproduction towards a conservative model can be expected in the next decades, which will lead to an increase in the proportion of older people in the populations of African countries. Yet in Europe, the ageing population process has not taken place simultaneously in particular countries. Therefore, the degree of ageing of individual populations within Europe itself also varies (Dudley, 1996; Rudakova et al., 2023).

Countries affected by population ageing are taking a variety of measures to secure the care needs of dependent older people based on the social policy model adopted, including the social assistance and health care systems. Comparative studies are being conducted on the performance of different solutions based on available resources (Costa-Font & Courbage, 2011; Poškutė,

& Greve, 2017). Building a long-term care system is a very complex process, and it is based on four main elements: beneficiaries, benefits, providers, and financing (Dintrans, 2020). While each system has its own specific conditions and is based on the use of available resources, the cultural context and social preferences for forms of support are also important. For years, there has been a debate in analyzing the costs of care as per the form of service provision (Chappell et al., 2004; Kaye et al., 2010). Residential forms are indicated as more expensive and less desirable for care recipients, but studies should not only be conducted at a macro level, but also include the individual needs of the dependent older person, the care potential of the family and the local environment, as well as the available services provided in a community-based form.

Currently, there are three main pillars of assistance in the system. The first one is the social care sector, where the cost of services is borne by recipients, family members and local governments. The health sector constitutes the second pillar, where certain services are financed exclusively by the National Health Fund (NFZ) and some are co-financed by the recipient. The last pillar includes services provided by the private sector, which has seen rapid growth in recent years. Each of the sub-systems presented has its own distinctive characteristics in terms of the organization and financing of services.

The aim of this study is to characterize determinants of the development of the silver economy in the area of residential care services in Poland. A forecast of the number of elderly people requiring assistance until 2026 was produced based on the results of research carried out as part of the PolSenior2 project. For the purpose of the analysis, we used statistical and financial data made available by the Ministry of Family and Social Affairs, the Ministry of Health the National Health Fund, the Supreme Chamber of Control, OECD, Eurostat, the Central Statistical Office, among other institutions, in addition to our own research results.

The first part of the article characterises the silver economy and the main issues related to the market of care services for dependent elderly. It presents basic demographic information describing the process of ageing among the Polish population and an estimate of the number of people requiring support in everyday functioning together with the projections until 2060. The next part of the article analyses the three main sectors of residential care, i.e. care services provided by the health care sector, social assistance and private entities. The final section presents the conclusions of these analyses.

1. The silver economy and long-term care

The silver economy is a part of the general economy and includes public and private spending related to the specific needs of people aged 50 and over. Since its emergence, the silver economy has been associated with the health and care services sector provided to the elderly and pension systems. However, it should be emphasised that the silver economy also includes those market segments providing services and products to address the needs of older people (European Commission, 2018). Nevertheless, many researchers point to a broader understanding of the concept of the silver economy, which also includes education, clothing and fashion, media, culture, inter-sectoral activities, gerontotechnologies and social innovation (Klimczuk, 2016; Reshetnikova et al., 2021).

Three main areas can be identified that are relevant when addressing the silver economy. The first is silver industries, which includes goods and services dedicated for seniors and support for older employees. The second important module is social innovation including, in particular, strategies, concepts, and organizational forms. The last is gerontology, which includes research and development of gerontotechnology (Krzyminiewska & Pondel, 2019).

Long-term care services can be classified as meeting basic, or even biological, needs. For elderly people with a low degree of self-care skills, care must often be provided from the beginning of the crisis (e.g., after a stroke, femoral fracture), the duration of which is sometimes counted in months and years (Alzheimer's disease). While we cannot eliminate the risks that affect the dependency of individuals, we can attempt to reduce them through preventive measures such as healthy eating, physical, social and intellectual activity and avoidance of stimulants, etc.

The amount of support provided is determined on an individual basis. It depends mainly on the medical condition, but also on the place of residence, preferences in terms of the organisation and provision of care, the income and accumulated capital of the dependent person and family members, etc. (Cardoso et al., 2012). The availability of services is determined by a number of factors, which we can also include those occurring at the macro level, in particular: the mechanism for organising and financing care services, the number of providers, and the forms and scope of services, which are not uniform.

The third important determinant is the existence of a permanent imbalance with regard to care services provided to dependent elderly people, where demand for services exceeds supply. The deepening imbalance may be further enhanced by the lack of a coherent and long-term social policy for such care, which includes services provided by public entities and the private market of services. When seeking to maintain the current level in meeting the needs of dependent people, the dynamics of investment demographic

and social changes must be monitored. Expenditure is required for the three main determinants necessary for the development of the care services: the care workforce, infrastructure, and mechanisms for financing benefits with private and public funds.

In home and community care, there is a notable shortage of direct care workers (Zagrodney et al., 2023). This problem is global and affects most societies with a growing elderly population (Fujisawa & Colombo, 2005; Prince et al., 2013). Many problems related to resources in long-term care were exacerbated during the COVID-19 pandemic. It is worth noting that the shortage of workers affected carers with different qualifications, including highly skilled workers such as nurses and doctors working in long-term care facilities (Grabowski & Mor, 2020; White et al., 2021; Xu et al., 2020).

The findings of studies on the organization of long-term care systems in EU countries indicate that two models prevail. The first, where responsibility for regulating care services rests with central authorities, and the second, where responsibility is shared between central and regional authorities. This is reflected in the distribution of funds, mainly including the sharing of care costs for both in-house and residential care, and the choice of service providers between public and private ones (Riedel et al., 2016). The solutions adopted for financing care services, which may be prospective or retrospective, have consequences for how the system functions and the inadequacies, which depend on a number of factors, including the wealth of a given population (Costa-Font et al., 2017). In some countries, e.g. the Nordic countries, formal care is more developed, with the state bearing a significant part of the burden. In the case of Western European countries, a model based on social-financial programmes is popular. Financial schemes that operate on the basis of a provision- and insurance-based system allow consumers of services to choose their provider, which contributes to the development of different segments of the care market. In contrast, Mediterranean and central-south-eastern countries are characterized by lower levels of formal care and social transfers. Care is mainly provided by family members or is privately funded (Damiani et al., 2011). In EU countries, an average of 1.8% of GDP was spent on long-term care services provided by health sector providers in 2020. The highest percentage was observed in the Netherlands and amounted to 3.24% of GDP, with the lowest percentage being in Slovakia at 0.03% of GDP. Polish spending in this area is relatively low, as it only constitutes 0.54% of GDP (Eurostat, 2023). According to forecasts, expenditure on health services related to long-term care is expected to increase to around 2.8% of GDP by 2070 for EU countries, and in some countries, e.g. Denmark, it may reach around 7% of GDP (OECD, 2022).

2. Ageing population and the demand for care services

Since OECD member states, which mainly include Western Europe and North America, entered the second demographic transition, the issue of securing care for a growing number of dependent elderly has become increasingly urgent. Attempts have been made for decades to estimate the needs for services as well as the costs in particular populations, taking into account the local characteristics of the current support system (Cardoso et al., 2012; Comas-Herrera et al., 2007; Hrynkiewicz & Halicka, 2022; Spasova et al., 2018, Spetz et al., 2015). In the case of Poland, we are facing a very dynamic process of population ageing. In 2022, people aged 65 and over accounted for 19.5% of the population; by 2060, the percentage of senior citizens will have risen to 32.6%. In turn, the percentage of senior citizens aged 85 and above will have increased from 2.1% in 2022 to more than 6% in 2060 (GUS, 2023a). The caregiving potential of families will decline, as is best illustrated by the ratio of prospective support for elderly parents, which is the quotient of the size of the subpopulation aged 50–64 by the size of the elderly in the same population (85 and older) (Freedman et al., 2024; Kowaleski & Majdzinska, 2012; Redfoot et al., 2013). In 2022, it stood at 8, while by 2060 it will drop to 2.9 (GUS, 2023a). One of the reasons for the dynamic increase in the proportion of older people in the population of Poland but also of other Central and Eastern European countries is the very low fertility rate, which oscillates around 1.3–1.5, and is significantly below the threshold for the simple replacement of generations (2.15). This situation has persisted since the transition period of the 1990s. In addition, many Central European countries, including Poland, recorded a negative migration balance, especially after the enlargement of the European Union in 2004. Migrants were predominantly people of working age, leaving their ageing parents and grandparents in their home countries.

The risk of dependency increases with age and is associated with the fact that an elderly person needs continuous assistance from family or other persons (Legdeur et al., 2018). Senior citizens affected by dependency can be divided into two groups: “requiring assistance” and “definitely requiring assistance”. The latter group includes those with limited ability to perform basic daily activities. The results of a survey conducted as part of the PolSenior2 project indicate that in the 60–64 age group, 6% of men and 9.8% of women require assistance. However, in the age category of 90 and older, 70.9% of men and 84.5% of women require assistance. In the case of seniors definitely requiring assistance in the 60–64 age group, such a need is expressed by 2.6% of men and 4.1% of women. In the oldest age category, the proportion rises to 46.6% of men and 61.3% of women (Błędowski, 2021).

Table 1 presents estimates of the number of senior citizens in each age category who require assistance. In 2022, the total number of such individuals is estimated to be 1.9 million people. In each age category, the predominance of women is noticeable, especially for those over 75 years of age. By 2060, the share of those in need of assistance will rise to 3.3 million people. In this group, the majority will be individuals aged 80 and over.

Table 1. Estimates* of the number of men and women aged 60 and older requiring assistance from other persons between 2022 and 2060 (in thousands) in Poland

Age	Year									
	2022		2030		2040		2050		2060	
	Sex									
	M	W	M	W	M	W	M	W	M	W
60–64	69	127	60	107	80	141	73	129	52	90
65–69	131	162	108	128	123	143	148	169	112	128
70–74	147	148	161	155	133	123	181	163	170	150
75–79	100	204	164	318	143	259	169	294	209	354
80–84	77	249	128	375	161	438	140	355	198	483
85–89	81	262	82	239	179	467	167	399	211	473
90+	48	175	65	222	114	330	180	481	184	451
Total	653	1327	766	1544	932	1900	1058	1990	1136	2130
Total M+W	1980		2310		2832		3048		3265	

Note: * Estimates based on the results of research conducted as part of the PolSenior2 project. Adopted risk levels by age category, for men: 60–64 years: 6%, 65–69 years: 11.6%, 70–74 years: 17%, 75–79 years: 21.7%, 80–84 years: 28.1%, 85–89 years: 52.4%, 90 years and older: 70.9%; for women: 60–64 years: 9.8%, 65–69 years: 11.7%, 70–74 years: 12.5%, 75–79 years: 28.2%, 80–84 years: 47.1%, 85–89 years: 70.7%, 90 years and over: 84.5% (Błędowski, 2021).

Source: based on (Błędowski, 2021; GUS, 2023a).

Table 2, in turn, shows estimates of the number of elderly people who definitely need assistance. Over the course of almost four decades, their number will increase from 0.9 million in 2022 to 1.6 million in 2060. This group is already dominated by senior citizens aged 80 and older in 2022, and this number will grow in the coming decades. Demand for care will increase as people born during the baby boom of the second half of the 20th century pass the second and third thresholds of old age. The analysis of the data presented (Tables 1 and 2) makes it possible to determine prospective demand for long-term care services in each decade. Forecasts facilitate preparing the

care system for the increasing burden, including the development of the silver economy in the area of care services.

Table 2. Estimates* of the number of men and women aged 60 and older definitely requiring assistance from other persons between 2022 and 2060 (in thousands) in Poland

Age	Year									
	2022		2030		2040		2050		2060	
	Sex									
	M	W	M	W	M	W	M	W	M	W
60–64	30	53	26	45	35	59	32	54	23	38
65–69	50	62	41	49	47	55	56	65	43	49
70–74	77	36	84	37	70	29	95	39	89	36
75–79	42	75	69	116	60	95	71	107	87	129
80–84	39	121	65	182	81	213	71	173	100	235
85–89	35	128	35	117	77	228	72	194	91	231
90+	31	127	42	161	75	239	118	349	121	327
Total	304	601	362	707	444	918	515	981	553	1045
Total M+W	905		1069		1362		1496		1599	

Note: * Estimates based on the results of research conducted as part of the PolSenior2 project. Adopted risk levels by age category, for men: 60–64 years: 2.6%, 65–69 years: 4.4%, 70–74 years: 8.9%, 75–79 years: 9.1%, 80–84 years: 14.2%, 85–89 years: 22.6%, 90 years and older: 46.6%.; for women: 60–64 years: 4.1%, 65–69 years: 4.5%, 70–74 years: 3%, 75–79 years: 10.3%, 80–84 years: 22.9%, 85–89 years: 34.5%, 90 years and over: 61.3% (Błędowski, 2021).

Source: based on (Błędowski, 2021; GUS, 2023a).

According to estimates published in an OECD report (2021a), 34% of people aged 65 and older report limitations when performing daily activities, and 16% describe these as significant. If these criteria are adopted, in 2022 there were 2.5 million elderly people reporting activity limitations and 1.1 million described them as significant. In 2060, their number would rise to 3.4 million and 1.6 million, respectively. Regardless of the scenarios adopted, the demand for long-term care services will grow rapidly. When analysing the issues of service development prospects, it is important to consider that the needs of elderly citizens will vary greatly. This will be due to health conditions, personal preferences in terms of the expected support, and the possibility of providing support based on available resources. Therefore, for the purposes of this discussion, it is worth distinguishing several major segments of the market for goods and services dedicated to dependent elderly, family and professional caregivers:

- Care services provided directly to dependent seniors in residential and community-based forms.
- Benefits for family caregivers.
- Equipment and solutions used when providing care.
- Equipment and solutions to enhance or maintain self-care skills.

Each of the areas identified has its own individual development prospects. Some are dependent on legal regulations for organization and financing, while others are mainly subject to factors occurring in the free market for goods and services.

3. Residential care

Services provided in the form of residential care should be reserved for dependents for whom safe and appropriate care in the form of health visiting service cannot be secured. This is the most expensive form and requires significant investment to make the delivery of services possible (Konetzka, 2014; Marek et al., 2012). Greater demand for this type of service is affected by the following factors:

1. Demographic: increase in the number of elderly people in the population; decline in the care potential of families; increase in the number of elderly individuals with low levels of independence; increase in the number of single-person households among the elderly; external long-term migration of working-age family members.
2. Systemic: insufficient supply of services provided in the community (Konetzka, 2014). In this case, people who could have benefited from community-based support, if it were available, end up in the residential care system. Lack of indirect forms of care in the form of day care for dependents and respite care.
3. Community: lack of support for family caregivers in the form of counselling, guidance, training and support groups; insufficient financial support for family caregivers when they give up their professional duties.

In the case of the first factor, it is largely independent and the influence from social policy is significantly limited. For the other two factors, on the other hand, this importance is dependent to a great extent on long-term care policies. There are three main pillars in residential care: facilities run by the health sector, including mainly medical care and nursing facilities; facilities run by the social welfare sector in the form of nursing homes; and private care facilities, i.e. nursing homes and family care homes.

3.1. The health sector

In 2022, there were 527 long-term care facilities operated by health sector entities and co-financed by the National Health Fund, with nearly 30,000 places (Table 3). Over the period of 20 years, the number of facilities has increased by 80 since 2012, and the number of places by more than 7,500. According to a report by the Supreme Chamber of Control, 61% of patients are aged 65 and older (Najwyższa Izba Kontroli, 2020). Therefore, the ratio of the number of available places in facilities to the number of senior citizens is important. In 2012, on a national scale, there were 4 places per 1,000 people aged 65 and older; for 2022, we can observe a slight increase to 4.1. It is worth noting that there are large differences between voivodships. The largest number of places in medical care and nursing facilities per number of residents is in Podkarpackie (6) and Dolnośląskie (5.7) voivodships, with Wielkopolskie (1.7) and Warmińsko-mazurskie (2.5) voivodships at the other end of the spectrum.

The reasons for the differences between the voivodships in terms of the development of health care infrastructure aiming to meet the treatment and care needs of dependent people, mainly those in advanced age, can be found in the social policy priorities in the field of health care pursued by local authorities (at the level of municipalities, districts and the voivodships), and the differences in the dynamics of ageing in particular regions. It is mainly local authorities that undertake initiatives related to the development of infrastructure, equipment and employment of the necessary staff in order to then apply for a contract for the provision of services with a public insurer. This type of investment requires the securing of substantial financial resources as early as at the establishment stage.

Care and nursing facilities are intended for dependent persons who have completed hospital treatment, but require specialized care provided by health sector personnel. Patients are referred to the facility by a doctor and have a score of up to 40 on a test conducted using the Barthel scale. The fee for the stay paid by elderly person amounts to up to 70% of their pension, with the remaining cost being covered by the National Health Fund. Family members are not required to co-pay.

In 2012, PLN 1,000,610,000 was allocated from the National Health Fund's budget to nursing and long-term care services, which accounted for 1.65% of overall spending on health services. In contrast, in 2022, spending on the range of services analysed here increased to PLN 1,892,479,000, which accounted for 2.79% of the National Health Fund's budget spent on health services (Ministerstwo Zdrowia, 2023). The rise in spending was influenced by an increase in the number of providers but also by rising costs of implementing medical and care procedures.

Table 3. Medical care and nursing facilities between 2012 and 2022 in Poland

Voivodship	2012			2022		
	Number of care and nursing facilities	Beds	Beds/1000 aged 65+	Number of care and nursing facilities	Beds	Beds/1000 aged 65+
Dolnośląskie	54	2652	6.4	64	3217	5.7
Kujawsko-pomorskie	29	1222	4.3	38	1659	4.4
Lubelskie	20	794	2.4	22	1275	3.2
Lubuskie	12	407	3.1	11	579	3.2
Łódzkie	28	1349	3.3	36	1943	3.9
Małopolskie	33	2439	5.1	39	3156	5.3
Mazowieckie	66	4128	5.2	67	5032	5.0
Opolskie	0	0	0.0	15	769	4.2
Podkarpackie	36	1818	6.3	40	2224	6.0
Podlaskie	21	536	3.0	19	727	3.4
Pomorskie	27	1088	3.7	28	1201	2.9
Śląskie	59	3002	4.3	63	3905	4.5
Świętokrzyskie	14	618	3.1	19	1039	4.2
Warmińsko--mazurskie	14	520	2.9	18	608	2.5
Wielkopolskie	21	728	1.6	28	1012	1.7
Zachodniopomorskie	13	440	1.9	20	936	2.9
TOTAL	447	21741	4.0	527	29282	4.1

Source: based on (Bank Danych Lokalnych, 2023; GUS, 2013, 2022).

It is a very expensive form of long-term care, not least because of the need to hire highly specialized personnel, including doctors, nurses and therapists. Due to the fact that the market for health care services has experienced a shortage of sufficient medical staff for years, this can be a very big obstacle to the development of this type of services, especially since, according to an OECD report, Poland has a very low ratio of doctors (2.4) and nurses (5.1) per 1,000 inhabitants (OECD, 2020). What is more, the development of this segment of services is largely dependent on the payer (the National Health Fund), which makes the valuation of services and determines the limits of funds allocated for contracting services in this area.

3.2. Social welfare sector

The expansion of the care infrastructure and the co-financing of services in social welfare homes is mainly the responsibility of local governments. In 2022, there were 822 facilities of this type operating across the country, with a combined total of more than 80,000 places (Table 4). On average, the facilities had 97 places at their disposal, although it is worth emphasizing that there are social welfare homes with 300 residents. Currently, the maximum size of an institution should not exceed 100 residents. Since 2012, only 20 new social welfare homes have been opened, and the number of places has increased by only 1,490 (MRiPS, 2023a, 2023b). Although the largest number of such homes is in Śląskie voivodeship (97) with a total of 8573 places, and Mazowieckie (92) voivodship with 9340 places, the highest level of security per 1,000 people aged 65 and over is recorded in the Opolskie voivodship with 16.4 and the Warmińsko-mazurskie voivodship with 15 (MRiPS, 2023b).

The cost of staying in the facility is covered primarily by the recipient (up to 70% of their pension, disability or allowance benefits), then ascendants and descendants (a surcharge applies above 300% of the income criterion in social assistance), with the shortfall being paid by the municipality. Municipal expenditure on surcharges for stays in social welfare homes is growing rapidly, which is mainly due to the increasing monthly cost of stay; in 2023 this was on average around PLN 6,000–6,500 (EUR 1.3–1.4).[2] In contrast, the average retirement benefit was PLN 3,300 gross (EUR 726) (ZUS, 2023b). Surcharges from family members obligated to pay welfare are at very low levels. Therefore, some municipalities try to limit the issuance of referrals to social welfare homes, thus avoiding future costs (Iwański, 2016). Of the 56,000 people admitted to social welfare homes under regulations introduced in 2004 (described above) who were in 2022 in the facilities, the fee for 54,000 had to be paid by municipalities. In 2012, municipalities spent PLN 810 million (EUR 178 million) to support stays in social welfare homes, which accounted for 0.58% of total local government expenditure. In 2022, on the other hand, PLN 2,237 million (EUR 492 million) was spent nationwide for this purpose, which constitutes a 176% increase over the decade. The share in municipal budgets has increased to 0.90% (Bank Danych Lokalnych, 2024; MRiPS, 2023b).

As there are noticeably low dynamics of investment carried out by local governments in terms of the building new social welfare homes, the private--public partnership model is a solution. Municipalities can provide land for investment and guarantee the referral of residents of their municipality to the facility once it obtains the status of a social welfare home of a certain type

[2] Calculation based on the average annual euro exchange rate in 2023 of EUR 1 = PLN 4.5436 (Ministerstwo Finansów, 2024).

Table 4. Social welfare homes in Poland – as of 2022

Voivodship	Number of facilities	Number of places	Average number of residents	Number of places per 1,000 persons aged 65+	Number of persons aged 65+
Dolnośląskie	61	5834	96	10.3	565557
Kujawsko-pomorskie	46	3982	87	10.6	376036
Lubelskie	46	4525	98	11.3	400554
Lubuskie	23	2236	97	12.2	183109
Łódzkie	55	6134	112	12.2	503516
Małopolskie	89	8114	91	13.7	592767
Mazowieckie	92	9340	102	9.3	1002043
Opolskie	28	3030	108	16.4	184335
Podkarpackie	52	4881	94	13.3	368208
Podlaskie	22	2177	99	10.2	213022
Pomorskie	40	4045	101	9.9	407857
Śląskie	97	8573	88	9.9	869576
Świętokrzyskie	36	3336	93	13.4	248774
Warmińsko--mazurskie	40	3656	91	15.0	244010
Wielkopolskie	63	6382	101	10.5	606730
Zachodniopomorskie	32	3870	121	12.0	322343
TOTAL	822	80115	97	11.3	7088437

Source: based on (GUS, 2023b; MRiPS, 2023b; data from records of 16 voivodship offices).

(there are seven main types of facilities for dependents, e.g., somatically ill, mentally ill). Meanwhile, a private investor would build the facility, equip it and run it, once it receives regulatory approvals for operation and meets the standards set forth in the relevant legislation (entry in the register kept by the voivodship office). Although the first investments of this type have already been completed in Poland, it is still not a popular model, despite its potential to provide opportunities for the expansion of care infrastructure in the form of social welfare homes, and to strengthen this market segment. Given the high demand for this form of care and the potential for development, this could be an attractive form of capital investment for private entities.

The return rate on this type of investment is long-term, due to the need for substantial resources during the construction stage, as well as equipping the

building and hiring the necessary staff. The investor must have a guarantee that the local government with which it is cooperating will refer residents in need of residential care to the facility for a specified period (e.g., 10–20 years). Since, in the vast majority of cases, municipalities pay a substantial surcharge for residents in nursing homes, if the monthly cost of care rises dynamically (through inflation, increases in labour costs, etc.), then municipalities may not have the necessary funds to cover the fee. In this case, the investment risk increases significantly. In order to reduce the risk, the investor may allocate part of the places to senior citizens referred by the municipalities on the basis of tasks set out in the Social Assistance Act. Another part of the resources could be dedicated to providing care services for private clients. In this type of facility, it may also be worth considering community-based activities, e.g., in the form of a rehabilitation equipment rental, a training facility, a shop with essential care items and community care services, etc. Expanding the business profile can help diversify revenue sources.

3.3. Private sector of residential long-term care

Although nominally there is a slight increase in the number of places in social welfare homes, the supply in relation to the number of elderly people, including dependents, has been steadily declining in recent years. This situation contributes to the development of the private market for residential care. Facilities providing 24-hour care for the disabled, chronically ill or elderly may be provided by business entities or non-government organizations for which this type of activity is included in the statute (Ustawa, 2004). In the case of business activity in the form of nursing homes or family care homes, the total cost of the stay is borne by the senior citizen or by family members who agree to partake in the costs. A contract is signed between the recipient and the provider, which specifies the scope of services and the terms of payment for additional services, e.g., medicines, diapers, additional rehabilitation treatments).

In 2023, there were 753 facilities providing 24-hour care for people with disabilities, chronic illnesses or the elderly nationwide (Table 5). The largest number is in Mazowieckie voivodship (151) and Śląskie voivodship (92), the lowest in Świętokrzyskie voivodship (19) and Lubuskie voivodship (14). If we consider the number of places in institutions per 1,000 people aged 65 and over, the most favourable ratio is in Pomorskie voivodship (7.3) and the worst in Podkarpackie voivodship (1.6). Facilities of this type are mostly smaller than social welfare homes, and house an average of 38 residents. The largest residential care facility run as a business or by a non-government organization is found in Małopolskie voivodship, with 195 places. There are also facilities run as family care homes, which care for 5–6 dependents.

Table 5. Facilities providing 24-hour care for persons with disabilities, chronic illnesses or the elderly as business and statutory activities

Voivodship	Number of facilities	Number of places	Average number of residents	Number of places per 1,000 persons aged 65+	Number of persons aged 65+
Dolnośląskie	61	2 402	39	4.2	565557
Kujawsko-pomorskie	29	885	31	2.4	376036
Lubelskie	20	667	33	1.7	400554
Lubuskie	14	373	27	2.0	183109
Łódzkie	36	1136	32	2.3	503516
Małopolskie	53	2264	43	3.8	592767
Mazowieckie	151	6682	44	6.7	1002043
Opolskie	32	1230	38	6.7	184335
Podkarpackie	20	607	30	1.6	368208
Podlaskie	54	940	17	4.4	213022
Pomorskie	83	2961	36	7.3	407857
Śląskie	92	3973	43	4.6	869576
Świętokrzyskie	19	856	45	3.4	248774
Warmińsko--mazurskie	22	868	39	3.6	244010
Wielkopolskie	43	1608	37	2.7	606730
Zachodniopomorskie	24	1091	45	3.4	322343
TOTAL	753	28543	38	4.0	7088437

Source: based on (GUS, 2023b; MRiPS, 2023b; data from records of 16 voivodship offices).

New facilities are established every year, and once they meet the standards, they are approved to provide residential care. However, it is worth pointing out that this form of business, although it has prospects for growth, does have certain limitations. The first of these is related to the need for substantial investment expenditures. As early as at the stage of constructing or adapting the building for care purposes, the investor must decide which group of clients it will target. The main issues here are the standard of rooms (single, double or triple), the size and number of common spaces, the area for recreation, rooms for rehabilitation, occupational therapy, and room equipment, etc. Although the regulations set minimum staffing requirements, some clients will look for facilities that provide a higher level of services in terms of reha-

bilitation, leisure activities, etc. Another important issue is the retention of care staff, including medical staff. There is a shortage of care workers in the market, and staffing problems are reported by both public and private facilities (Iwański, 2019; Kałuża-Kopias, 2018). This problem concerns most countries in the European Union, leading care staff to migrate to countries where wages in this sector are higher (Facchini, 2022; Leiber et al., 2019). Salaries in this market segment are not high, and are close to the minimum wage in some facilities in Poland. Providing salaries at a level that will allow staff to be retained feeds into increased costs, which, in the case of free market entities, is reflected in the price of the services offered. The monthly estimated cost of staying in a private facility ranges from PLN 4,000 (EUR 880) to over PLN 10,000 (EUR 2,200), depending on the standard.

Conclusions

Residential long-term care represents an important segment of the silver economy in terms of demographics, and the three main market areas distinguished in this field should enjoy strong prospects for growth. However, we can identify several factors that are of key importance for the care services market to develop. First of all, it is essential to initiate efforts to train nursing staff, both in the area of social assistance and health care. Many nursing homes are already facing a shortage of nurses, for whom this type of facility is not an attractive place to work in terms of salary. Allowing social welfare homes to contract medical services with the National Health Fund could be one of the solutions to this issue; in fact, this is a solution that managers in this sector have been advocating for years. Furthermore, it is important to educate professional caregivers who will be able to provide proper and safe care for dependents. However, care professions are not attractive in the labour market, which is mainly due to low salaries and high level of professional responsibility (chronically ill, dying people). Indeed, staff shortages in long-term care have worsened in recent years in many countries (Chen et al., 2023; Scales, 2021)

Considering that in the coming years the purchasing power of the average pension can be expected to decline and the cost of care to rise, it seems necessary to introduce additional mechanisms for co-funding care services, e.g. a solution similar to German ones in the form of long-term care insurance (Freudenberg, 2015; Przybyłowicz, 2017; Sawulski et al., 2019). While there were draft bills in Poland that intended to introduce an insurance model co-financing long-term care, they did not meet with the approval of legislative representatives (Poselski projekt, 2018). What is worth emphasising is that the bill adopted in 2023 on the Support Benefit for Persons with Disabilities

excludes from support those residing in residential care facilities, regardless of their type and operating authority (Ustawa, 2023). The result of a 2016 study in Poland (with a nationwide sample of 1,026 respondents) indicated that 65 respondents would consider taking out private care insurance that would pay out in the event of dependency (Jurek, 2019).

The cost of care will continue to rise in years to come, which is primarily due to rising labour costs. This is influenced by the rapid growth in Poland's minimum wage in recent years (PLN 2600 (EURO 572) gross in 2020 to a projected level of PLN 4300 (EURO 946) gross in mid-2024 (ZUS, 2023a). Rising costs are also impacted by high inflation. Thus, expenses for food, energy and the building maintenance in general are increasing. If the monthly cost of a stay continues to grow dynamically, the availability of services will decrease, especially for entities run as a business. Social welfare homes will also be affected, as municipalities will limit referrals to facilities if the amounts resulting from surcharges on stays exceed the financial capacity of poorer local governments. In the case of services provided by health sector entities, with rising costs, the supply of services may decrease, especially if the public payer does not include rising costs in its contracts.

References

Bank Danych Lokalnych. (2024). *Revenue of gminas and cities with powiat status budgets*. https://bdl.stat.gov.pl/bdl/dane/podgrup/tablica

Błędowski, P. (2021). Potrzeby opiekuńcze. In P. Błędowski, T. Grodzicki, M. Mossakowska & T. Zdrojewski (Eds.), *PolSenior 2 Badanie poszczególnych obszarów stanu zdrowia osób starszych, w tym jakości życia związanej ze zdrowiem* (pp. 913–930). Wydawnictwo Gdańskiego Uniwersytetu Medycznego.

Cardoso, T., Oliveira, M. D., Barbosa-Póvoa, A., & Nickel, S. (2012). Modeling the demand for long-term care services under uncertain information. *Health Care Management Science*, *15*, 385–412. https://doi.org/10.1007/s10729-012-9204-0

Chappell, N. L., Dlitt, B. H., Hollander, M. J., Miller, J. A., & Carol McWilliam, C. (2004, June). Comparative costs of home care and residential care. *The Gerontologist*, *44*(3), 389–400.

Chen, M., Goodwin, J. S., Bailey, J. E., Bowblis, J. R., Li, S., & Xu, H. (2023). Longitudinal associations of staff shortages and staff levels with health outcomes in nursing homes. *Journal of the American Medical Directors Association*, *24*(11), 1755–1760.

Comas-Herrera, A., Wittenberg, R., Pickard, L., & Knapp, M. (2007). Cognitive impairment in older people: Future demand for long-term care services and the associated costs. *International Journal of Geriatric Psychiatry*, *22*(10), 1037–1045. https://doi.org/10.1002/gps.1830

Costa-Font, J., & Courbage, C. (2011). *Financing long-term care in Europe*. Palgrave Macmillan.

Costa-Font, J., Courbage, C., & Zweifel, P. (2017). Policy dilemmas in financing long-term care in Europe. *Global Policy*, *8*, 38–45.

Damiani, G., Farelli, V., Anselmi, A., Sicuro, L., Solipaca, A., Burgio, A., Iezzi, D. F., & Ricciardi, W. (2011). Patterns of long-term care in 29 European countries: Evidence from an exploratory study. *BMC Health Services Research*, *11*, 1–9.

Dintrans, P. V. (2020). Designing long-term care systems: Elements to consider. *Journal of Aging & Social Policy*, *32*(1), 83–99.

Dudley, K. (1996). Demographic transition theory. *Population Studies*, *50*(3), 361–387. https://doi.org/10.1080/0032472031000149536

European Commission, Directorate-General for Communications Networks, Content and Technology. (2018). *The silver economy: Final report*. Publications Office. https://data.europa.eu/doi/10.2759/685036

Eurostat. (2023). *Long-term care (health) expenditure*. https://ec.europa.eu/eurostat/databrowser/explore/all/popul?lang=en&subtheme=hlth.hlth_care.hlth_sha11.hlth_sha11_sum&display=list&sort=category&extractionId=tps00214

Facchini, C. (2020). Caring for non-self-sufficient older people in Italy: From a familistic system to the immigrant live-in careworker model. *Public Sciences & Policies*, *6*(2), 149–168.

Freedman, V. A., Agree, E. M., Seltzer, J. A., Birditt, K. S., Fingerman, K. L., Friedman, E. M., Lin, I. F., Margolis, R., Park, S. S., Patterson, S. E., Polenick, C. A., Reczek, R., Reyes, A. M., Truskinovsky, Y., Wiemers, E. E., Wu, H., Wolf, D. A., Wolff, J. L., & Zarit, S. H. (2024). The changing demography of late-life family caregiving: A research agenda to understand future care networks for an aging US Population. *The Gerontologist*, *64*(2), 1–7. https://doi.org/10.1093/geront/gnad036

Freudenberg, C. (2015). *Adequacy and fiscal sustainability: a long-term evaluation of public pension schemes in Hungary and Poland* (Doctoral dissertation). Albert-Ludwigs-Universität Freiburg.

Fujisawa, R., & Colombo, F. (2009). *The long-term care workforce: Overview and strategies to adapt supply to a growing demand*. OECD Health Working Papers, 44. https://doi.org/10.1787/225350638472

Grabowski, D. C., & Mor, V. (2020). Nursing home care in crisis in the wake of COVID-19. *JAMA*, *324*(1), 23–24.

GUS. (2013). *Zdrowie i ochrona zdrowia w 2012 r*. https://stat.gov.pl/obszary-tematyczne/zdrowie/zdrowie/zdrowie-i-ochrona-zdrowia-w-2012-r-,1,3.html

GUS. (2022). *Zdrowie i ochrona zdrowia w 2021 roku*. https://stat.gov.pl/obszary-tematyczne/zdrowie/zdrowie/zdrowie-i-ochrona-zdrowia-w-2021-roku,1,12.html

GUS. (2023a). *Prognoza ludności na lata 2023–2060*. https://demografia.stat.gov.pl/bazademografia/Prognoza_2023_2060.aspx

GUS. (2023b). *NSP 2021 – wyniki ostateczne*. https://stat.gov.pl/spisy-powszechne/nsp-2021/nsp-2021-wyniki-ostateczne/

Hrynkiewicz, J., & Halicka, M. (2022). *Niesamodzielna starość*. Rządowa Rada Ludnościowa, Główny Urząd Statystyczny. https://kd.stat.gov.pl/images/publikacje/02_kd_niesamodzielna_starosc_internet.pdf

Iwański, R. (2016). *Opieka długoterminowa nad niesamodzielnymi osobami starszymi*. Cedewu.

Iwański, R. (2019). System pomocy społecznej a starzenie się populacji – wyzwania i bariery w zakresie opieki długoterminowej. *Studia z Polityki Publicznej / Public Policy Studies*, *1*(21), 39–52.

Jurek, Ł. (2019). Zapotrzebowanie na prywatne ubezpieczenie opiekuńcze w Polsce. *Rozprawy Ubezpieczeniowe. Konsument na Rynku Usług Finansowych*, *33*(3), 68–80.

Kałuża-Kopias, D. (2018). Imigranci w systemie opieki nad osobami starszymi. *Rynek Pracy*, *3*(166), 36–45.

Kaye, H. S., Harrington, C., & LaPlante, M. P. (2010). Long-term care: who gets it, who provides it, who pays, and how much? *Health Affairs*, *29*(1), 11–21.

Klimczuk, A. (2016). Modele ‚srebrnej gospodarki' w Unii Europejskiej w ujęciu porównawczym. Próba wprowadzenia do dyskusji. *Problemy Zarządzania*, *14*(2), 41–59.

Konetzka R.T. (2014). The hidden costs of rebalancing long-term care. *Health Service Research*, *49*(3), 771–777. https://doi.org/10.1111/1475-6773.12190

Kowaleski, J., & Majdzińska, A. (2012). Miary i skale zaawansowania starości demograficznej. In A. Rossa (Ed.), *Wprowadzenie do gerontometrii* (pp. 7–34). Wydawnictwo Uniwersytetu Łódzkiego.

Krzyminiewska, G., & Pondel, H. (2019). *Silver economy and the development policy of rural areas in Poland*. In Hradec Economic Days: Double-blind Peer Reviewed. Proceedings part I of the International Scientific Conference Hradec Economic Days.

Legdeur, N., Heymans, M. W., Comijs, H. C., Huisman, M., Maier, A. B., & Visser, P. J. (2018). Age dependency of risk factors for cognitive decline. *BMC Geriatrics*, *18*, 1–10.

Leiber, S., Matuszczyk, K., & Rossow, V. (2019). Private labor market intermediaries in the Europeanized live-in care market between Germany and Poland: A typology. *Zeitschrift für Sozialreform*, *65*(3), 365–392.

Marek, K. D., Stetzer, F., Adams, S. J., Popejoy, L. L., & Rantz, M. (2012). Aging in place versus nursing home care: Comparison of costs to Medicare and Medicaid. *Research in Gerontological Nursing*, *5*(2), 123–129.

Ministerstwo Finansów. (2024). Średnie arytmetyczne średnich kursów wybranych walut obcych *w 2023 r*. https://www.gov.pl/web/finanse/srednie-arytmetyczne

Ministerstwo Zdrowia. (2023). *Finanse NFZ*. https://www.nfz.gov.pl/bip/finanse-nfz/

MRiPS. (2023a). *Statystyka za rok 2013*. https://www.gov.pl/web/rodzina/statystyka-za-rok-2013

MRiPS. (2023b). *Statystyka za rok 2022*. https://www.gov.pl/web/rodzina/statystyka-za-rok-2022

Najwyższa Izba Kontroli. (2020). *Dostępność opieki długoterminowej finansowanej ze środków NFZ*. https://www.nik.gov.pl/kontrole/P/19/061/

OECD. (2020). *OECD health statistics 2020*. https://www.oecd-ilibrary.org/social-issues-migration-health/data/oecd-health-statistics_health-data-en

OECD. (2021a). *Health at a glance 2021: OECD indicators*. OECD Publishing. https://doi.org/10.1787/ae3016b9-en

OECD. (2021b). *Health Statistics 2021, OECD historical population data and projections*. https://stats.oecd.org/Index.aspx?DataSetCode=POP_PROJ

OECD. (2022). *Health at a glance: Europe 2022: State of health in the EU Cycle. Public expenditure projections for health and long-term care*. https://www.oecd-ilibrary.org/sites/5b352bdf-en/index.html?itemId=/content/component/5b352bdf-en

Poselski projekt. (2018). Poselski projekt ustawy o pomocy osobom niesamodzielnym. https://www.sejm.gov.pl/sejm8.nsf/PrzebiegProc.xsp?nr=2771

Poškutė, V., & Greve, B. (2017). Long-term care in Denmark and Lithuania—a most dissimilar case. *Social Policy & Administration*, *51*(4), 659–675. https://doi.org/10.1111/spol.12318

Prince, M., Prina, M., & Guerchet, M. (2013). *World Alzheimer report 2013: Journey of caring: An analysis of long-term care for dementia*. Alzheimer's Disease International.

Przybyłowicz, A. (2017). *Ubezpieczenie pielęgnacyjne w Republice Federalnej Niemiec*. Zakład Ubezpieczeń Społecznych.

Redfoot, D., Feinberg, L., & Houser, A. N. (2013). *The aging of the baby boom and the growing care gap: A look at future declines in the availability of family caregivers* (vol. 12). AARP Public Policy Institute.

Reshetnikova, L., Boldyreva, N., Perevalova, M., Kalayda, S., & Pisarenko, Z. (2021). Conditions for the growth of the "silver economy" in the context of sustainable development goals: Peculiarities of Russia. *Journal of Risk and Financial Management*, 14(9), 401. https://doi.org/10.3390/jrfm14090401

Riedel, M., Kraus, M., & Mayer, S. (2016). Organization and supply of long-term care services for the elderly: A bird's-eye view of old and new EU Member States. *Social Policy & Administration*, *50*(7), 824–845.

Rudakova, E. K., Ustinkin, S. V. & Samoylova, M. P. (2023). Analysis of demographic processes in Europe: Demographic transition or depopulation? In S. G. Maximova (Ed.), *Complex social systems in dynamic environments* (pp. 323–331). Series: Lecture Notes in Networks and Systems, vol. 365. Springer. https://doi.org/10.1007/978-3-031-23856-7_30

Sawulski, J., Magda, I., & Lewandowski, P. (2019). Czy polski system emerytalny zbankrutuje. *IBS Policy Paper*, (2), 1–11.

Scales, K. (2021). It is time to resolve the direct care workforce crisis in long-term care. *The Gerontologist*, *61*(4), 497–504.

Spasova, S., Baeten, R., & Vanhercke, B. (2018). Wyzwania opieki długoterminowej w Europie. *Eurozdrowie*, *24*(4), 7–12.

Spetz, J., Trupin, L., Bates, T., & Coffman, J. M. (2015). Future demand for long-term care workers will be influenced by demographic and utilization changes. *Health Affairs*, *34*(6), 936–945. https://doi.org/10.1377/hlthaff.2015.0005

Ustawa. (2024). Ustawa z dnia 12 marca 2004 r. o pomocy społecznej (Dz.U. z 2021 r. poz. 2268 ze zm. https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=wdu20040640593

Ustawa. (2024). Ustawa z dnia 7 lipca 2023 r. o świadczeniu wspierającym (Dz.U. 2023 poz. 1429. (access:8.10.2023) https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id= WDU20230001429

WHO. (2022). *Ageing and health*. https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

White, E. M., Wetle, T. F., Reddy, A., & Baier, R. R. (2021). Front-line nursing home staff experiences during the COVID-19 pandemic. *Journal of the American Medical Directors Association*, *22*(1), 199–203.

Xu, H., Intrator, O., & Bowblis, J. R. (2020). Shortages of staff in nursing homes during the COVID-19 pandemic: What are the driving factors? *Journal of the American Medical Directors Association*, *21*(10), 1371–1377.

Zagrodney, K. A. P., King, E. C., Simon, D., Nichol, K. A., & McKay, S. M. (2023). Economic evidence for home and community care investment: The case for Ontario personal support workers' wage parity. *Healthcare Policy*, *19*(1), 23–31. https://doi.org/10.12927/hcpol.2023.27161

ZUS. (2023). *Minimalne i przeciętne wynagrodzenie*. https://www.zus.pl/baza-wiedzy/skladki-wskazniki-odsetki/wskazniki/minimalne-i-przecietne-wynagrodzenie

ZUS. (2023). *Struktura wysokości emerytur i rent*. https://www.zus.pl/baza-wiedzy/statystyka/opracowania-tematyczne/struktura-wysokosci-emerytur-i-rent

The registers of long-term care facilities published by the voivodship offices

Dolnośląski Urząd Wojewódzki. (2023). Rejestr placówek zapewniających całodobową opiekę osobom niepełnosprawnym, przewlekle chorym lub osobom w podeszłym wieku, w tym na podstawie przepisów o działalności gospodarczej. Retrieved October 8, 2023 from https://bip.duw.pl/bip/rejestry-ewidencje-arc/wydzial-zdrowia-i-polit/52,Wydzial-Zdrowia-i-Polityki-Spolecznej.html

Kujawsko-Pomorski Urząd Wojewódzki. (2023). Rejestr placówek całodobowej opieki. Retrieved October 8, 2023 from https://bip.bydgoszcz.uw.gov.pl/75/rejestr-placowek-calodobowej-opieki.html

Lubelski Urząd Wojewódzki. (2023). Wojewódzki rejestr placówek zapewniających całodobową opiekę. Retrieved October 8, 2023 from https://luwwlublinie.bip.gov.pl/wydzial-polityki-spolecznej-rejestry/wojewodzki-rejestr-placowek.html

Lubuski Urząd Wojewódzki. (2023). Rejestr placówek zapewniających całodobową opiekę. Retrieved October 8, 2023 from https://bip.lubuskie.uw.gov.pl/polityka_spoleczna/rejestry

Łódzki Urząd Wojewódzki. (2023). Rejestr placówek zapewniających całodobową opiekę osobom niepełnosprawnym, przewlekle chorym lub osobom w podeszłym wieku za 2023 rok. Retrieved October 8, 2023 from https://www.gov.pl/web/uw-lodzki/rejestr-placowek-zapewniajacych-calodobowa-opieke-osobom-niepelnosprawnym-przewlekle-chorym-lub-osobom-w-podeszlym-wieku2

Małopolski Urząd Wojewódzki. (2023). Rejestr domów pomocy społecznej oraz placówek zapewniających całodobową opiekę. Retrieved October 8, 2023 from https://bip.malopolska.pl/muw,a,3261,rejestr-domow-pomocy-spolecznej-oraz-placowek-zapewniajacych-calodobowa-opieke.html

Mazowiecki Urząd Wojewódzki. (2023). Rejestr placówek zapewniających całodobową opiekę osobom niepełnosprawnym, przewlekle chorym lub osobom w podeszłym wieku, w tym prowadzonych na podstawie przepisów o działalności gospodarczej. Retrieved October 8, 2023 from https://www.gov.pl/web/uw-mazowiecki/rejestr-placowek-zapewniajacych-calodobowa-opieke-osobom-niepelnosprawnym-przewlekle-chorym-lub-osobom-w-podeszlym-wieku-w-tym-prowadzonych-na-podstawie-przepisow-o-dzialalnosci-gospodarczej

Opolski Urząd Wojewódzki. (2023). Rejestr placówek zapewniających całodobową opiekę osobom niepełnosprawnym, przewlekle chorym lub osobom w podeszłym wieku. Retrieved October 8, 2023 from https://www.gov.pl/web/uw-opolski/rejestr-placowek-zapewniajacych-calodobowa-opieke-osobom-niepelnosprawnym-przewlekle-chorym-lub-osobom-w-podeszlym-wieku

Podkarpacki Urząd Wojewódzki. (2023). Rejestr placówek zapewniających całodobową opiekę osobom niepełnosprawnym, przewlekle chorym lub osobom w podeszłym wieku w ramach działalności gospodarczej. Retrieved October 8, 2023 from https://rzeszow.uw.gov.pl/dla-klienta/pomoc-spoleczna/wykaz-jednostek-udzielajacych-wsparcia-potrzebujacym

Podlaski Urząd Wojewódzki. (2021). Rejestr placówek zapewniających całodobową opiekę osobom niepełnosprawnym, przewlekle chorym lub osobom w podeszłym wieku. Retrieved October 8, 2023 from https://puw.bip.gov.pl/prowadzone-rejestry/rejestr-placowek-zapewniajacych-calodobowa-opieke-osobom-niepelnosprawnym-przewlekle-chorym.html

Pomorski Urząd Wojewódzki. (2023). Rejestr placówek zapewniających całodobową opiekę osobom niepełnosprawnym, przewlekle chorym lub osobom w podeszłym wieku. Retrieved October 8, 2023 from https://uwgdansk.bip.gov.pl/rejestry-wps/rejestr-placowek-zapewniajacych-calodobowa-opieke-osobom-niepelnosprawnym-przewlekle-chorym-lub-osobom-w-podeszlym-wieku.html

Śląski Urząd Wojewódzki. (2023). Placówki Całodobowej Opieki. Retrieved October 8, 2023 from https://suw.bip.gov.pl/placowki-pomocy-spolecznej/wykaz-zarejestrowanych-placowek-zapewniajacych-calodobowa-opieke-osobom-niepelnosprawnym-przewlekle-chorym-lub-w-podeszlym-wieku-prowadzonym-w-ramach-dzialalnosci-statutowej-lub-gospodarczej-w-wojewodztwie-slaskim.html

Świętokrzyski Urząd Wojewódzki. (2023). Wykazy i rejestry placówek. Retrieved October 8, 2023 from https://www.kielce.uw.gov.pl/pl/urzad/polityka-spoleczna/wykazy-i-rejestry-placo/6569,Wykazy-i-rejestry-placowek.html

Warmińsko-Mazurski Urząd Wojewódzki. (2023). Rejestry prowadzone w Wydziale Polityki Społecznej. Retrieved October 8, 2023 from https://www.gov.pl/web/uw-warminsko-mazurski/rejestry-prowadzone-w-wydziale-polityki-spolecznej

Wielkopolski Urząd Wojewódzki. (2023). Wydział Polityki Społecznej. Retrieved October 8, 2023 from https://www.poznan.uw.gov.pl/rejestry-ewidencje-i-archiwa/wydzial-polityki-spolecznej

Zachodniopomorski Urząd Wojewódzki. (2023). Placówki zapewniające całodobową opiekę osobom niepełnosprawnym, przewlekle chorym lub w podeszłym wieku. Retrieved October 8, 2023 from https://www.gov.pl/web/uw-zachodniopomorski/placowki-zapewniajace-calodobowa-opieke-osobom-niepelnosprawnym-przewlekle-chorym-lub-w-podeszlym-wieku

Aims and Scope

The **Economics and Business Review** is a quarterly journal focusing on theoretical, empirical and applied research in the fields of Economics and Corporate and Public Finance. The Journal welcomes the submission of high quality articles dealing with micro, mezzo and macro issues well founded in modern theories and relevant to an international audience. The EBR's goal is to provide a platform for academicians all over the world to share, discuss and integrate state-of-the-art Economics and Finance thinking with special focus on new market economies.

The manuscript

1. Articles submitted for publication in the **Economics and Business Review** should contain original, unpublished work not submitted for publication elsewhere.
2. Manuscripts intended for publication should be written in English, edited in Word in accordance with the **APA editorial** guidelines and sent to: secretary@ebr.edu.pl**.** Authors should upload two versions of their manuscript. One should be a complete text, while in the second all document information identifying the author(s) should be removed from papers to allow them to be sent to anonymous referees.
3. Manuscripts are to be typewritten in **12' font in A4 paper** format, one and half spaced and be aligned. Pages should be numbered. Maximum size of the paper should be up to 20 pages.
4. Papers should have an abstract of about 100-150 words, keywords and the Journal of Economic Literature classification code (**JEL Codes**).
5. Authors should clearly declare the aim(s) of the paper. Papers should be divided into numbered (in Arabic numerals) sections.
6. **Acknowledgements** and references to grants, affiliations, postal and e-mail addresses, etc. should appear as a separate footnote to the author's name a, b, etc and should not be included in the main list of footnotes.
7. **Footnotes** should be listed consecutively throughout the text in Arabic numerals. Cross-references should refer to particular section numbers: e.g.: See Section 1.4.
8. **Quoted texts** of more than 40 words should be separated from the main body by a four-spaced indentation of the margin as a block.
9. **References** The EBR 2022 editorial style is based on the **7th editio**n of the Publication Manual of the American Psychological Association (**APA**). For more information see APA Style used in EBR guidelines.
10. **Copyrights** will be established in the name of the **EBR publisher**, namely the Poznań University of Economics and Business Press.

More information and advice on the suitability and formats of manuscripts can be obtained from:

Economics and Business Review
al. Niepodległości 10
61-875 Poznań
Poland
e-mail: secretary@ebr.edu.pl
www.ebr.edu.pl

Subscription

Economics and Business Review (EBR) is published quarterly and is the successor to the Poznań University of Economics Review. The EBR is published by the Poznań University of Economics and Business Press.

Economics and Business Review is indexed and distributed in Scopus, Claritave Analytics, DOAJ, ERIH plus, ProQuest, EBSCO, CEJSH, BazEcon, Index Copernicus and De Gruyter Open (Sciendo).

Subscription rates for the print version of the EBR: institutions: 1 year – €50.00; individuals: 1 year – €25.00. Single copies: institutions – €15.00; individuals – €10.00. The EBR on-line edition is free of charge.