Vol. 15 No. 2 / 2024
ISSN: 1847-9375

Business Systems Research

A Systems View accross Technology & Economics

Special Issue: "Novel Solutions and Novel Approaches in Operational Research" co-published with the Slovenian Society INFORMATIKA – Section for Operational Research (SSI-SOR)

Impressum

Focus and Scope

Business Systems Research Journal (BSR) is an international scientific journal focused on improving the competitiveness of businesses and economic systems. BSR examines a wide variety of decisions, processes, and activities within the actual business setting and the systems approach framework. Theoretical and empirical advances in business systems research are evaluated regularly. Special attention is paid to educational, social, legal, and managerial aspects of business systems research. In this respect, the BSR journal fosters the exchange of ideas, experience, and knowledge between regions with different technological and cultural traditions, in particular in transition countries. Papers submitted for publication should be original theoretical and practical papers. The journal also publishes case studies describing innovative applications and critical reviews of theory.

Abstracted/indexed in:

Baidu Scholar, Cabell's Whitelist, CEJSH (The Central European Journal of Social Sciences and Humanities), CNKI Scholar, CNPIEC – cnpLINKer, Dimensions, DOAJ, EBSCO (relevant databases), EBSCO Discovery Service, EconBiz, Engineering Village, ERIH PLUS (European Reference Index for the Humanities and Social Sciences), Genamics JournalSeek, Google Scholar, Hrcak, Inspec, J-Gate, JournalGuide, JournalTOCs, Microsoft Academic, Naver Academic, Naviga (Softweco), Norwegian Register for Scientific Journals, Series and Publishers, Primo Central (ExLibris), ProQuest (relevant databases), Publons, QOAM (Quality Open Access Market), ReadCube, Research Papers in Economics (RePEc), SCImago (SJR), SCOPUS, Semantic Scholar, Sherpa/RoMEO, Summon (ProQuest), TDNet, TEMA Technik und Management, Ulrich's Periodicals Directory/ulrichsweb, WanFang Data, Web of Science - Emerging Sources Citation Index, WorldCat (OCLC)

Publisher
IRENET, Society for Advancing Innovation and Research in Economy

ISSN
Business systems research (Online) = ISSN 1847-9375

Editorial Office
e-mail: ***bsreditors@gmail.com***
Web:
http://***www.bsrjournal.org***; http://***hrcak.srce.hr/bsr***; https://***sciendo.com/journal/BSRJ***

Business Systems Research

Special issue in Novel Solutions and Novel Approaches in Operational Research

Special Issue Articles

Editorial for the Special Issue: "Novel Solutions and Novel Approaches in Operational Research"

co-published with the ***Slovenian Society INFORMATIKA – Section for Operational Research (SSI-SOR)***

Samo Drobne
University of Ljubljana, Faculty of Civic and Geodetic Engineering, Ljubljana, Slovenia
Lidija Zadnik Stirn
University of Ljubljana, Biotechnical Faculty, Ljubljana, Slovenia
Berislav Žmuk
University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia

Abstract

This special issue of Business Systems Research (SI of the BSR) is being co-published by the Slovenian Society INFORMATIKA – Section for Operational Research (SSI -SOR). It focuses on recent advances in Operations Research and Management Science (OR / MS), with a particular emphasis on linking OR / MS with other areas of quantitative and qualitative methods in the context of a multidisciplinary framework. The ten papers that were chosen for this Special Issue of the BSR present advancements and new techniques (methodology) in the field of Operations Research (OR), as well as their application in a variety of fields, including risk management, mathematical programming, game theory, gravity, spatial analysis, logistics, circular economy, continuous improvement, sustainability, e-commerce, forecasting, Gaussian processes, linear regression, multi-layer perceptron, and machine learning.

Keywords: gravity model; multi-criteria decision-making; inconsistent pairwise comparisons; semiparametric spatial autoregression; cluster analysis; greenhouse gas emissions; maritime transport; internal migration; investor sentiment; age management.

JEL classification: C44, C61, C63, D81
Paper type: Editorial

Citation: Drobne, S., Zadnik Stirn, L., Žmuk, B. (2024). Editorial for the special issue: "Novel solutions and novel approaches in Operational Research", Business Systems Research, 15(2), 1-7.
DOI: doi.org/10.2478/bsrj-2024-0013

Editorial process

Operations research (OR), sometimes referred to as management science (MS), is a systematic methodology for decision-making that seeks to determine the most efficient means of constructing or managing a system, particularly in contexts necessitating resource allocation (Winston, 2003; Boucherie et al., 2021). Operations Research (OR) is a discipline dedicated to decision support, primarily aimed at creating tools and methodologies that assist decision-makers in problem-solving and judgement formation. Data analysis, simulation, modelling methodologies, and software applications are the decision-support topics of Operations Research (Mladenić et al., 2003; Rubio and Jiménez-Parra, 2014).

Operations Research (OR) is applicable in several domains, such as industrial engineering, management, business, manufacturing, government, healthcare, transportation, geographic information systems, scheduling, marketing, inventory, and other sectors (Cochran et al., 2011). The applications of Operations Research facilitate the clear and adaptable articulation of complex issues within a real-world setting, integrating both quantitative (e.g., financial ratios) and qualitative aspects into the evaluation process (Figueira et al., 2005).

The application of Operations Research has substantially enhanced organisational efficiency, resulting in increased productivity and social welfare. The International Federation of Operational Research Societies (IFORS) and the Association of European Operational Research Societies (EURO) serve as umbrella bodies for operational research societies globally. Collectively, these two associations encompass over 50 national societies, including the Slovenian Society INFORMATIKA - Section for Operational Research (SSI-SOR). SSI-SOR primarily concentrates on the organisation and coordination of international symposia.

The 17th International Symposium on Operations Research, often called SOR'23, was held in Bled, Slovenia, from September 20th to September 22nd, 2023. SOR'23 was a scientific event in the field of Operations Research. It was another in the usual series of international OR conferences held every two years and hosted by SSI-SOR in Slovenia. The main goal of SOR'23 was to promote the knowledge, interest and education of OR in Slovenia, Europe, and worldwide. In addition, it was agreed at SSI-SOR to collaborate with other disciplines to find a middle ground between the breadth of theoretical knowledge in OR and the understanding of theory, techniques, and problems in other disciplines, both within and outside OR. SOR'23 was attended by 198 individuals from different research institutes, universities, government agencies, private and public companies, and 19 countries worldwide, both in person and online. The conference included 96 papers or abstracts, with 198 authors and co-authors contributing to their production. After a blind peer review process conducted by two independent reviewers from the SOR'23 Programme Committee and reviewers nominated by SSI-SOR, the articles were finally approved for publication.

As a result of the decision taken at SOR'23 to publish the special issue (SI) of the BSR, the call for papers for this SI was already published at this conference in Bled in September 2023. The invitation was addressed to all those who had registered for SOR'23 and to other researchers from the field of OR. The submitted papers should present current breakthroughs and novel approaches in OR methodologies and models and their practical applications in economics, business, finance, organisation, management, social sciences, environment, agriculture, education and transport, among others.

Fourteen submissions were received. Several articles are expanded versions of concise papers delivered at SOR’23 and subsequently published in the Proceedings of SOR’23 (Drobne et al., 2023). The submissions to the BSR's Special Issue were initially assessed anonymously by the guest editors and subsequently by two experts. This special issue of the BSR has 10 distinct essays from various writers. They have continually emphasised model development and modelling, enhancing their practical orientation. Furthermore, they exceed mere algorithm presentation by enhancing them with the newest advancements in optimisation, simulation, and decision analysis.

The selected contributions span various developments and techniques in operations research (OR), economics, spatial science, and business, with practical applications across different sectors. Methodologically, the papers cover topics such as gravity models, multi-criteria decision-making, semiparametric spatial autoregression, cluster analysis, behavioural finance, and mathematical programming. These approaches are applied to pressing issues like environmental sustainability, labour market dynamics, age management, agricultural policies, and market asymmetries. Case studies include a wide geographic range, with contributions focusing on the European Union, Slovenia, Croatia, Central and Eastern Europe, and beyond. Together, these papers provide interdisciplinary insights into economic, social, and environmental challenges, offering policy recommendations and strategic frameworks for improving sustainability and efficiency across various domains.

The achievements of the BSR's SI are due to the collective work that has been done. The guest editors would like to thank the authors for their thoughtful and well-written contributions and the reviewers for their careful consideration of the contributions and their insightful and helpful comments. Last but not least, the guest editors would like to express their sincere gratitude and appreciation to the Editor-in-Chief, Professor Mirjana Pejić Bach, PhD, for asking us to serve as guest editors of the BSR's SI.

Contributions

The purpose of the papers published in BSR, following BSR's objectives and editorial policy, is to present original theoretical and empirical advances in business and economic systems using a wide range of methodological approaches, mainly from operations research/analytics, management science, and statistics. This is done to fulfil BSR's mission and comply with BSR's editorial policy. These objectives have been achieved with ten papers BSR has accepted for this SI.

In the first paper, entitled "The Impact of a New Container Port on the Greenhouse Gas Pollution", Bogataj, Campuzano-Bolarin, and Moreno Nicolás examine how the construction of a new container terminal affects greenhouse gas emissions, including CO2. Using a gravity model within a global supply chain, the study evaluates the environmental impact of maritime transport and the added emissions from poor railway connections. Results indicate that while the new terminal can reduce emissions by cutting time spent on routes and in ports, insufficient rail infrastructure undermines these gains. The authors show how the optimal capacity of a new port can be calculated using a multi-level gravity model, predicting its effect on pollution both at the port and along transportation routes to consumers.

In the article "Selection Procedure of the Approximation Methods for Deriving Priorities: A Case of Inconsistent Pairwise Comparisons", Vesna Čančer presents a method to evaluate the accuracy of approximation methods used in multi-criteria decision-making. The study compares the results of different approximation methods to the eigenvalue method, particularly in cases where pairwise comparisons exhibit inconsistency. Using mean absolute deviation (MAD) and mean absolute percentage deviation (MAPD), the research shows that the geometric mean method is the most accurate for deriving priorities. This paper contributes to decision-making literature by guiding users on choosing appropriate methods when deriving priorities from inconsistent comparisons.

In the third article "Design of Social Infrastructure and Services Taking into Account Internal Migration by Age Cohort", Samo Drobne and Marija Bogataj explore the impact of demographic changes on urban and regional planning, focusing on migration patterns by age cohort. Using the gravity model applied at Slovenian NUTS 2 and NUTS 3 spatial levels, the study investigates how factors such as distance, wages, and the capacity of care homes influence internal migration. Key findings reveal that distance is less significant for those aged 65-74, while wages mainly affect younger cohorts. The study emphasizes the need for adapted social infrastructure and services, particularly for older adults, to support the growing silver economy.

In the fourth article "Beyond Parametric Bounds: Exploring Regional Unemployment Patterns Using Semiparametric Spatial Autoregression", Andrea Furková and Peter Knížat address the limitations of traditional econometric models in capturing nonlinear relationships between economic variables. The authors propose a semiparametric spatial autoregressive model to account for both spatial effects and nonlinearities. Using penalised basis splines, they demonstrate how this approach provides greater flexibility in modelling local variations. The empirical study, focusing on regional unemployment across the European Union, reveals significant spatial dependence between regions. The authors conclude that semiparametric models offer improved accuracy over parametric models, providing a better understanding of regional unemployment dynamics.

In the fifth article "Sentiment and Stock Characteristics: Comprehensive Study of Individual Investor Influence on Returns, Volatility, and Trading Volumes", Aleš Kresta, Jialei Xiong, and Bahate Maidiya explore the relationship between individual investor sentiment and stock market variables. Using data from the American Association of Individual Investors (AAII) sentiment index, the study examines 480 components of the Standard & Poor's 500 index. The findings show a positive correlation between investor sentiment and stock returns, while a negative relationship is observed with volatility and trading volume. The research highlights the significant influence of individual investor sentiment on market behaviour, contributing to the field of behavioural finance.

In the sixth article "Age Management Practices and Benefits in Organisation: An Evaluation of the Effect of Economic Sector, Organisation Size, and Family Business Status", Terezie Krestová, Aleš Kresta, and Lucie Bestová examine the factors influencing age management practices within organisations. The study explores whether the economic sector, organisation size, and family business status affect the selection of age management strategies and their observed benefits. Based on a questionnaire survey and chi-square tests, the authors find that the economic sector impacts the choice of age management practices, while the organisation's size affects the benefits gained. The findings offer practical guidance for organisations in selecting appropriate age management strategies to maximize potential outcomes.

In the seventh article, "School-to-Work Transition in the Youth Labor Market in Central and Eastern Europe: A Cluster Analysis Approach", Tomislav Korotaj, James Ming Chen, and Nataša Kurnoga explore youth labour market dynamics in eleven central and Eastern European countries from 2008 to 2021. Using hierarchical clustering and multidimensional scaling, the study evaluates wage ratios, early departure from education or training, and the share of youth not in employment, education, or training. The results reveal distinct clusters, including the Visegrád countries, the Baltics, and the Balkans, each facing unique challenges. The findings emphasize the role of historical and geographical ties in shaping youth labour market outcomes, with Poland, Estonia, and Bulgaria emerging as outliers in their respective regions.

In the eighth article, "Strategic Categorization of Dairy Cow Farms in Croatia using Cluster Analysis", Maja Petrač, Krunoslav Zmaić, and Jaka Žgajnar examine the challenges faced by the milk processing sector in Croatia, characterized by a bipolar structure with large-scale producers and struggling small to medium-sized farms. The study aims to categorize typical dairy farms using cluster analysis to provide insights for policy formulation. Employing both hierarchical and non-hierarchical clustering techniques on data from the Croatian Agency for Agriculture and Food, the researchers identified 16 clusters of relatively homogeneous farms. These clusters offer a detailed understanding of Croatia's dairy sector and provide a foundation for targeted investments aimed at improving farm efficiency, economic viability, and sustainability.

In the ninth article "Estimating Asymmetric Fuel Price Responses in Croatia", Karol Szomolányi, Martin Lukáčik, and Adriana Lukáčiková examine the asymmetry in the transmission of oil prices to retail fuel prices in Croatia. The study employs various econometric models, including the Linex approach, which uses a non-linear adjustment cost function and the generalised method of moments. Results from standard methods are mixed, while the Linex approach reveals clear price asymmetries, particularly in how fuel prices rise faster than they fall. The study concludes that the Linex approach is effective in detecting price asymmetries, even in large datasets with frequent price changes, aligning with findings from similar studies worldwide.

In the tenth article "Decision-Making Model to Support Agricultural Policies in Realizing Economic and Social Sustainability", Jaka Žgajnar and Lidija Zadnik Stirn present a mathematical programming model to aid in the development of the Common Agricultural Policy (CAP) Strategic Plan for 2023-2027, focusing on the beef sector. The linear programming model enables ex-ante analysis by simulating farm-level production plans and aggregating outcomes at the sector level. The study highlights the challenges faced by the beef sector, including high costs, low efficiency, and structural limitations. Results show that production-related payments are crucial to mitigate negative trends, particularly for larger farms, underscoring the model's effectiveness in supporting policy design for enhanced economic and social sustainability.

It can be concluded that the high-quality and timely topics of the SI of BSR papers are of interest to both scientific and professional audiences, as they may influence both theory and applications.

Ljubljana, Zagreb, September 2024

Guest Editors of SI BSR

Samo Drobne
Lidija Zadnik Stirn
Berislav Žmuk

References

1. Boucherie, R. J., Braaksma, A., & Tijms, H. C. (2021). *Operations Research: Introduction to Models and Methods*, World Scientific, New Jersey.
2. Cochran, J. J., Cox, L.A., Keskinocak, P., Kharoufeh, J. P., Smith, C. (2011). *Wiley Encyclopedia of Operations Research and Management Science*, Wiley, New Jersey.
3. Drobne, S., Zadnik Stirn, L., Kljajić Borštnar, M., Povh, J., & Žerovnik, J. (2023). *The 17th International Symposium on Operational Research SOR'23, Proceedings*, Bled, Slovenia, September 20-22, 2023. Ljubljana: Slovenian Society Informatika (SSI). Section for Operational Research (SOR).
4. Figueira, J., Greco, S., & Ehrgott, M. (2005). *Multicriteria Decision Analysis*, Springer, New York.
5. Mladenić, D., Lavrač, N., Bohanec, M., & Moyle, S. (2003). *Data Mining and Decision Support*, Kluwer, Boston.
6. Rubio, S., & Jiménez-Parra, B. (2014). Reverse logistics: Overview and challenges for supply chain management, *International Journal of Engineering Business Management*, 6(6-12), 1-7.
7. Winston, W. L. (2003). *Operations Research: Applications and Algorithms*, Cengage Learning, Duxbury.

About the editors

Samo Drobne is a member of the Management Board of Slovenian Society INFORMATIKA – Section for Operational Research (SSI-SOR); Secretary of Slovenian Society INFORMATIKA – Section for Operational Research (SSI-SOR); Co-editor of several proceedings of the international symposia on operations research in Slovenia (Proceedings of SOR); Member of Editorial board of Geodetski Vestnik – an open access journal of the Association of Surveyors of Slovenia; Member of Editorial board of Journal Communications - Scientific Letters of the University of Žilina. The editor can be contacted at **samo.drobne@fgg.uni-lj.si.**

Lidija Zadnik Stirn is a president of the Slovenian Society INFORMATIKA – Section for Operational Research (SSI-SOR); member of the board of the Slovenian Society INFORMATIKA (SSI); Representative of SSI-SOR in International Federation of Operational Research Societies (IFORS); Representative of SSI-SOR in Association of European Operational Research Societies (EURO); Co-editor of Central European Journal of Operations Research (CEJOR); Co-editor of several proceedings of the international symposia on operations research in Slovenia (Proceedings of SOR); Member of Editorial board of Croatian Operational Research Review (CrORR). The editor can be contacted at **lidija.zadnik@bf.uni-lj.si.**

Berislav Žmuk is an Associate Professor who graduated with a major in Accounting, postgraduate Statistical Methods for Economic Analysis and Forecasting, and gained his PhD in Business Economics at the Faculty of Economics and Business, University of Zagreb. Currently, he is an Associate Professor at the Department of Statistics, Faculty of Economics and Business, the University of Zagreb, where he teaches the following subjects: Statistics, Business Statistics, Business Forecasting and Introduction to Economic Statistics. He served as a president of the Croatian Statistical Association. His main research fields include applications of statistics in business and economy, survey methodology, and statistical quality control. The editor can be contacted at **bzmuk@efzg.hr**.

The Impact of a New Container Port on the Greenhouse Gas Pollution

David Bogataj
Group of Numerical Simulations in Industrial Engineering, Campus Muralla del Mar, Polytechnic University of Cartagena, Spain
Alma Mater Europaea- ECM, Slovenia
Francisco Campuzano-Bolarin, José Andrés Moreno Nicolás
Group of Numerical Simulations in Industrial Engineering, Campus Muralla del Mar, Polytechnic University of Cartagena, Spain

Abstract

Background: Large vessels that call at European ports will have to pay for their CO2 emissions from transporting cargo that enters or goes from a European port since January 2024. The costs will increase with increasing global trade. This results in a higher pollution level, including greenhouse gas (GHG) emissions like CO2. **Methods/ Approach:** Based on the gravity model embedded in a global supply chain, we developed a model to evaluate maritime transport pollution in case a new, sufficiently large container port becomes operational. Additionally, we consider how lousy railway connections to European customers increase transportation costs and pollution. **Results:** The approach to the well-connected sequences of gravity models in the intercontinental maritime chains evaluates the differences in quantities of cargo between ports when a new port is opened, and the waiting time does not change. We also highlight that poor rail connections can reduce this positive effect. **Conclusions:** We showed how it is possible to estimate the optimal capacity of a new port with a multi-level gravity model and how this would affect the pollution around the port and on the routes from the port to the final consumers.

Keywords: gravity model; supply chain; seaport; pollution; intercontinental transport; container terminal; railway; CO2 emissions

JEL classification: A13; C10; J11; J26
Paper type: Research article

Received: Jan 30, 2024
Accepted: Jul 15, 2024

Acknowledgements: The research leading to these results received funding from the grant PDC2022-133957-I00 funded by MCIN/AEI /10.13039/501100011033 and by European Union Next Generation EU/PRTR.

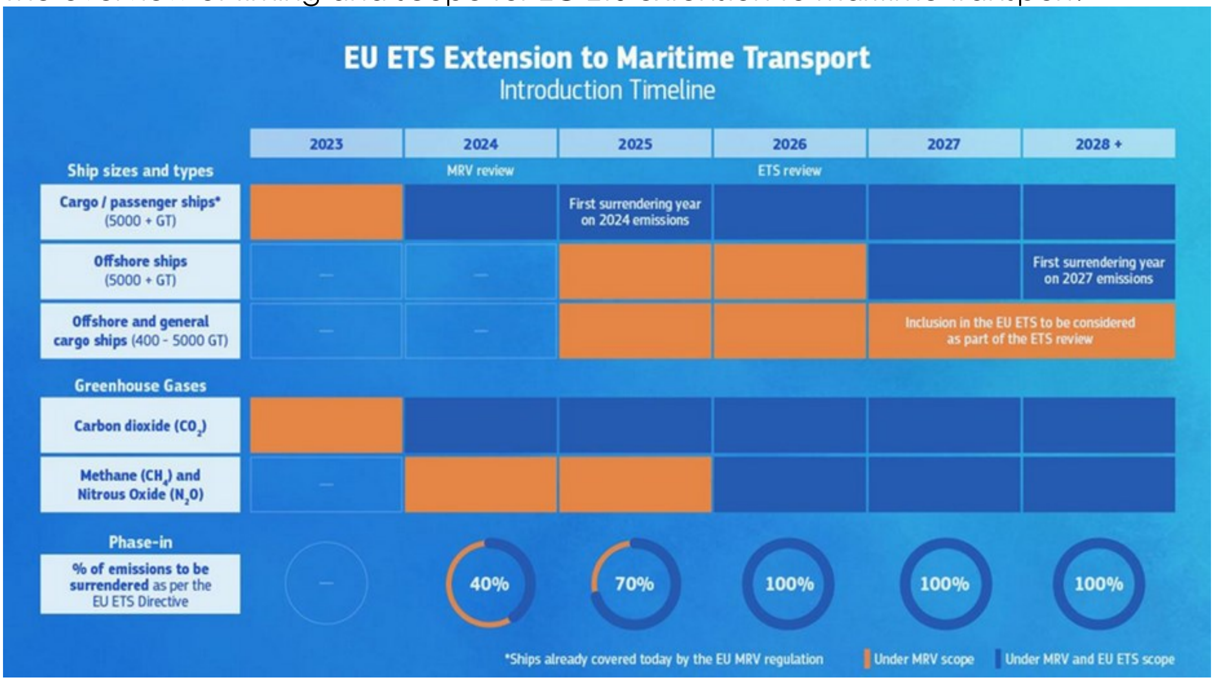
Citation: Bogataj, D., Campuzano-Bolarin, F., & Moreno Nicolás, J. A. (2024). The Impact of a New Container Port on the Greenhouse Gas Pollution. *Business Systems Research, 15*(2), 8-20.
DOI: doi.org/10.2478/bsrj-2024-0014

Introduction

The European Union Emissions Trading System (EU ETS), including the MRV (monitoring, reporting and verification) of vessels' emissions, is becoming one of the EU's main tools for reducing greenhouse gas emissions in maritime transport. The Commission has adopted several implementing acts and directives that determine the necessary rules and methods for the system's successful operation to ensure the timely inclusion of maritime transport in the EU ETS as early as January 2024 (European Commission, 2024).

Sea routes and ports strongly impact pollution, especially with exhaust gases. For example, the results of Barberi et al. (2021) demonstrate the contingency between emissions and port infrastructures and show how proper investments can help reduce pollutant emissions in ports and maritime transportation in general. Port emissions are generated by manoeuvring and waiting in lines, including loading and unloading of ships, onshore operations using energy for loading, unloading, and warehousing cargo, and port vehicle traffic. Barberi et al. (2021) reported that estimating emissions in ports can be tricky due to the multitude of co-existing sources. The exhaust emissions from ships in maritime transport are calculated using activity-based and fuel-based methods.

Figure 1
The overview of timing and scope for EC ETS extension to maritime transport.



Source: European Commission, 2024

In our paper, we shall assume that the fuel structure and activity-based structure on the lines from other continents to Europe and from Europe to other continents do not change and that only the length of all sea routes and, therefore, the time spent on routes and in ports is reduced by investments and operations of a new port. The exhaust emission in the ports will be calculated as in Barberi et al. (2021), considering the same structure of activities as in the Chinese port that is also on the main intercontinental road of vessels (see Bogataj et al., 2024). The paper is an extended version of the paper presented at the SOR 2023 conference (Bogataj, D., & Campuzano-Bolarín, 2023).

The paper consists of six sections. After the short literature review in the Introduction, the method to evaluate the location advantages of a new port for international cargo is presented in the second section. The method of calculating CO_2 and other emissions in case of additional ports is presented in the third section. The fourth section describes challenges for equalising technical standards on Spanish and other European railways. The last section gives the conclusions and some directions for further study.

Literature review

Gravity models (GMs) have been known in social science since 1963 (see the literature review of Bınová, 2015), but as a sequence of sub-models in a model of the supply chain was presented first in 2018 (see Table 1). The idea of a three-stage sequencing of GM is given only in Bogataj et al. (2022, 2024). The articles of Bogataj et al. also tie all gravity models together well. The Web of Science Core Collection (WoS) found the first article linking the topics of "supply chain" (SC) and the "gravity model" (GM) only at the beginning of the second decade of this millennium, and only 41 articles followed the first publication. Among them, only three papers include the topic "maritime transport" to the topic "supply chain" (Chang et al., 2021; Gani, 2021; Randrianarisoa and Gillen, 2022). In the basic approach, the formulation of GM is as follows:

$$T_{ij} = \alpha P_i P_j / d_{ij}^2 \qquad (1)$$

where T_{ij} is the flow between origin j and destination i and d_{ij} is the distance between the origin and destination, P_i and P_j are the sizes of populations in the destination and origin. In some recent papers, authors replaced P_i and P_j by gross domestic product GDP_i and GDP_j and later by logistics quality indicator (M) as the product of the logistics capability parameter τ_i and the logistics volume L_i ($M_i = \tau_i L_i$ and $M_j = \tau_j L_j$) of hubs, capturing reliability, responsiveness, assurance, empathy, and tangibles. However, power in such models is still equal to 1 or at least symmetrical in many papers (Wei and Lee, 2021). In recent papers, the GDP_i was replaced by production quantity Q_i and GDP_j by attractiveness A_j of a port:

$$T_{ij} = \alpha Q_i^{\beta} A_j^{\gamma} / d_{ij}^{\delta} \qquad (2)$$

We will say that (2) is the basic equation also if Q and A are replaced by GDP_i, $\tau_i L_i$ or P_i . In the models of supply chains except by Bogataj et al. (2022, 2024) and Randrianarisoa and Gillen (2022), the distance d_{ij} has remained geographical. However, in papers considering pollution in a supply chain, like in Randrianarisoa and Gillen (2022), we can find first explanations why it is better to replace geographical distance (road, train, or Euclidean distance) with the time-consuming distance in sea routes and in waiting lines in a port, like in some GMs of population migrations and commuting (Drobne and Bogataj, 2014, 2015, 2017, 2020, Bogataj and Drobne, 2011, Drobne et al, 2011, Janež et al., 2016, 2018). The success of container transport is due to the reduction in the duration of port calls, which means the reduction of waiting and service time in ports (Slack et al., 2018).

Although vessels now spend less time in port than at sea, it is still a cost and pollution factor. Randrianarisoa and Gillen (2022) studied only one stage of GM, which cannot answer the questions on the impact of a new node in a sea chain. Their paper proposes a novel way to reduce sulphur emissions in international transport by focusing on the role of logistics improvements and vessel size. When transportation intensity increases, pollution is calculated by adding a new number of vessels. The basic GM with the modified factors is extended. There are no solutions about where to locate a new node to reduce pollution. GDPs per capita and not GDPs itself at origin and destination have been two basic variables in the cited paper. The accessibility factor is divided into the natural and created part, where the geo-distance is part of this structure. The time spent factor is added extra and therefore is less fit for our purposes. Only four articles in WoS deal with a supply chain with two consecutive sections (Table 1).

Table 1
Articles in WOS with topics »supply chain" and »gravity model" where authors also mention maritime transport or any other sea transport in two or more stages

Year	Author	Mode	Purposes and methods	Stages
2018	Liu et al.	Any	A food-based virtual blue, green, and grey water network is built here. The gravity model and linear programming optimisation were used to model the interprovincial trade of Chin. There is no question about the location of a node.	2
2018	Wei et al.	Rail and road transport	Two-stage GM of inland logistics refers to the connectivity between the Maritime Silk Road and the Silk Road Economic Belt. There is no question about where to locate a seaport.	2
2021	Wei and Lee	Railway Express	This study employs a hybrid method with an improved entropy-weighted TOPSIS, a matching logistics capability and demand model, and a cross-border logistics gravity model considering the Silk Road Economic Belt. There is no question about where to locate a node or the impact of its capacity.	2
2022	Bogataj et al.	Vessel	The article shows how to improve the supply chain if a new port is included in the supply of perishable goods.	3

Source Web of Science

According to the findings of Nordås and Piermartini (2004), the log expression of GM of cargo (one-stage model) from j to i (M_{ij}) can be presented as follows:

$$M_{ij} = a_0 \cdot y_i^{a_1} \cdot y_j^{a_2} \cdot d_{ij}^{a_3} \cdot border_{ij}^{a_4} \cdot lang_{ij}^{a_5} \cdot island_{ij}^{a_6} \cdot landlock_{ij}^{a_7} \cdot (1+t)_{ij}^{a_8} \cdot infr_i^{a_9} \cdot infr_j^{a_{10}} \cdot T_i^{a_{11}} \cdot T_j^{*\,a_{12}} \cdot lat_i^{a_{13}} \cdot lat_j^{a_{14}} \quad (3)$$

where the first three variables give a relatively high correlation coefficient, and according to the high correlation of time distance with other variables (Rietveld et al., 1999), the other variable could be included in the distance if we replace the geographical distance d_{ij} with the time-spending distance τ_{ij} (time assumed in direct proportion to distance). In this case, we leave the velocity factor v, which is equal in all directions, at the intercept: $a = v \cdot a_0$. In this case, we may write:

$$M_{ij} = (a_0 v) \cdot y_i^{a_1} \cdot y_j^{a_2} \cdot \tau_{ij}^{a_3} \rightarrow lnM_{ij} = \ln a + a_1 \ln y_i + a_2 \ln y_j + a_3 \ln \tau_{ij} \quad (4)$$

Methodology for evaluation of port advantages

The starting point of this research is the work conducted by Nordås and Piermartini (2004), who estimated the powers (regression coefficients of linearisation) in maritime and road transport as given in Table 2. Regarding high R^2 for regressions in Table 2, we used these values to evaluate the potential flow of containerised cargo through Cartagena port when constructed, and flows will optimally adapt to the new network possibilities.

Table 2
Value of main regression coefficients in the gravity model (p < 0.01).

Regression coefficient	Indicator	Road infrastructure	Maritime infrastructure
a1	GDP importer	0.94	0.80
a2	GDP exporter	1.12	0.91
a3	distance between I and j	-1.22	-0.71
adjusted R2		> 0.65	> 0.70

Source: Nordås and Piermartini (2004)

Continuing the procedure like in the article by Bogataj et al. (2024), in this paper, we added the waiting time for cargo unloading from vessels and loading on trucks as τ_k ; when revised (2) to (5):

$$M_{ki} = a \cdot y_k^{0.80} \cdot y_i^{0.91} \cdot (\tau_{ki} + \beta\tau_k)^{-0.71} \quad (5)$$

and for road transport from port k to EU city l:

$$M_{lk} = b \cdot y_l^{0.94} \cdot y_k^{1.12} \cdot (\tau_{kl} + \gamma\tau_k)^{-0.71} \ . \quad (6)$$

In (5) and (6) β and γ are the proportionality coefficient between the costs of the hour of waiting and unloading or loading in the port and the hours of sailing of the container ship and truck, respectively, per unit of cargo. We may take the average waiting and service time of ships at ports of some EU and neighbouring countries from UNCTAD (2018) and Slack et al. (2018), wherein port call and performance statistics, including time spent at ports, are given. In our numerical example, we used $\beta = 1$ and $\gamma = 1$. World merchandise trade volume 2000 - 2021, including projections, are available at WTO (2021). Sea time distances between the biggest ports from a list of ports on other continents and EU ports are available at the World Shipping Council (2021). We took the data for the year 2021 and their time series. The distances between EU central places and EU ports are available at Google Maps. Time distances for trucks have been calculated under the assumption that the average speed of trucks is 70 km/h.

The average waiting and service time of ships at ports of the EU countries is grouped in our paper in three groups: group I: 0.3-0.6 days: Gibraltar, Norway, and Denmark; group II: 0.61-0.8 days: Croatia, Sweden, Lithuania, Slovenia, Spain, Portugal, Poland, Finland, Latvia, France Netherlands; group III: 0.81+ days: others. 13 the biggest ports have been considered. In Table 3, we calculated the average time distances of 1 twenty-foot-long container (TEU) in the case of $(\boldsymbol{\tau_{lk}} + \boldsymbol{\tau_k})$ o the EU continental member states (l) and the leading 12 EU ports (k), in case of being transported by trucks. From Table 3, based on the GM approach, the optimal cargo sharing between ports in 2039, if the port of Cartagena will be completed by then, is calculated as given in Table 4. Cartagena as the potential port was added.

Table 3
Time distance in hours between metropolises in EU continental member states and the main EU ports transported by truck*.

Port / City	CA	HB	BH	ANT	ZB	VA	AL	BA	LH	MA	RT	GD	PI
Vienna	38	17	17	19	20	33	44	30	24	24	21	29	29
Brussels	31	13	12	5	6	27	35	22	10	18	7	24	35
Sofia	45	32	32	35	36	43	53	38	40	40	36	30	33
Prague	37	13	14	17	19	33	43	27	21	23	17	16	32
Berlin	38	9	10	15	16	35	44	30	35	26	14	13	34
Copenhagen	44	11	13	19	21	40	48	35	24	31	18	19	43
Tallinn	61	34	35	40	41	57	68	53	44	48	38	23	54
Athens	58	41	41	44	45	55	65	50	47	45	44	40	30
Madrid	10	36	36	27	26	9	13	13	23	20	28	47	41
Helsinki	65	35	38	44	45	63	73	57	50	53	43	27	60
Paris	27	18	18	10	10	23	30	18	8	15	11	29	35
Zagreb	35	22	23	23	24	32	41	27	27	21	24	22	25
Budapest	40	20	21	23	25	36	45	31	27	25	24	19	29
Rome	32	28	29	26	28	28	38	24	27	18	28	32	13
Riga	55	28	30	34	36	53	63	48	40	43	33	18	50
Luxemburg	28	14	13	8	9	25	35	20	12	16	10	23	31
Vilnius	53	25	26	30	32	48	58	44	36	40	30	14	45
Amsterdam	34	11	10	7	8	29	37	25	14	21	6	21	36
Warsaw	38	17	18	23	24	41	51	36	28	31	22	10	38
Lisbon	17	40	40	32	32	17	13	21	28	26	34	50	47
Bucharest	53	33	33	36	37	48	58	43	40	38	37	31	40
Stockholm	53	20	22	29	30	50	58	44	34	40	28	27	51
Ljubljana	33	21	21	21	23	29	40	25	24	19	22	22	23
Bratislava	38	18	18	21	23	35	45	30	25	26	22	16	29

CA-Cartagena; BH-Bremerhaven; ANT-Antwerp, ZB-Zeebrugge; VA-Valencia; AL-Algeciras; HB-Hamburg; BA-Barcelona; LH-Le Havre; MA-Marseille; RT-Rotterdam; GD-Gdansk; PI-Piraeus
*Source: Authors' work based on Google Maps. *Waiting time in ports is added.*

Table 4
Expansion of the annual cargo handling by 2039 in the main EU ports [in 10^6 TEU].

If Cartagena Port is built						
Cartagena	Hamburg	Bremen h.	Antwerp	Zeebrugge	Valencia	Algeciras
7,9	5,5	9,2	-0,6	11,1	3,4	1,6
Barcelona	Le Havre	Marseilles	Rotterdam	Gdansk	Gioia Tauro	
6,1	9,5	9,2	0,1	8,9	5,1	
If Cartagena port will not be completed						
Cartagena	Hamburg	Bremen h.	Antwerp	Zeebrugge	Valencia	Algeciras
0,0	6,1	10,3	0,0	12,4	3,8	1,8
Barcelona	Le Havre	Marseilles	Rotterdam	Gdansk	Gioia Tauro	
6,8	10,6	10,2	0,1	9,9	5,7	

Source: Authors' work.

In Table 4, the data are in 10^6 TEU, as logistics and shipping companies usually measure the quantity of cargo. They load up to 24000 kilograms (24 metric tons) of cargo in a TEU. An empty container weighs 2280 kilograms (2.24 metric tons). Hence, the total weight of a fully laden twenty-foot container will be 26280 kilograms (26.28 metric tons. Bogataj et al. (2024) presented detailed methods and data derivation based on the sequential GM models of global supply chains.

Potential increase of yearly emissions in the EU's biggest ports by 2039 if Cartagena port would be completed

The optimal allocation of cargo along individual routes on the graph has been discussed countless times from the point of view of economic efficiency but less so from the point of view of pollution. However, the development and application of gravity models for these purposes were also less frequently used in these studies.

Using the GM approach, Oesingmann (2022) studied the impact of the EU ETS on aviation demand; Hintermann and Ludwig (2023) pointed to informational transaction costs that increase when trading across national borders, also showing the usefulness of the GM, but not exposing the maritime transport. Bart (2011) has shown that municipal emissions can be calculated as a share of total national road transport emissions with the support of a GM. Pothen and Hübler (2018) studied climate and trade policies, focusing on their interactions, and the GM approach was also explored. In WoS, however, we do not find any authors who would connect the pollution problem in maritime transport, or only in ports, with the newly introduced EU ETS in naval affairs. In our paper, following the data in Table 4, the expected pollution in the biggest EU ports can be estimated using the GM approach.

Let us take for the EU the biggest ports with the same structure of emissions of nearly 125,000 ships, like Qingdao in 2016 (Sun et al., 2018), where container ships' pollution is approximately 2.8/5. [NO_x, CO, HC, CO_2, SO_2, $PM_{2,5}$] = 10^3[16.7, 1.7, 0.7, 1,275, 11.9, 0,96].

Table 5
Increase the yearly emissions in the EU, the most prominent ports, by 2039 if the Cartagena port is completed and technology does not change [ton].

	Cartagena	Hamburg	Bremen h.	Antwerp	Zeebrugge	Valencia	Algeciras
NOx	7798	5415	9034	-579	10888	3355	1609
CO	805	559	932	-60	1124	346	166
HC	331	230	383	-25	462	142	68
CO2	596342	414064	690857	-44257	832629	256540	123019
SO2	5556	3858	6437	-412	7757	2390	1146
PM25	447	310	518	-33	624	192	92
	Barcelona	**Le Havre**	**Marseilles**	**Rotterdam**	**Gdansk**	**Gioia Tauro**	
NOx	6023	9368	8995	78	8760	5032	
CO	621	967	928	8	904	519	
HC	256	398	382	3	372	214	
CO2	460571	716361	687857	6001	669854	384810	
SO2	4291	6674	6409	56	6241	3585	
PM25	345	537	516	4	502	288	

Source: Sun et al. (2018) and own calculations.

Here, 1 TEU will be, on average, equal to nearly 30 metric tons. The yearly emission in 2039 (when the Cartagena port would be completed) in the EU's biggest ports will increase for the values in Table 5 if the technology does not improve. The realised data forecasted in Tables 4 and 5 will be subject to the MRV procedures established inside the European Emissions Trading System, as described in the Introduction.

Challenges for equalising technical standards on Spanish and other European railways

Since the average weight of a 20 FT TEU is about 2,500kg (a container 20 feet long, 8'0" wide, and usually 8'6" high), we can calculate the expected quantity of CO2 yearly emissions if the average reference emissions amount to 52.7 gram of CO2 per tonne-kilometre (gCO2/km), as reported by Transport & Environment 2021.

Table 6
Yearly CO_2 pollution on roads by trucks to the countries potentially using the Cartagena port by 2039. (Annual capacity in 10^6 kg)

Country	Distance [102 km]	MI. C [103 ton]	Pollution [ton] of CO2	Country	Distance [102 km]	MI. C [103ton]	Pollution [ton] of CO2
Austria	24	450	56,916	Italy	20	2,232	235,253
Belgium	19	667	66,787	Latvia	36	27	5,122
Bulgaria	30	65	10,277	Lithuania	34	42	7,526
Croatia	22	77	8,927	Luxembo urg	18	110	10,435
Czechia	23	272	32,969	Netherlan ds	21	1,005	111,223
Denmark	28	305	45,006	Poland	30	602	95,176
Estonia	39	4.5	4,624	Portugal	9	580	27,509
Finland	40	152	32,042	Romania	30	182	28,774
France	16	3,665	309,033	Slovak Rep.	25	170	22,398
Germany	25	3,495	46,0466	Slovenia	21	72	7,968
Greece	34	135	24,189	Spain	4.5	4,997	118,504
Hungary	25	170	22,398	Sweden	34	375	67,193

Source: EUROSTAT 2022, Via Michelin, 2024 – to the metropolis (rounded).

The results are given in Table 6, which shows that the yearly pollution of transporting the cargo from the port of Cartagena to individual EU member states would be approximately 1,8 million tons. It is approximately 1/3 of annual CO2 emission in the port of Cartagena, as given in Table 5. In equation (5) and other calculations, we assumed that cargo transport from European ports to EU member states will only be carried out by road. While this is almost essential for perishable goods, heavy, non-perishable cargo is more suitable for rail transport. Carried out by the road, *the average emissions amount to 52.7 grams of CO_2 per tonne-kilometre (gCO2/tkm) of cargo. These data constitute the overall baseline for the following year's* (2025) *emissions targets of a 15%* reduction in Europe (Transport & Environment, 2021).

Regarding the relative costs of pollution per tonne-kilometre transport of cargo, we can also expect, in the future, that the cost of rail transport will not exceed a quarter of the cost of road transport (Banfi et al., 2000). For rail transport, it is often expected to present only 10% of costs caused by road transport if there are no specific obstacles in the way. However, Spain has problems here, as the standards of its railway networks (the electrification technology and the track gauge) are incompatible with other European standards. It causes problems at the borders with France.

According to the EU Directives (European Commission, 2013), considering Regulations No 1316/2013 and 1316/2013, The freight corridor RFC6, which connects Southwestern European countries with Eastern European countries, belongs to the TEN-T network as defined (European Commission, 2013). RFC6, as all the lines of the TEN-T Corridors Core Network must be standardised regarding gauge (or mixed gauge) by 2030, it should have 25 kV AC electrification, a single European signalling system and train speed control ERTMS (European Rail Traffic Management System). The Global System for Railway Mobile Communications (GSM-R) was also introduced for TEN-T.

Figure 2
Cartagena port on the RCF6 corridor goes through Spain and France, continuing to Kyiv.

Source: Mediterranean Rail Freight Corridor (2024) and authors elaboration, AJOT (2023).

The Cartagena port would become one of the most important nodes inside the Mediterranean Corridor (MC), the longest in the European Union. MC ranks second in freight transport and serves countries representing 18% of the European population, but it is easily linked to other EU corridors. We calculated the advantages of cargo being transhipped at Cartagena for the case of road transport only, leading to the optimal solution, as given in Table 5. The Rail Freight Corridor 6 (RFC6), which belongs to MC, has two separate routes: the Central Mediterranean Corridor and the Coastal Mediterranean Corridor (see Figure 2). Cargo on these corridors must cross the French border if it continues to other European countries. RFC6 is the most connecting freight corridor, linked to freight corridors RFC1, 2, 3, 4, 5, 7, 10 and 11.

Crossing the border from Spain to France is costly due to the differences in the electrification technology and the track gauge that requires special transhipments. When the new EU technological standards for railways are also achieved in Spain, the advantage of the port of Cartagena over other ports, calculated based on the sequential gravity models on the global supply chain, will be definitive for all modes of transport.

Conclusion

To regulate the rising environmental impact of sea transport, the EU also integrated ports into the European Emissions Trading System (EU ETS) in 2024. Vessels' CO_2 emission is now included in the EU ETS, which is essential for reducing greenhouse gas emissions in maritime transport. That is why it is necessary, when deciding on constructing a new port or expanding old ports in the intercontinental goods transport network, to assess the expected pollution in the ports and the routes to the final consumer.

The optimal allocation of cargo along individual routes regarding pollution was rarely studied, especially in case of a development or application of gravity models. We found only four articles in the WOS that include both the "gravity model" and "Emission Trading System" (Oesingmann, 2022; Hintermann and Ludwig, 2023; Bart, 2011; and Pothen and Hübler, 2018). They studied environmental trade policies, and the GM approach was also used. However, nobody considered maritime transport and port pollution.

In the article, we showed how it is possible to estimate the optimal capacity of a new port with a multi-level gravity model and how this would affect the pollution around the port and on the routes from the port to the final consumers. In particular, we focused on the problem of uncoordinated technical standards of railways in Spain and elsewhere in Europe, which raises the cost of rail transport. The calculations assumed that the structure of ships in Mediterranean liner transport is similar to that of ships in Chinese ports on the same intercontinental line. When the technical standards are fulfilled, direct transport could reduce pollution on transport routes by as much as 70%-90%.

This paper has not considered future technical improvements that could influence pollution, which could be the subject of further research. From the data collected and displayed, it is possible to study the impact of pollution on health, the built and natural environment, and climate change, which is still a wide-open field of research.

References

1. AJOT. (2023). The first direct freight rail connection with France comes into operation at the Port of Barcelona. Ports & Terminals. Available at https://www.ajot.com/news/the-first-direct-freight-rail-connection-with-france-comes-into-operation-at-the-port-of-barcelona (November 9, 2023).
2. Maibach, M., & Banfi, S. (2000). External Costs of Train Transport Accident, Environmental and Congestion Costs in Western Europe. Zürich/Karlsruhe, INFAS, IWW. Available at: http://worldcat.org/isbn/274610184X (December 12, 2023)
3. Barberi, S., Sambito, M., Neduzha, L., & Severino, A. (2021). Pollutant Emissions in Ports: A Comprehensive Review. *Infrastructures, 6*(8). 114. https://doi.org/10.3390/infrastructures6080114
4. Bart, I. (2011). Municipal emissions trading: reducing transport emissions through cap-and-trade. *Climate Policy, 11*(1). 813-828. https://doi.org/10.3763/cpol.2009.0019
5. Bınová, H. (2015). Modified Method of Gravity Model Application for Transatlantic Air Transportation, *Neural Network World, 25*(2). 203-217. https://doi.org/10.14311/nnw.2015.25.011
6. Bogataj, M., & Drobne, S. (2011). Decision making in competitiveness of regions, *Croatian Operational Research Review, 2*, 194-207.
7. Bogataj, D., Bogataj, M., Campuzano-Bolarin, F., & Nicolás, J. A. M. (2022). Location Advantages of the Container Port for Perishable Goods in the Murcia Region. *IFAC-PapersOnLine, 55*(10). 2701-2706. https://doi.org/10.1016/j.ifacol.2022.10.118
8. Bogataj, D., & Campuzano-Bolarín, F. (2023). How Would the Construction of a New Container Port in Cartagena Influence the Reduction of Greenhouse Gas Pollution. In: Drobne et al. (Eds). SOR '23 Proceedings of 17th Symposium of Operational Research in Slovenia. Bled: SDI-SOR, 363-366.
9. Bogataj, D., Campuzano-Bolarin, F., Moreno Nicolás, J. A., & Sethi, S. P. (2024). Gravity modelling of intercontinental supply chains considering a new port location. *Central European Journal of Operations Research*. https://doi.org/10.1007/s10100-024-00909-0
10. Chang, Y.-T., Jo, A., Choi, K.-S., & Lee, S. (2021). Port efficiency and international trade in China. *Transportmetrica A: Transport Science, 17*(4). 801-823. https://doi.org/10.1080/23249935.2020.1817169

11. Drobne, S., & Bogataj, M. (2022). Migration Flows through the Lens of Human Resource Ageing. *Business Systems Research Journal, 13*(3). 47-62. https://doi.org/10.2478/bsrj-2022-0024
12. Drobne, S., & Bogataj, M. (2017). The impact of public investments in facilities on the potential housing market for older persons. *Facilities, 35*(7/8). 422-435. https://doi.org/10.1108/f-09-2016-0091
13. Drobne, S., & Bogataj, M. (2015). Optimal Allocation of Public Service Centres in the Central Places of Functional Regions. *IFAC-PapersOnLine, 48*(3). 2362-2367. https://doi.org/10.1016/j.ifacol.2015.06.441
14. Drobne, S., & Bogataj, M. (2014). Regions for Servicing Old People: Case study of Slovenia. *Business Systems Research Journal, 5*(3). 19-36. https://doi.org/10.2478/bsrj-2014-0017
15. Drobne, S; Bogataj, M; Zupan, M., & Lisec, A. (2011). Dynamics and local policy in commuting: attractiveness and stickiness of Slovenian municipalities, SOR'11 PROCEEDINGS: The 11th International Symposium on Operational Research in Slovenia, 323-328
16. European Commission. (2013). Trans-European Transport Network (TEN-T). Available at: https://transport.ec.europa.eu/transport-themes/infrastructure-and-investment/trans-european-transport-network-ten-t_en (30. September 2023)
17. European Commission. (2024). Climate Action. Available at https://climate.ec.europa.eu/eu-action/transport/reducing-emissions-shipping-sector/faq-maritime-transport-eu-emissions-trading-system-ets_en (12 January 2024).
18. Hintermann, B., & Ludwig, M. (2023). Home country bias in international emissions trading: Evidence from the EU ETS. *Resource and Energy Economics, 71*, 101336. https://doi.org/10.1016/j.reseneeco.2022.101336
19. Janež, P., Drobne, S., & Bogataj, M. (2018). Forecasting Dynamics of Daily Commuting to Work to Other Municipality in the Case of Changing Taxation Policies. *Lecture Notes in Management and Industrial Engineering*, 105-112. https://doi.org/10.1007/978-3-319-58409-6_12
20. Janež, P., Bogataj, M., & Drobne, S. (2016). Impact of the real estate taxation and municipal revenue on dynamics of internal migration: Case study for City Municipal of Ljubljana. *Geodetski vestnik, 60*(04), 644-684. https://doi.org/10.15292/geodetski-vestnik.2016.04.644-684
21. Mediterranean Rail Freight Corridor. (2024) https://www.medrfc.eu/
22. Nordås, H.K., & Piermartini, R. (2004). Infrastructure and Trade. Staff Working Paper ERSD-2004-04. WTO - Economic Research and Analysis Division: Geneva.
23. Oesingmann, K. (2022). The effect of the European Emissions Trading System (EU ETS) on aviation demand: An empirical comparison with the impact of ticket taxes. *Energy Policy, 160*, 112657. https://doi.org/10.1016/j.enpol.2021.112657
24. Pothen, F., & Hübler, M. (2018). The interaction of climate and trade policy. *European Economic Review, 107*, 1-26. https://doi.org/10.1016/j.euroecorev.2018.04.004
25. Randrianarisoa, L. M., & Gillen, D. (2022). Reducing emissions in international transport: A supply chain perspective. *Transportation Research Part D: Transport and Environment, 102*, 103074. https://doi.org/10.1016/j.trd.2021.103074
26. Rietveld, P., Zwart, B., van Wee, B., & van den Hoorn, T. (1999). On the relationship between travel time and travel distance of commuters. *The Annals of Regional Science, 33*(3), 269-287. https://doi.org/10.1007/s001680050105
27. Slack, B., Comtois, C., Wiegmans, B., & Witte, P. (2018). Ships time in port. *International Journal of Shipping and Transport Logistics, 10*(1), 45. https://doi.org/10.1504/ijstl.2018.088322
28. Sun, X., Tian, Z., Malekian, R., & Li, Z. (2018). Estimation of Vessel Emissions Inventory in Qingdao Port Based on Big data Analysis. *Symmetry, 10*(10), 452. https://doi.org/10.3390/sym10100452

29. Transport & Environment. (2021). Easy Ride: why the EU truck CO2 targets are unfit for the 2020s. Available at https://www.transportenvironment.org › 2021/10 PDF. (10 December 2023).
30. UNCTAD. (2018). Review of Maritime Transport 2018. United Nations publication. Sales No. E.18.II.D.5. UNCTAD: New York and Geneva.
31. Wei, H., & Lee, P. T.-W. (2021). Designing a coordinated horizontal alliance system for China's inland ports with China railway express platforms along the Silk Road Economic Belt. *Transportation Research Part E: Logistics and Transportation Review, 147*, 102238. https://doi.org/10.1016/j.tre.2021.102238
32. World Shipping Council. (2021). Top 50 ports. Available at www.worldshipping.org/top-50-ports. (20 May 2022).
33. WTO. (2021). Indices of the world merchandise trade volume from 2000 to 2021, including projections. Available at https://www.wto.org/english/news_e/pres20_e/pr862_e.htm#:~:text=World%20merchandise%20trade%20volume%20is%20forecast%20to%20fall,in%202020%20before%20rising%20by%204.9%25%20in%202021 (22 May 2022).

About the authors

David Bogataj is a Full Professor of Economics - Management at Alma Mater Europaea, Head of the Systems Engineering and Economics Department and Head of the doctoral study on Social Gerontology at Alma Mater Europaea - ECM/ Slovenija. He teaches management, infrastructure courses and actuarial-mathematical models in demography. He is also researching the global supply chain modelling and simulation in Industrial Engineering at the Technical University of Cartagena, Numerical Simulation Group in Industrial Engineering. He is actively involved in several international and national research projects. He is a Member of the Advisory Board of the scientific journals Clinical Interventions in Aging and International Journal of Production Management and Engineering. The author can be contacted at **dbogataj@actuary.si.**

Francisco Campuzano-Bolarín is a Full Professor at the Technical University of Cartagena, the Department of Business Economics secretary, teaching Decision Models, Simulation of Logistics Systems, and Business Organization and Management. He is also a member of the Numerical Simulation Group in Industrial Engineering. He is actively involved in international and national research projects. He is a Member of the Advisory Board of the DYNA scientific journal and the International Journal of Production Management and Engineering. The author can be contacted at **francisco.campuzano@upct.es.**

José Andrés Moreno Nicolás is a Professor Titular at the Technical University Cartagena, teaching Noise and Vibration in Machines, Transportation Engineering, Design of Machine Elements II, Noise and Vibration Control, and Ergonomic Design and Product Reliability. He is actively involved in international and national research projects. He is a leading researcher for Grup I+D, Department of Mechanical Engineering, and Head of the Simulation Group at Industrial Engineering. The author can be contacted at **josea.moreno@upct.es.**

Selection Procedure of the Approximation Methods for Deriving Priorities: A Case of Inconsistent Pairwise Comparisons

Vesna Čančer
University of Maribor, Faculty of Economics and Business, Slovenia

Abstract

Background: When pairwise comparisons are used to express preferences for alternatives or judgments on criteria's importance, several methods can be used to derive priorities in multi-criteria decision-making. In the case of inconsistency, different methods give different results. **Objectives:** The main goal of this paper is to present the procedure of measuring the accuracy of the selected approximation methods based on pairwise comparisons compared to the priorities obtained by the eigenvalue method. It also aims to illustrate the procedure on the numerical example characterised by acceptable inconsistency. **Methods/Approach:** The presented procedure is based on a prescriptive approach, the fixed ratio scale, reciprocal pairwise comparison matrices, and consistency ratio. Mean absolute deviation and mean absolute percentage deviation are used to measure accuracy. **Results:** The first result is the theoretical statement of the priorities' accuracy measurement procedure. The results of the numerical example characterised by the preferences of strength slight to strong plus show that, on average, the most accurate approximation method is the geometric mean method. **Conclusions:** The research contributes to the literature on prescriptive approaches to decision-making. The results can show potential users which approximation method to use and lecturers which of them to include in the curriculum portfolio.

Keywords: accuracy; analytic hierarchy process; approximation method; pairwise comparisons; priority; simulation

JEL classification: C44, C60, C63
Paper type: Research article

Received: Jan 25, 2024
Accepted: May 5, 2024

Acknowledgements: This research is a part of the research program P5-0023: Entrepreneurship for Innovative Society, supported by the Slovenian Research Agency.

Citation: Čančer, V. (2024). Selection Procedure of the Approximation Methods for Deriving Priorities: A Case of Inconsistent Pairwise Comparisons. *Business Systems Research, 15*(2), 21-30.
DOI: doi.org/10.2478/bsrj-2024-0015

Introduction

In multi-criteria decision-making, pairwise comparisons (Kuske et al., 2019) are recognised as a useful indirect way of expressing preferences for alternatives and judgments of criteria's importance whose advantages are reflected in a growing number of applications in solving complex problems (e.g., Chakrabortty, Abdel-Basset, & Ali, 2023; Čančer et al., 2023; Koczkodaj et al, 2016; Promentilla et al., 2018). According to Koczkodaj et al. (2016), pairwise comparison method is one of a few valid methods for processing subjective data. Based on pairwise comparisons of preferences to various alternatives and the importance of various criteria performed by the decision-maker, the pairwise comparison matrices are built. Grzybowski (2012) identifies the right eigenvalue method, which is later in this paper called the eigenvalue method, as one of the most popular, commonly used, and recommended for deriving priorities. According to Choo and Wedley (2004), the eigenvalue method also satisfies the condition of correctness in error-free cases. However, quality and purpose-developed computer programs that support decision-making procedures, including this method, are often inaccessible to users due to several factors such as price, incompatibility of computer programs with the operating system, etc. Users can, and in these cases, also use approximation methods for creating appropriate spreadsheets with available programs, such as Excel. The question of how to choose the method to calculate the priorities (Ishizaka, 2019) arises.

Kazibudzki (2019) pointed out that when judgments or preferences are perfectly consistent, i.e., cardinally transitive, all approximation procedures coincide, and the quality of the prioritization process is exemplary. According to Koczkodaj et al (2016), the lack of consistency in the pairwise comparison matrices is the main challenge in terms of realistic inputs. As the human judgments and preferences are rarely perfectly consistent (Kazibudzki, 2019), the results are often subject to considerable inaccuracy and the quality of the priorities derived from pairwise comparison matrices with different approximation methods may vary. By adapting Chen's (2020) definition of inaccuracy to our problem it can be concluded that inaccuracy means that priorities do not reflect the real relative importance levels of criteria or preference levels of alternatives. This is a critical issue since approximation methods are often used in practice to approximate priorities. If the matrix of expressed judgments or preferences is inconsistent, different approximating methods give different priorities, i.e., weights and local alternatives' values. When selecting approximation methods, it is, therefore, appropriate to assess the accuracy of the obtained priorities.

The purpose of this paper is, therefore, to develop the procedure for measuring the accuracy of the approximation methods for priorities derivation based on inconsistent pairwise comparisons (see, e.g., Choo & Wedley, 2004; Saaty, 2012) compared to the priorities obtained by the eigenvalue method. This paper aims also to compare the accuracy of the approximation methods on the numerical example of the inconsistent pairwise comparisons matrix with the elements expressing slight to strong plus preferences to alternatives.

For this purpose, mean absolute deviation (*MAD*) and mean absolute percentage deviation (*MAPD*) (Bastič, 2003; Render et al., 2021) were adjusted and used as the measures of accuracy. For example, Grzybowski (2012) already used *MAD* when comparing simulation results of several prioritization techniques, but for the analysis of the rounding impact and the errors of human nature. The eigenvalue method, on the other hand, is used in the analytic hierarchy process (AHP) method. Many scientists, practitioners, and students from various scientific and professional fields agree that AHP is a simple and versatile multicriteria method (e.g., Čančer et al., 2023; Ishizaka, 2019; Promentilla et al., 2018) that helps individuals and groups solve important comprehensive decision problems. It provides a systematic decision-making procedure (Promentilla et al., 2018). Koczkodaj et al (2016) pointed out that AHP is not the only representation of pairwise comparisons. However, the AHP method had a significant impact on the pairwise comparisons research (Koczkodaj et al, 2016; Ágoston & Csató, 2022), which also applies to the research presented in this paper.

This paper aims to answer the following research questions:

- RQ1: How can we combine the simulation of inconsistent pairwise comparison matrices and customized accuracy measures to the procedure of selecting the most accurate approximation method for deriving priorities? and
- RQ2: Which approximation method considered in case of inconsistencies and slight to strong plus judgments or preferences gives the most accurate priorities?

The rest of the paper is structured as follows. The methodological section presents the basics of the eigenvalue method of the AHP and the selected approximation methods for priorities derivation. It also presents the selected accuracy measures and a description of the procedure for assessing the accuracy of the approximation methods for priorities derivation. The next section illustrates the developed procedure with a numerical example. The paper concludes with the main findings, limitations, and further research possibilities.

Methodology

Under the term priority, we understand the weights of criteria and local values of alternatives. When decision-makers cannot determine criteria weights and local alternatives' values directly they can use indirect methods based on ordinal, interval, and ratio scales (Belton & Stewart, 2002; Čančer, 2012). Judgments on the importance of the criteria and preferences for alternatives concerning a single criterion can be expressed by pairwise comparisons based on a ratio scale (Saaty, 2008, 2012). The linguistic equivalents to numerical values of the fundamental AHP scale (Saaty, 2008) were used in this research: 1 means that none of the two criteria compared is more important or none of the two alternatives compared is favoured; 2 means that the criterion is slightly more important than the compared one or the alternative is slightly more preferred than the compared one; the linguistic explanation of judgment or preference strength 3 is moderate, of 4 moderate plus, of 5 strong, of 6 strong plus, of 7 very strong, of 8 very, very strong, and of 9 extreme. Reciprocal values should be used when the criterion is less important or the alternative is less preferred than the compared one.

Eigenvalue method for priorities derivation

Let us summarize the basics of the eigenvalue method for the priorities' derivation (Saaty & Sodenkamp, 2010). Judgments on criteria's importance, expressed by pairwise comparisons, are the ratios of the criteria weights that indicate that criterion i is a_{ij} times more important than criterion j. Similarly, preferences for alternatives concerning each criterion, expressed by pairwise comparisons, are the ratios of the local values that indicate that the alternative A_i is a_{ij} times more preferred than alternative A_j:

$$a_{ij} = \frac{p_i}{p_j}, \qquad (1)$$

where p_i is the weight of the ith criterion or the local value of the ith alternative, and p_j is the weight of the jth criterion or the local value of the jth alternative, i = 1, 2, …, k,

$j = 1, 2, \ldots, k$. Based on pairwise comparisons, we can write a square matrix A. Let A be the matrix of expressed judgments on the criteria's importance as well as the matrix of expressed preferences to alternatives with the following characteristics: $a_{ij} > 0$, $a_{ij} = 1/a_{ji}$, $a_{ii} = 1$, and $a_{im} \times a_{mj} = a_{ij}$. The latter characteristic, the so-called transitivity, applies only in the case of complete consistency. In this case, $Ap = kp$, or $(A - kE)p = 0$, which is a homogenous system of k linear equations with k unknown variables. It has infinitely many solutions because the rows in matrix A are proportional. In practice, the consistency is usually incomplete, so we get the system:

$$Ap = \lambda p, \qquad (2)$$

where λ is the eigenvalue of matrix A and p is the eigenvector of matrix A. If and only if $k = \lambda$, the consistency is complete. λ is determined so that (2) has infinitely many solutions. We obtain a polynomial of the k^{th} level. At λ_{max}, we calculate a particular solution so that $\sum_{m=1}^{k} p_m = 1$. The smaller the difference $|\lambda_{max} - k|$, the more consistent a decision-maker. Consistency index $CI = \frac{\lambda_{max} - k}{k-1}$ can be used as a measure of inconsistency. However, in this paper, the consistency ratio

$$CR = \frac{CI}{R}, \qquad (3)$$

where R is the random index of inconsistency, obtained experimentally considering k (Ishizaka, 2019; Saaty & Sodenkamp, 2010), is used as a measure of inconsistency. A decision-maker is reasonably consistent if $CR \leq 0.1$.

Approximation methods for priorities derivation

Priorities can be derived using several approximation methods (Choo & Wedley, 2004). Ease of use was a fundamental criterion for including the assessment of the accuracy of the following approximation methods in the paper:

I. Divide the sum of the values in each row with the sum of all values in matrix A.
II. Calculate the reciprocal value of the sum of the values in each column in matrix A.
III. Calculate priorities as the average of priorities calculated by I and II.
IV. First, add the values in each column in matrix A. Then divide each entry in each column by the total of that column to obtain the normalized matrix which permits meaningful comparison among elements. Finally, calculate the average over the rows by adding the values in each row of the normalized matrix and dividing the rows by the number of entries in each. This is the so-called approximative eigenvector method based on normalization (Saaty, 2012).
V. First, calculate the geometric mean of a row in the pairwise comparison matrix A. That geometric mean is the priority value of the factor indicated by the row. Normalize the priorities by dividing each priority value by the sum of all priorities that is obtained from the geometric mean. This is the so-called geometric mean method (Choo & Wedley, 2004; SpiceLogic Inc, 2022).

Measures of accuracy

We adjusted the selected measures of forecast accuracy (Bastič, 2003; Render et al, 2021) to the measures of the accuracy of priorities derived by approximation methods based on pairwise comparisons. To see how accurate the priorities were, the priorities obtained with approximation methods were compared to the priorities obtained with the eigenvalue method, which in this paper is assumed as an exact method. The error is defined as the difference between the priority obtained with the exact method and the priority obtained with an approximation method. The adjusted measures are as follows.

Mean absolute deviation (*MAD*) is computed by taking the sum of the absolute values of the individual errors and dividing it by the number of errors:

$$MAD = \frac{1}{r}\sum_{l=1}^{r}|p_l^e - p_l^a|, \qquad (4)$$

where p_l^e is the exact l^{th} priority and p_l^a is the approximate l^{th} priority, and *r* is the number of simulations regarding *CR* (3).

Mean absolute percentage deviation (*MAPD*) is calculated by taking the sum of the absolute values of the individual errors, dividing it by the sum of exact priorities, and multiplying by 100:

$$MAPD = \frac{\sum_{l=1}^{r}|p_l^e - p_l^a|}{\sum_{l=1}^{r} p_l^e} \times 100. \qquad (5)$$

According to exact priorities, it reflects the mean percentage of absolute individual errors.

The procedure of measuring the accuracy of the approximation methods for deriving priorities

The procedure of measuring the accuracy of the approximation methods for priorities derivation is based on the simulation:

- Initiate from the perfectly consistent pairwise comparisons matrix with expressed judgments on the criteria's importance or preferences for alternatives. Then change a particular element in matrix *A* so that the inconsistency increases. In this procedure, *CR* is used as a discrete variable. As we want to measure the priorities' accuracy when a decision-maker is acceptably consistent ($CR \leq 0.1$), the *CR*'s values from 0.01 to 0.1 are considered.
- Obtain the priorities' values with the exact method and with the approximation methods considered.
- Calculate the accuracy measures *MAD* (4) and *MAPD* (5) for each priority and approximation method. For each approximation method, the average of the *MAD* values is calculated, as well as the average of the *MAPD* values. The approximation method where the mean *MAD* and *MAPD* values are the lowest should be used to prepare the multi-criteria decision-making basis.

The computer program Expert Choice can be used to obtain the priorities' values with the exact method, and the computer program Excel can be used to obtain the priorities' values with approximation methods.

Results

Let us illustrate the procedure of measuring the accuracy of the approximation methods for priorities derivation with a numerical example based on an extensive real-life problem of selecting the most appropriate video-conferencing system. The following video-conferencing systems for medium room were included as alternatives: MeetingBar A30 (Yealink, 2023) – Alternative 1, Panacast 50 (Jabra[GN], 2023) –

Alternative 2, and Poly Studio X50 (Plantronics Inc, 2024) – Alternative 3. Based on the IT expert help that considered the characteristics of alternatives, preferences for alternatives concerning 'camera' are as follows: Alternative 1 is 3 times – moderately more preferred than Alternative 2 and 6 times – strongly plus more preferred than Alternative 3, and Alternative 2 is twice – slightly more preferred than Alternative 3.

The initial matrix A is given as follows:

$$\begin{bmatrix} 1 & 3 & 6 \\ {}^{1}/_{3} & 1 & 2 \\ {}^{1}/_{6} & {}^{1}/_{2} & 1 \end{bmatrix}. \qquad (6)$$

$CR = 0$ shows that (6) is a perfectly consistent matrix. Then we changed the element a_{23} so that the inconsistency increased: 2.5 (CR = 0.01), 3 (CR = 0.02), 3.5 (CR = 0.03), 3.75 (CR = 0.04), 4 (CR = 0.05), 4.25 (CR = 0.06), 4.51 (CR = 0.07), 4.74 (CR = 0.08), 5.02 (CR = 0.09), 5.26 (CR = 0.1). In (6), the reciprocal values of a_{23} must be calculated for a_{32}, as well.

Table 1 presents the values of priorities p_1, p_2, and p_3, obtained with the exact and five approximation methods described in the previous section and rounded to three decimal places.

Table 1
Values of Priorities Obtained with Exact and Approximation Methods

Method	Consistency Ratio										
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
	p_1										
Exact	0.667	0.661	0.655	0.649	0.647	0.644	0.642	0.639	0.637	0.635	0.633
I	0.667	0.649	0.632	0.614	0.605	0.597	0.589	0.580	0.573	0.564	0.557
II	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667
III	0.667	0.658	0.649	0.640	0.636	0.632	0.628	0.623	0.620	0.616	0.612
IV	0.667	0.660	0.653	0.646	0.643	0.639	0.636	0.633	0.630	0.627	0.624
V	0.667	0.661	0.655	0.649	0.647	0.644	0.642	0.639	0.637	0.635	0.633
	p_2										
Exact	0.222	0.237	0.250	0.261	0.266	0.271	0.275	0.279	0.283	0.288	0.291
I	0.222	0.249	0.274	0.297	0.308	0.318	0.329	0.339	0.348	0.359	0.367
II	0.222	0.227	0.231	0.233	0.234	0.235	0.236	0.237	0.237	0.238	0.239
III	0.222	0.238	0.252	0.265	0.271	0.277	0.282	0.288	0.293	0.298	0.303
IV	0.222	0.238	0.251	0.263	0.268	0.274	0.279	0.284	0.288	0.293	0.297
V	0.222	0.237	0.250	0.261	0.266	0.271	0.275	0.279	0.283	0.288	0.291
	p_3										
Exact	0.111	0.102	0.095	0.090	0.087	0.085	0.083	0.081	0.080	0.078	0.076
I	0.111	0.102	0.095	0.089	0.087	0.085	0.083	0.081	0.079	0.077	0.076
II	0.111	0.105	0.100	0.095	0.093	0.091	0.089	0.087	0.085	0.083	0.082
III	0.111	0.103	0.097	0.092	0.090	0.088	0.086	0.084	0.082	0.080	0.079
IV	0.111	0.102	0.096	0.091	0.089	0.087	0.085	0.084	0.082	0.081	0.079
V	0.111	0.102	0.095	0.090	0.087	0.085	0.083	0.081	0.080	0.078	0.076

Note: I, II, III, IV, V – approximation methods
Source: Author's calculations

It can be concluded that the higher the CR, the more the priorities obtained differ from those at CR = 0 (Table 1). This applies to all the methods used, except for II, when used for the calculation of p_1, as the first column in (6) does not change in simulations.

The values of *MAD* and *MAPD* were calculated by (4) and (5) so that the approximation values of priorities were considered at *CR* = 0.01 to *CR* = 0.1 (Table 1), *r* = 10. At *CR* = 0, the exact value is equal to the approximation value, regardless of the approximation method used. The values of accuracy measures *MAD* and *MAPD* of priorities obtained with approximation methods from I to V are given in Table 2.

Table 2
Values of Accuracy Measures for Priorities Obtained with Approximation Methods

Approximation method	Mean Absolute Deviation				Mean Absolute Percentage Deviation			
	p_1	p_2	p_3	Mean	p_1	p_2	p_3	Mean
I	0.0482	0.0487	0.0003	0.0324	7.482	18.030	0.350	8.62
II	0.0228	0.0354	0.0053	0.0212	3.539	13.106	6.184	7.61
III	0.0128	0.0066	0.0024	0.0073	1.987	2.444	2.801	2.41
IV	0.0051	0.0034	0.0019	0.0035	0.792	1.259	2.217	1.42
V	0	0	0.0010	0.0003	0	0	1.167	0.39

Source: Author's calculations

The results in Table 2 show that on average, approximation method V, i.e., the geometric mean method, gives the most accurate priorities' values: mean absolute deviation is 0.0003 (*MAD* = 0.0003), and the sum of the absolute values of individual errors present 0.39 % of the sum of the exact priorities (*MAPD* = 0.39). Moreover, this method gives perfectly accurate values of p_1 and p_2 (*MAD* = *MAPD* = 0), and among the considered approximation methods, the second most accurate value of p_3 (*MAD* = 0.001, *MAPD* = 1.167). The second most accurate approximation method is, on average, the approximative eigenvector method based on normalization (IV, *MAD* = 0.0035, *MAPD* = 1.42), followed by the approximation method based on the average of priorities obtained with approximation methods I and II (III, *MAD* = 0.0073, *MAPD* = 2.41), and the approximation method based on the reciprocal value of the sum of the values in each column of *A* (II, *MAD* = 0.212, *MAPD* = 7.61). The least accurate approximation method is I which is based on the ratio of the sum of the values in each row and the sum of all values in the matrix *A* (*MAD* = 0.0324, *MAPD* = 8.62). The same order of accuracy of the approximation methods also applies to p_1 and p_2. For calculating p_3 values, however, the approximation method I is the most accurate, followed by the approximation methods V, IV, III, and II.

Discussion and Conclusion

The research work presented in this paper resulted in the theoretical statement of the priorities' accuracy measurement scheme, based on pairwise comparisons. The methodological part answered the first research question. Beginning with a perfectly consistent pairwise comparison matrix with expressed judgments on the criteria's importance or preferences for alternatives, the procedure for selecting the most accurate approximation method for deriving priorities includes several sequential steps. Initially, a simulation is conducted to gradually increase the inconsistency until reaching the *CR* at which the pairwise comparison matrix is still acceptably consistent. Following this, priorities are derived using the eigenvalue method and selected approximation methods. Then, accuracy measures are calculated, and the approximation method with the minimal mean accuracy values is identified. The accuracy measures *MAD* and *MAPD* were adapted to this problem.

The procedure has been applied to a numerical example to illustrate its applicability. The results of the considered numerical example can help us answer the second research question. In case of inconsistencies and slight to strong plus judgments or preferences, on average, the geometric mean method gives the most accurate priorities among the selected approximation methods for deriving priorities. This is in line with the summarization (Ishizaka, 2019) that simulations did not identify significant differences between the geometric mean and eigenvalue method.

The procedure for measuring the accuracy of the approximation methods for deriving priorities is useful for advising on the determination of a portfolio of approximation methods to those users who express their judgments and preferences by pairwise comparisons but do not have access to computer programs in which the calculation of priorities is based on the eigenvalue method. The results can show potential users which approximation method to use, and lecturers which approximation methods to include in the curriculum portfolio.

In this research, the accuracy measures were limited to *MAD* and *MAPD*. A well-known accuracy measure is an average error, known as bias, which tells whether the priorities' values obtained with approximation methods tend to be too high or too low and by how much; it may be negative or positive (Render et al, 2021). Because the negative errors can cancel out the positive ones, it is not a good measure of the actual size of the errors (Render et al, 2021); for this reason, it has been omitted. This research is limited to AHP pairwise comparisons using a positive reciprocal matrix. Focused on the research problem considered in this paper we did not use accuracy measures in other research fields (e.g., Vrigazova, 2020, 2021) or efficiency measures (Abele-Nagy et al., 2018; Chen, 2020). The numerical example was limited to the matrix of order 3 x 3, i.e., the lowest order in which inconsistencies can arise, and to relatively large ratios between matrix *A* elements. This research does not deal with the fuzzy analytic hierarchy process where approximation methods have been extensively applied to determining the weights of criteria in MCDM problems (Chen, 2020).

Further research can be oriented toward matrices of higher order, with different, smaller, and larger ratios between matrix A elements in several simulations. Further research possibilities also arise in adapting the presented procedure to examine the accuracy of fuzzy analytic hierarchy process methods according to the consistency ratio.

References

1. Abele-Nagy, K., Bozóki, S., & Rebák, Ö. (2018). Efficiency analysis of double perturbed pairwise comparison matrices. *Journal of the Operational Research Society, 69*(5), 707–713. https://doi.org/10.1080/01605682.2017.1409408
2. Ágoston, K. C., & Csató, L. (2022). Inconsistency thresholds for incomplete pairwise comparison matrices. *Omega, 108*, 102576. https://doi.org/10.1016/j.omega.2021.102576
3. Bastič, M. (2003). *Izvedbeni management: Optimizacijski modeli [Operations Management: Optimization Models,* in Slovenian] (2nd ed.). Maribor, Slovenia: University of Maribor, Faculty of Economics and Business.
4. Belton, V., & Stewart, T. J. (2002). Multiple Criteria Decision Analysis. https://doi.org/10.1007/978-1-4615-1495-4
5. Chakrabortty, R. K., Abdel-Basset, M., & Ali, A. M. (2023). A multi-criteria decision analysis model for selecting an optimum customer service chatbot under uncertainty. *Decision Analytics Journal, 6*, 100168. https://doi.org/10.1016/j.dajour.2023.100168
6. Chen, T. (2020). Enhancing the efficiency and accuracy of existing FAHP decision-making methods. *EURO Journal on Decision Processes, 8*(3-4), 177-204. https://doi.org/10.1007/s40070-020-00115-8

7. Choo, E. U., & Wedley, W. C. (2004). A common framework for deriving preference values from pairwise comparison matrices. *Computers & Operations Research, 31*(6), 893-908. https://doi.org/10.1016/s0305-0548(03)00042-x
8. Čančer, V. (2012). Criteria weighting by using the 5Ws & H technique. *BSRJ, 3*(2), 41-48. https://doi.org/10.2478/v10305-012-0011-3
9. Čančer, V., Tominc, P., & Rožman, M. (2023). Multi-Criteria Measurement of AI Support to Project Management. *IEEE Access, 11*, 142816-142828. https://doi.org/10.1109/access.2023.3342276
10. Grzybowski, A. Z. (2012). Note on a new optimization based approach for estimating priority weights and related consistency index. *Expert Systems with Applications, 39*(14), 11699-11708. https://doi.org/10.1016/j.eswa.2012.04.051
11. Ishizaka, A. (2019). Analytic Hierarchy Process and Its Extensions. *Multiple Criteria Decision Making*, 81-93. https://doi.org/10.1007/978-3-030-11482-4_2
12. Jabra[GN]. (2023, November 8). *Jabra PanaCast 50 Tech Sheet 08112023*. Retrieved from https://www.jabra.com/supportpages/jabra-panacast-50#/#8200-232
13. Kazibudzki, P. T. (2019). The Quality of Ranking during Simulated Pairwise Judgments for Examined Approximation Procedures. *Modelling and Simulation in Engineering, 2019*, 1-13. https://doi.org/10.1155/2019/1683143
14. Koczkodaj, W. W., Mikhailov, L., Redlarski, G., Soltys, M., Szybowski, J., Tamazian, G., Wajch, E., & Yuen, K. K. F. (2016). Important Facts and Observations about Pairwise Comparisons (the special issue edition). *Fundamenta Informaticae, 144*(3-4), 291-307. https://doi.org/10.3233/fi-2016-1336
15. Kuske, C., Soltys, M., & Kułakowski, K. (2019). Approximating consistency in pairwise comparisons. *Procedia Computer Science, 159*, 814–823. https://doi.org/10.1016/j.procs.2019.09.240
16. Plantronics. (2024). *Poly Studio X50*. Retrieved from https://www.poly.com/us/en/products/video-conferencing/studio/studio-x50
17. Promentilla, M. A. B., Aviso, K. B., Lucas, R. I. G., Razon, L. F., & Tan, R. R. (2018). Teaching Analytic Hierarchy Process (AHP) in undergraduate chemical engineering courses. *Education for Chemical Engineers, 23*, 34–41. https://doi.org/10.1016/j.ece.2018.05.002
18. Render, B., Stair, R. M., Hanna, M. E., & Hale, T. S. (2021). *Quantitative Analysis for Management* (13th ed). Harlow, England: Pearson.
19. Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences, 1*(1), 83. https://doi.org/10.1504/ijssci.2008.017590
20. Saaty, T. L. (2012). *Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World* (3rd ed.). Pittsburgh, PA: RWS Publications.
21. Saaty, T. L., & Sodenkamp, M. (2010). The Analytic Hierarchy and Analytic Network Measurement Processes: The Measurement of Intangibles: Decision Making under Benefits, Opportunities, Costs and Risks. *Applied Optimization*, 91-166. https://doi.org/10.1007/978-3-540-92828-7_4
22. SpiceLogic. (2022). *Analytic Hierarchy Process Software: AHP Calculation Methods*. Retrieved from https://www.spicelogic.com/docs/ahpsoftware/intro/ahp-calculation-methods-396
23. Vrigazova, B. (2021). The Proportion for Splitting Data into Training and Test Set for the Bootstrap in Classification Problems. *Business Systems Research Journal, 12*(1), 228-242. https://doi.org/10.2478/bsrj-2021-0015
24. Vrigazova, B. (2020). Tenfold Bootstrap as Resampling Method in Classification Problems. *ENTRENOVA - ENTerprise REsearch InNOVAtion*, 6(1), 74–83. Retrieved from https://hrcak.srce.hr/ojs/index.php/entrenova/article/view/13435
25. Yealink. (2023). *MeetingBar A30 Datasheet*. Retrieved from https://support.yealink.com/en/portal/knowledge/show?id=6459d53b0a11b818a75424c6

About the author

Vesna Čančer, Ph.D. in economic and business sciences, is a full professor of quantitative methods in business science at the University of Maribor's Faculty of Economics and Business. Her research focuses primarily on decision analysis with an emphasis on multi-criteria decision-making, research methods, and management science. She was the leader of several research projects for the application of multi-criteria methods in business practice. She also transfers the results of her research work into pedagogical work. She is a member of the forum Women in Society Doing Operational Research and Management Science at the Association of European Operational Research Societies and the Section for Operational Research of the Slovenian Society Informatika. The author can be contacted at **vesna.cancer@um.si.**

Design of Social Infrastructure and Services Taking into Account Internal Migration by Age Cohort

Samo Drobne
University of Ljubljana, Faculty of Civil and Geodetic Engineering, Ljubljana, Slovenia
Marija Bogataj
Institute INRISK, Trebnje, Slovenia

Abstract

Background: European cities and regions are facing depopulation and an ageing population, leading to a shift in the demand and supply of goods and giving rise to the silver economy. This demographic change has an impact on urban and regional planning, which is influenced by both internal and external migration. **Objectives:** Based on the hypothesis that the attractiveness of locations also depends on the age of the inhabitants, the paper investigates the gravitational effects on the intensity of migration flows by age cohorts. **Methods/Approach:** This study examines how factors that influence the retention or attraction of people towards specific areas affect migration between age groups at different hierarchical spatial levels, using the gravity model implemented at the Slovenian spatial levels NUTS 2 and NUTS 3. **Results:** Distance is least important for the 65-74 age group, while wages influence only the youngest cohorts. The capacity of care homes has a significant influence on the attractiveness of older cohorts to move between NUTS 2 regions. There is a high correlation between the factors at the municipal and NUTS 3 levels for the population aged 75+. The factors at NUTS 2 and NUTS 3 levels show a strong correlation for those under 65. **Conclusions:** These results can form a basis for the development of the silver economy as they show the need for adapted infrastructures and services for older adults. As the age structure is changing, authorities should adapt infrastructures and services to the different levels of central places/regions. The growing number of older people makes research into optimal solutions for long-term care a crucial factor for the silver economy.

Keywords: gravity model; infrastructure; attractiveness; demography; gerontology
JEL classification: A13; C10; J11; J26

Paper type: Research article

Received: Feb 19, 2024
Accepted: May 5, 2024

Citation: Drobne, S., & Bogataj, M. (2024). Design of Social Infrastructure and Services Taking into Account Internal Migration by Age Cohort. *Business Systems Research, 15*(2), 31-46.
DOI: doi.org/10.2478/bsrj-2024-0016

Introduction

The shrinking of cities and regions is a growing phenomenon in all European cities and regions and, at the same time, a popular topic in international regional and urban research, the first publications of which were discussed in detail by Oswalt and Bundes (2005). In the post-industrial society, the number of industries and employees has declined in most industrialised cities in developed countries. Some cities, such as Detroit, even went bankrupt due to the migration of industry. While the decline of the first cities was related to the loss of production factors, especially raw materials (Manaos), which was followed by suburbanisation with the availability of private vehicles and quick access to city centres (Li et al., 2024), today the fundamental problem is the ageing of the population and demographic decline in general. In some countries, however, the contradiction between the decline in urban population and the increase in building land is also becoming a pressing problem for many smaller cities (Li et al., 2022).

Simulation and optimisation models for building land acquisition are an important part of decision-making models to support land policy and a fundamental way to ensure that even smaller cities can make sustainable use of land resources (Li et al., 2024). With the development of spatial planning technology, optimisations in decision-making on the location and timing of investments and, above all, simulations, the theoretical foundations and methodological approaches for the optimal allocation of building land - housing and social infrastructure - are constantly being updated. However, these processes do not take sufficient account of the optimal spatial development in relation to the different age structures of the population, which require different living environments and, above all, different social infrastructures in the neighbourhoods where they live. Many optimisation models based on multi-criteria linear programming (Das et al., 2015), system dynamics and Markov chain processes (Fang et al., 2019; Wang et al., 2022a Wang et al., 2022b; Liang et al., 2018) solve the question of what the dynamics of property construction should be but do not take into account the importance of the attractiveness of the location in relation to the age structure of the population.

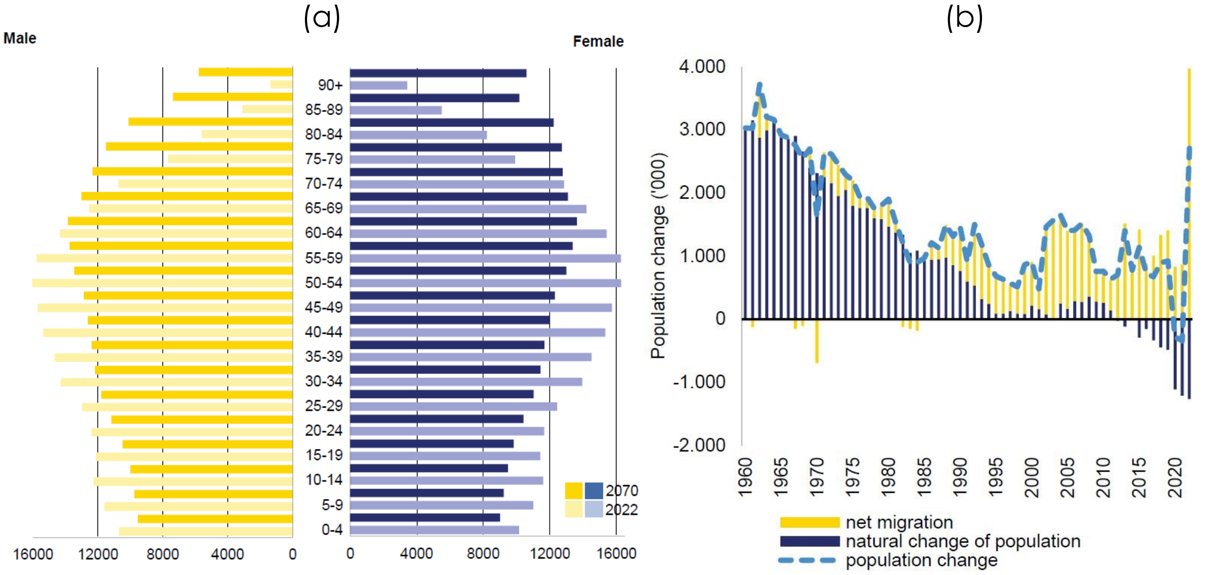
As Li et al. (2024) note, optimisation models of spatial and temporal allocation do not focus on the selection of indicator factors. Therefore, in our study, we focused on the issue of location attractiveness, i.e. factors that attract population from elsewhere to the chosen spatial unit or ensure that residents do not move away. Knowing the value of these indicators will make it much easier to successfully plan land utilisation and meet the demand for housing and social infrastructure in their vicinity.

The question arises as to whether investments in space and other factors have characteristically different effects on the intensity of growth or shrinkage of an area and thus influence migration flows with regard to the age of the inhabitants. We hypothesise that the attractiveness of locations also depends on the age of the inhabitants. We have investigated the gravitational effects on the intensity of migration flows by age cohorts and the spatial organisation of territorial units.

Population ageing and shrinking areas

European cities and regions are ageing and shrinking (European Commission - EC, 2023). The EC has summarised the EUROPOP2023 population projections published by Eurostat in March 2023. Based on these demographic fundamentals for the age-related expenditure projections for the 27 EU Member States and Norway, the EC forecasts fertility, mortality and net migration for the years 2022–2100. Figure 1a shows that the number and percentage of 65+ (men) and 75+ (women) cohorts will increase significantly by 2070, and the population decline in the EU is shown in Figure 1b.

Figure 1
European Union: (a) Population by age group and sex, 2022 and 2070 (thousands), (b) Population change, 1960–2020.



Source: European Commission (2023), based on Eurostat data.

The population decline affects population density, which varies by state and region and between urban and rural areas. Migration also influences this decline or growth. Immigration from non-European countries mitigates the sharp population decline in the EU member states. The impact on the EU as a whole is shown in Figure 1b, but there are significant differences between individual countries. In the period 1960-2020, Spain, France, Germany and Italy recorded the largest net inflows, while Poland, Bulgaria, Romania, Portugal, Lithuania and Croatia recorded the largest outflows. However, the differences between rural and urban areas within the individual countries are large. The main reasons for immigration or emigration can be analysed using the gravity approach. The theoretical principle of the gravity model is twofold:

- the degree of interaction is directly proportional to the size of the masses (number of inhabitants, economic power of the areas) and
- the degree of interaction decreases with the distance that separates them.

The first papers on the topics of "gravity model" and "migration" appeared in the Web of Science (WoS) Core Collection in 1966 and the second in 1970 (Christian & Braden, 1966; Johnston, 1970).

The dynamics of population density in countries and regions have a significant impact on the economy. As shrinking and ageing areas have fewer human resources and consume less productive output specific to younger populations, they require more products and services typical of older inhabitants. Therefore, the structure of age cohorts should influence changes in spatial planning and product supply. Eurostat confirms the thesis of Angel et al. (2010, 2011) that shrinkage dynamics are lower in larger cities than in small towns and rural areas. With the industrial transition to Industry 4.0 and Industry 5.0, declining fertility and ageing are significantly changing the demographic structure (Bogataj et al., 2019a, 2019b, 2020a, 2020b; Calzavara et al., 2020). The relative shrinkage of LAU2 spatial units in the EU and the European Economic Area (EEA) in the first decade of this millennium is presented in Drobne and Bogataj (2022). From this paper, we can see that in almost 41 per cent of EU and other EEA countries, even more than 40 per cent of urban LAUs 1 are shrinking demographically. We can also see that the proportion of depopulated LAUs 1 is higher in the east than in the western EU member states.

However, all these articles lack insight into migration by age structure and an answer to the question of how strongly individual factors, from investment in built space to the organisation of social infrastructure, influence individual age cohorts and how changes in age structure dictate the approach to spatial plans. From an organisational perspective, it is also important to answer the question of how hierarchical spatial structures influence migration. Therefore, migrations along different levels of spatial hierarchies are observed here.

We use NUTS and LAU abbreviations in this article. NUTS (Nomenclature of Territorial Units for Statistics) regions and LAU (Local Administrative Units) are classification systems used by the EU for statistical and administrative purposes. NUTS is a hierarchical system that divides countries into different levels of administrative regions for statistical reporting. NUTS regions are categorised into three levels: NUTS 1 (larger regions), NUTS 2 (smaller regions), and NUTS 3 (sub-regions or counties). LAU represents the lowest level of administrative divisions within the NUTS framework. LAU is divided into two levels: LAU 1 and LAU 2. LAU 1 corresponds to municipalities or equivalent local administrative units with a higher level of administrative authority. LAU 2 refers to smaller administrative units, such as districts or neighbourhoods, within the larger municipalities.

Literature review

In the journals indexed by the Web of Science (WoS), the keyword "shrinking city" is relatively new and was first mentioned in 2005 (Groth & Corijn, 2005). In the first decade of this millennium, there were only six articles with this keyword. The first to address the issue of migration in shrinking cities was Hillmann (2009), but he, too, only observed movements within the city. In this first decade, only Hanhörster (2009) pursued the idea of analysing (internal) migration flows in shrinking cities.

Although the gravity model in migration research is discussed in 454 articles in journals indexed by WoS, this topic relates more to shrinking or growing cities or regions. Only six articles address shrinking and growing cities or regions in the context of population ageing, all published in the last five years (Drobne et al., 2019; Arends-Kuenning et al., 2019; Bogataj et al., 2019b; Lin, 2020; Drobne & Bogataj, 2022; Zuo et al., 2023).

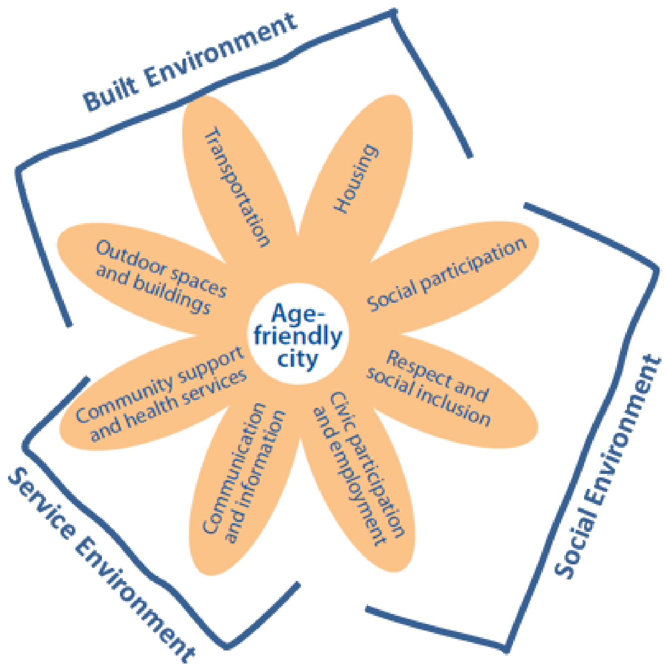
In the second decade, 159 articles were published. So far, there are 215 articles from the WoS Core Collection. The authors focus more on the renewal of industrial areas (Rienow et al., 2014). As highlighted in these articles by Wolff et al. (2018), the trend towards lower population density is more notable for small and medium-sized urban LAUs. Wolff et al. (2018) have also shown that the population has been shrinking faster in the last two decades of this millennium. The cases of high negative structural dynamics can be seen in the post-socialist states of Eastern Europe due to falling birth rates and emigration, but also in the post-industrial areas of Western Europe due to falling birth rates and changes in industrial activities. Ageing is expected to continue in the third and fourth decades in the East and West (European Commission, 2023); government and industry should therefore prepare for this phenomenon.

The articles on the challenges posed by the ageing of the European population and the related issues of migration of older citizens date mainly from the last decade when we started to become aware of the fact that as the population ages, the environment and services that the population prefers are also changing due to ageing. De Jong et al. (2020) have highlighted the role of welfare in location choice, which influences intra-European migration decisions across the European life course.

Some American authors also emphasised the importance of looking for optimal solutions for a longer time horizon. Especially when planning according to the wishes and needs of the ageing population, it should not be neglected that the wishes of the younger cohorts regarding the spatial standard in old age differ from the wishes of the older age groups.

Black and Hyer (2020) examined differences in the requirements of older cohorts, including baby boomers and older generations, for age-friendly communities where most residents are 50+ years old. The results of the investigation show significant differences between all cohorts in all WHO domains (see Fig. 2), with significant differences in preferences for housing, outdoor spaces, employment and participation in social activities (World Health Organisation, 2007). A certain difference between baby boomers and older generations was also found in Slovenia (Bogataj et al., 2020c), which shows that spatial comfort and adaptation of housing to older adults is more important for baby boomers than for older generations. This should influence the dynamics of building specialised housing units.

Figure 2
Eight domains essential for age-friendly communities are grouped into three clusters.



Source: World Health Organization (2007) and authors' elaboration.

Researchers report on the impact of parameters such as housing rents and prices, home ownership, national housing policy, demographic figures and economic conditions on mobility. In the first decade of this millennium, we find work by Engelhardt (2003) looking at equity constraints, Smith and Smith (2007) and Cunningham and Engelhardt (2008) looking at the effects of housing capital-gains taxation on migration, and many others looking at macro-level migrations, Ferreira et al. (2010) looking at the effects of housing bankruptcies on migration intensity, Conway and Houtenville (2003) focusing on older adults.

If we select the topics "older adults" and "migration", we find 638 articles in the Web of Science Core Collection (WoS) that examine the migration behaviour of older adults. Suppose we select the topics "gravity model", "older adults", and "migration"; only one article in WoS deals with the migration of the oldest citizens (Gu et al., 2022). In this article, a push-pull analysis is proposed. The results of a Poisson pseudo-maximum likelihood gravity model in this article have shown that the quantity and quality of health services have push and pull effects on older adults' decision to migrate, in addition to two factors, namely the influence of family needs and the cost effect that affects regional economic development. Karpestam (2018), not listed in WoS, also examined the gravity model in migration studies of different age cohorts in connection with new buildings. He found that new housing has recently become less accessible for certain age cohorts, even in Sweden. He investigated how the characteristics of new housing affect inter-community mobility for different age cohorts by using a gravity model that models migration as a function of origin and destination. His results show that new construction in the new millennium has affected migration within commuting regions more than between commuting regions. He also found significant negative effects on net migration into new builds from the other areas. The impact was stronger for young adults than for older adults.

The first report on migration in Slovenia, which also included older adults (focussing on their tourist activities and second residence), was presented by Bogataj and Drobne (2011). These were the results of the ESPON ATTREG project (Espon, 2013), which aimed to investigate the motivation and behaviour of migration flows and daily commuting of students, tourists, older people migrating to their second home, students and other "part-time" commuters – but especially the behaviour of human resources in gross migrations and daily commuting – between regions. At that time, one of the key elements of the European Commission's cohesion policy in 2010 was the contribution of new transport infrastructure development to regional economic development, so the paper focussed on distances in the gravity model. Bogataj and Drobne (ibid.) used the gravity approach to analyse gross migration in Slovenia (annual average 2000-2006) also as a function of GDP. The regression model of gross migration GM between the regions at the NUTS 3 level in Slovenia yielded the following equation:

$$GM_{ij} = 3.89 \cdot 10^{-5} \cdot P_i^{0.84} \cdot P_j^{0.83} \cdot d_{ij}^{-1.34} \cdot K_{GDP,i}^{0.90} \cdot K_{GDP,j}^{1.71},\ R^2 = 0.85 \quad (1)$$

where P_i and P_j are the population size in the origin and destination regions respectively, d_{ij} is the distance between the origin and destination regional centres, and $K_{GDP,i}$ and $K_{GDP,j}$ are the ratio between the GDP per capita in the NUTS 3 region and the GDP per capita at national level.

In the studies by Bogataj and Drobne, the list of factors was later expanded, but the examination of the individual age cohorts according to the hierarchical levels was not the subject of consideration. In our presentation here, we have therefore examined the effects of various factors on the migration of citizens from Slovenian municipalities to other regions at the NUTS 2 and NUTS 3 levels, considering various age cohorts, which structure will change in the next years rapidly.

Note that at the NUTS 1 level, Slovenia is the whole country, at the NUTS 2 level, Slovenia is divided into two cohesion regions, and at the NUTS 3 level, Slovenia is divided into 12 statistical regions. At a hierarchical level below NUTS 3, Slovenia is divided into 212 municipalities (LAUs).

Methodology

Based on the hypothesis that the attractiveness of locations also depends on the age of the inhabitants, the gravitational effects on the intensity of migration flows are examined according to spatial organisation and age cohorts. We analysed the influence of some factors on migrants according to different age cohorts and at different spatial levels in the normalised spatial interaction model as proposed by Drobne and Bogataj (2022) and modified here:

$$M_{ij}^{(c,s)} = k\, K(d_{ij})^{\beta} \prod_r K(r)_i^{\gamma(r)}\, K(r)_j^{\alpha(r)}, \tag{2}$$

where $M_{ij}^{(c,s)}$ is the estimated intensity of migration flows of age cohort c at a spatial level from a municipality of origin i to a municipality of destination j; age cohorts were defined as $c =$ 0–65, 66–74, 75+; analyses were conducted for three spatial levels, $s =$ municipal level, NUTS 3 level, NUTS2 level; k is the constant of proportionality; $K(d_{ij})$ is the coefficient of the shortest distance by state road network between the centre of origin municipality i and the centre of destination municipality j; $K(r)_i$ and $K(r)_j$ are coefficients of factors r in origin i or destination j, defined as the value of factor in municipality i and municipality j, respectively, divided by the average value of this factor in Slovenia (see Table 1).

At the municipal level, we have analysed the migration flows between municipalities in the same (statistical) NUTS 3 region; at the NUTS 3 level, we have analysed the inter-municipal migration flows between NUTS 3 regions but in the same NUTS 2 region; and at the NUTS 2 level, we have analysed the migration flows between municipalities of different NUTS 2 regions.

We analysed the internal migration flows between municipalities in Slovenia as an average for 2020/2021. Model (1) was linearized and solved with IBM SPSS using ordinary least squares (OLS) regression analysis for three age cohorts and three spatial levels and the significant values of α, β and γ were compared.

To better illustrate the relationships between the factors influencing different cohorts for different hierarchical relationships between origin and destination of migration flows, we summarised the power of a factor in immigration and emigration and calculated the ranks of the sums. The preferences of the different age groups were then analysed using a correlation analysis. As the ranks are integers, we used the formula (3) to calculate the Spearman correlation coefficient, which measures the degree of similarity between two ranks:

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{3}$$

Table 1
Factors in spatial interaction model (2) as choosen by authors.

Notation	Factor value	Additional description	Source
$M'^{(c,s)}_{ij}$	Number of migrants in age cohort c for year y from municipality of origin i to municipality of destination j	Average of yearly values for 2020/2021	SORS and authors' calculation
$M^{(y,s)}_{ij}$	Estimation of the number of migrants in age cohort c for year y from municipality of origin i to municipality of destination j	The estimation of the real value regarding model (1)	Authors' calculation
$K(d_{ij})$	Coefficient of the fastest time-spending distance between municipal centre of origin i and municipal centre of destination j	The ratio between the factor value for a pair of municipal centres and the average factor value for Slovenia for 2021	SIR and authors' calculation
$K(POP_{\circ})$	Coefficient of the number of inhabitants in the municipality	The ratio between the factor value for the municipality and the average factor value for Slovenia for 2021	SORS and authors' calculation
$K(UEMP_{\circ})$	Coefficient of registered unemployment rate in the municipality	The ratio between the factor value for the municipality and the average factor value for Slovenia for 2021	SORS and authors' calculation
$K(GEAR_{\circ})$	Coefficient of gross earning per capita in the municipality	The ratio between the factor value for the municipality and the average factor value for Slovenia for 2021	SORS and authors' calculation
$K(NDWE_{\circ})$	Coefficient of number of dwellings per 1000 inhabitants in the municipality	The ratio between the factor value for the municipality and the average factor value for Slovenia for 2021	SORS and authors' calculation
$K(PDM2_{\circ})$	Coefficient of average price per m² of dwelling in the municipality	The ratio between the factor value for the municipality and the average factor value for Slovenia for 2020/2021	SMARS and authors' calculation
$K(MREV_{\circ})$	Coefficient of municipal revenue per capita	The ratio between the factor value for the municipality and the average factor value for Slovenia for 2021	MFRS and authors' calculation
$K(AGEI_{\circ})$	Coefficient of ageing index in municipality	The ratio between the factor value for the municipality and the average factor value for Slovenia for 2021	SORS and authors' calculation
$K(HELD_{\circ})$	Coefficient of capacity of care homes in the municipality	The ratio between the factor value for the municipality and the average factor value for Slovenia at end of 2021	CSSI and authors' calculation
$K(SR_{\circ})$	Coefficient of number of single rooms in care homes	The ratio between the factor value for the municipality and the average factor value for Slovenia at end of 2021	CSSI and authors' calculation
$K(TAP_{\circ})$	Coefficient of number of temporary accommodation places	The ratio between the factor value for the municipality and the average factor value for Slovenia at end of 2021	CSSI and authors' calculation
$K(DCP_{\circ})$	Coefficient of number of day-care places	The ratio between the factor value for the municipality and the average factor value for Slovenia at end of 2021	CSSI and authors' calculation
$K(SDWE_{\circ})$	Coefficient of number of dwellings serviced by the care provider	The ratio between the factor value for the municipality and the average factor value for Slovenia at end of 2021	CSSI and authors' calculation
$K(COVCH_{\circ})$	Coefficient of the index of coverage of the care needs of elderly people in care homes	The ratio between the factor value for the municipality and the average factor value for Slovenia at end of 2021	CSSI and authors' calculation

Notes: SORS - Statistical Office of the Republic of Slovenia (SORS, 2023), SIR - Slovenian Infrastructure Agency (SIR, 2023), SMARS - Surveying and Mapping Authority of the Republic of Slovenia (SMARS, 2023), MFRS - Ministry of Finance of the Republic of Slovenia (MFRS, 2023), CSSI - Community of Slovenian Social Institutions (CSSI, 2023).
Source: authors' elaboration.

Results

Table 2 shows the regression coefficients of the linearized model (2). Presentation is given according to three different hierarchical spatial levels (LAU 1 municipal level, NUTS 3 region level and NUTS 2 region level) and according to three age cohorts of 0–65, 66–74 and 75+ years; the values of the regression coefficients for which the p-value is greater than 0.05 are given in brackets.

Table 2

Statistics of standardised coefficients α, β and γ for cohorts 0–65, 66–74, and 75+ year-olds at municipal, NUTS 3 and NUTS 2 hierarchical spatial levels (internal migration in Slovenia in 2020/2021; ANOVA p-value <0.001 for all).

Spatial level	LAU 1 level			NUTS 3 level			NUTS 2 level		
Age cohort	0–65	66–74	75+	0–65	66–74	75+	0–65	66–74	75+
R	0.881	0.734	0.727	0.751	0.645	0.618	0.779	0.660	0.618
R^2	0.776	0.539	0.528	0.563	0.416	0.382	0.607	0.436	0.382
Adjusted R^2	0.775	0.528	0.518	0.561	0.402	0.361	0.605	0.422	0.362
SE	0.691	0.603	0.693	0.749	0.562	0.556	0.707	0.545	0.555
No. of obs.	3,697	1,154	1,255	6,424	1,153	833	6,171	1,094	826
ANOVA stat. F	471.249	48.782	50.904	305.553	29.667	18.393.	351.654	30.545	18.300
β	-.575	-.371	-.452	-.378	-.206	-.374	-.338	-.310	-.402
$\gamma(POP)$	.473	.595	.571	.446	.464	.398	.583	.536	.586
$\alpha(POP)$	.422	.301	.194	.434	.258	.264	.563	.343	(.167)
$\gamma(UEMP)$	-.025	(.034)	(-.048)	(-.007)	(.038)	(-.032)	(.003)	(-.011)	(.074)
$\alpha(UEMP)$	(-.015)	(-.017)	(-.053)	(.006)	(.001)	(-.016)	.024	(.017)	(.040)
$\gamma(GEAR)$	.018	(.038)	.048	.043	.106	.078	.054	.122	(.060)
$\alpha(GEAR)$	.017	(.006)	(.003)	.027	(.021)	(.024)	.071	(.027)	(.033)
$\gamma(NDEW)$	.097	.170	.092	.105	.146	.124	.144	.108	(.096)
$\alpha(NDEW)$	.154	.176	.103	.152	.227	.130	.189	.214	(.030)
$\gamma(PDM2)$	-.038	(-.034)	-.064	.031	.076	(.026)	.024	(.007)	-.133
$\alpha(PDM2)$	(-.014)	(.029)	-.078	.065	.097	(.053)	.027	(-.006)	-.109
$\gamma(MREV)$	.068	.052	.072	(.002)	(-.038)	(.010)	(-.003)	(-.034)	(-.011)
$\alpha(MREV)$	.023	(.039)	(.005)	(-.001)	(-.016)	.063	(-.012)	(-.004)	(.031)
$\gamma(AGEI)$	(-.002)	(.026)	.059	(.014)	(.057)	.087	(.013)	.104	(.020)
$\alpha(AGEI)$	(-.016)	(-.009)	(-.015)	.032	(.024)	(.075)	(.020)	(-.009)	(.045)
$\gamma(HELD)$	(.015)	(-.069)	(-.064)	(-.027)	(-.069)	(-.070)	-.079	(.011)	(.131)
$\alpha(HELD)$	(.042)	.178	.588	(-.052)	(-.049)	.165	-.051	(.077)	.291
$\gamma(SR)$	(-.019)	(.036)	(.024)	(.042)	(.099)	(.035)	.063	(-.003)	(-.118)
$\alpha(SR)$	(-.032)	(.002)	-.156	.075	(.124)	(-.031)	.061	(-.036)	(-.078)
$\gamma(TAP)$	.055	(.007)	(.020)	.020	-.123	(-.060)	-.032	-.078	(-.009)
$\alpha(TAP)$	.035	(.010)	(.013)	.023	-.113	-.106	-.032	(-.023)	(-.031)
$\gamma(DCP)$	(.021)	(.001)	(.002)	(.014)	(-.044)	(.014)	(.015)	(-.052)	(-.098)
$\alpha(DCP)$	(.025)	(.032)	(-.034)	(.000)	(-.050)	(-.082)	(-.020)	(-.063)	(-.064)
$\gamma(SDWE)$	.021	.060	(.035)	.099	.058	.104	.070	(.065)	(.060)
$\alpha(SDWE)$	(.013)	(-.007)	(.022)	.090	.070	(.033)	.071	(.039)	(.035)
$\gamma(COVCH)$	.038	-.066	-.081	-.062	-.130	(-.073)	-.069	-.104	-.115
$\alpha(COVCH)$	.050	(-.003)	(-.022)	-.072	(-.043)	(-.024)	-.082	(-.016)	(-.030)

Note: The values of the regression coefficient where p-value > 0.05 are in parentheses.
Source: Authors' work

The adjusted R2 of the regression models are relatively high for inter-municipal flows (51.8% to 77.5%) and slightly lower for migration at the NUTS 3 (56.1% to 36.1%) and NUTS 2 (60.5% to 36.1%) levels. At each spatial level, they are highest for the youngest cohort and lowest for the oldest cohort - which correlates strongly with the number of observations in particular. Overall, the distance of migration, the population and the number of dwellings have a statistically significant influence on migration for all cohorts considered and at all three spatial levels. The statistical significance of the other factors varies depending on the cohort and spatial level.

To better understand the relationships between the factors affecting different cohorts in the context of migration flows, we have summarised the influence of each factor on immigration and emigration. The ranks of the sum of the powers are shown in Table 3 and the impacts of the different factors are plotted in Figure 3 for the municipal (LAU 1) and regional (NUTS 3 and NUTS 2) levels.

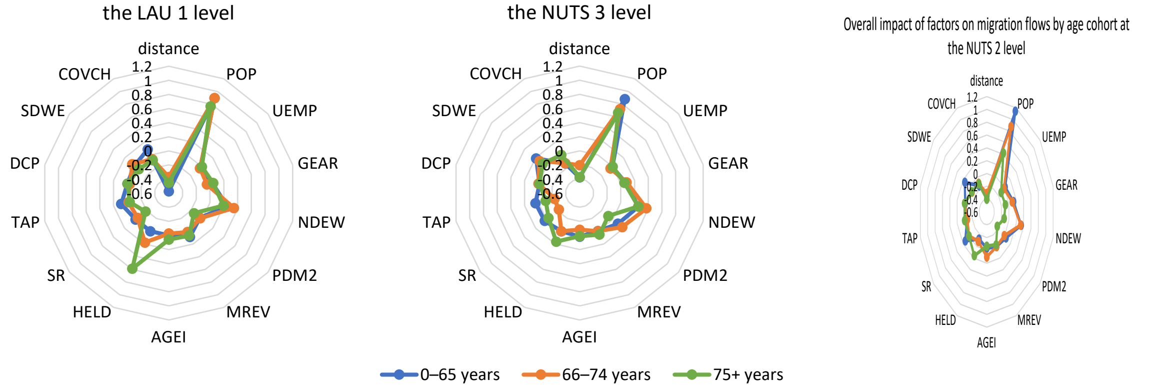
Table 3
A rank of the sum of the power of factors (immigration and emigration) for internal migration in Slovenia in 2020/2021 for different hierarchical spatial levels and different cohorts.

Spatial level	LAU 1 level			NUTS 3 level			NUTS 2 level		
Age cohort	0–65	66–74	75+	0–65	66–74	75+	0–65	66–74	75+
Distance	1	1	1	1	3	1	1	1	1
Population	14	14	14	14	14	14	14	14	14
Unemployment	3	5	8	4	6	6	8	6	5
Gross earning	9	3	9	10	10	10	11	11	7
Number of dwellings per capita	13	13	12	13	13	13	13	13	3
Dwelling price	2	7	3	11	12	3	9	7	2
Municipal revenue per capita	12	10	11	5	7	9	7	8	9
Ageing index	4	8	10	7	5	8	5	12	8
Care homes capacity	4	12	13	6	7	12	3	4	13
Single rooms in care homes	4	4	2	8	2	4	10	8	10
Temporary accommodation places	11	9	6	9	1	2	4	5	10
Day-care places	4	6	7	3	7	5	6	8	10
Dwellings serviced by the care provider	8	11	5	12	11	11	12	2	6
Coverage of the care needs of elderly people in care homes	10	2	4	2	4	6	2	3	4

Source: Authors' work

An examination of the bottom and top 10% of the ranks - overall the 1/5th strongest influencers of migration - confirms many of the findings from the literature that migration distance and population have the strongest influence on the decision to migrate. This is true for all three cohorts at all three spatial levels - with one exception for migrants between NUTS 3 regions in the middle age cohort (66-74 years), where the factor number of temporary accommodation places in the municipality plays a more important role than distance. With the exception of the oldest cohort, the number of dwellings in the municipality also has a significant influence on the decision to move (applies to all spatial levels considered).

Figure 3
The overall impact of factors on migration flows by age cohort at the level of municipalities (LAU 1), NUTS 3, and NUTS 2 regions.



Notes: POP - population, UEMP - unemployment, GEAR - gross earning, NDEW - number of dwellings per capita, PDM2 - dwelling price, MREV - municipal revenue per capita, AGEI - ageing index, HELD - care homes capacity, SR - single rooms in care homes, TAP - temporary accommodation places, DCP - day-care places, SDWE - dwellings serviced by the care provider, COVCH - coverage of the care needs of elderly people in care homes.
Source: Authors' work

Overall, the social infrastructures and services considered in the study also have a significant impact, particularly on the two oldest cohorts (66–75 and 75+). At the level of moves between municipalities within the same NUTS 3 region, the capacity of care homes, the number of single rooms in care homes and the general index of coverage of the care needs of older people in care homes are important factors; at the level of moves between NUTS 3 regions, the number of single rooms in care homes and the number of temporary accommodation places are important factors. At the level of moves between NUTS 2 regions, the number of flats serviced by the care provider is important in addition to the capacity of care homes. The general index of coverage of care needs of older people in care homes also has a significant impact on the decision to move over longer distances (inter-regional moves) for the youngest cohort considered (0–65 years).

Table 4
The rank correlation of the intensity of factors influencing migration between cohorts at the same spatial level.

0–65 and 66–74			66–74 and 75+			0–65 and 75+		
LAU 1	NUTS 3	NUTS 2	LAU 1	NUTS 3	NUTS 2	LAU 1	NUTS 3	NUTS 2
0.53	0.69	0.62	0.73	0.70	0.32	0.53	0.55	0.13

Source: Authors' work

Table 5
The rank correlation of the intensity of factors influencing migration of the same cohorts between two spatially designed levels.

LAU 1 level			NUTS 3 level			NUTS 2 level		
0–65	66–74	75+	0–65	66–74	75+	0–65	66–74	75+
0.49	0.57	0.83	0.83	0.50	0.39	0.44	0.40	0.56

Source: Authors' work

A comparison of preferences between the age cohorts was carried out using a correlation analysis. For this purpose, we calculated the Spearman correlation coefficient, which measures the degree of similarity between two rankings (two columns in Table 3). The results are available in Tables 4 and 5. We can see a significant correlation in the ranking of factors between the population aged 66-74 and 75+ at municipal and NUTS3 levels. There is also a significantly similar ranking between the municipal and NUTS 3 levels for the 75+ cohort and between the NUTS 3 and NUTS 2 levels for the 0-65-year-old cohort.

Conclusions

The gravity model for goods, finance or population in economic geography is based on Newton's findings. Newton's law states that two bodies always attract each other with a force proportional to the mass at the point of origin and the destination and that the force of attraction decreases proportionally to the square of the distance between these objects. In international economics, this law was introduced by Walter Isard (1954), who analysed bilateral trade flows as a function of economic size and distance between two spatial units – countries. Initial results have already shown that trade tends to decrease with increasing distance. Today, these aspects are also integrated into the global supply chain (Bogataj et al., 2024). The study of migration flows also came to similar results. The first study of this kind was published in 1966 by Cristian and Braden in the Web of Science Core Collection (WoS). 455 articles dealing with migration using the gravity approach can be found as results in the WoS. Having established that populations in Europe and other developed countries are ageing rapidly, the results of our research provide empirical support to a growing number of publications refuting the classification of migrants as a monolithic group without considering their age, highlighting differences in the requirements of older cohorts, including baby boomers and older generations, for age-friendly communities where most residents are over 50 years of age (e.g. Karpestam 2018; Black & Hyer, 2020; De Jong et al., 2020). Study results show significant differences between all cohorts in all WHO domains (see Figure 3), with substantial differences in preferences for housing, open space, employment and participation in social activities (World Health Organisation, 2007). Our results show differences between cohorts.

The research question was whether investments in space and other factors have significantly different effects on the intensity of migration flows depending on the age of the inhabitants. The answer is positive. In this study, we investigated the influence of various factors on the migration pattern of three different age cohorts among Slovenian municipalities and regions at two regional spatial levels. Particularly, we have studied the impact of the fastest time-spending distance between municipal centres, the number of inhabitants in the municipality, the registered unemployment rate, the gross earning per capita in the municipality, the number of dwellings per 1000 inhabitants, the average price per m^2 of dwelling in the municipality, municipal revenue per capita, ageing index in the municipality, the capacity of the care homes, coefficient of the number of single rooms in the care homes, the number of temporary accommodation places, the number of day-care places, the number of dwellings serviced by the care provider, and the index of coverage of the care needs of older people in the care homes in the municipality. All factors have been measured in the origin and destination. The flows have been considered between municipalities, NUTS 2 regions and statistical regions. The importance of these factors has been ranked for all levels of spatial units. There has been a very low correlation of ranking factors between cohorts 0–65 years old and 75+ years old inhabitants for migrations on the NUTS 2 level; the correlation of rank r_s was only 0.13, which means that priorities which attract the migrants or retain them are very different between population 0–65 years old and population 75+ in case of NUTS 2 migrations. We can conclude that the same is true for cohorts 6674 and 75+ years old inhabitants (r_s is only 0.32). It proves that younger cohorts have priorities, unlike the very old cohorts, when they wish to migrate to other NUTS 2 levels. A low correlation of ranks was also found between cohorts 75+ for migrations on NUTS 2 and NUTS 3 levels. When they migrate to other NUTS 2 regions, they change the ranking of priorities. The details are given in Tables 4 and 5.

Our findings contribute to the understanding of migration dynamics. They can provide policymakers and spatial planners with information on how to meet the needs of different age groups better as their share in the population structure changes over the years. The research results are based on the existing hierarchy of spatial units. The results can be improved by using location data of the individual places of origin and destination (geocodes of the residential units and destinations of the residents). A first improvement would also be possible by considering into account the delimitation of functional regions, which is the subject of our further research.

Acknowledgements

The authors acknowledge the financial support from the Slovenian Research Agency through the research project Hierarchical Design and Financing of Social Infrastructure of Smart Silver Villages (L7-3188), Evaluation of the sustainable development of urban space through the parameters of social infrastructure development and life satisfaction (J5-3112) and from research core funding Earth Observation and Geoinformatics (P2-0406).

References

1. Angel, S., Parent, J., Civco, D. L., & Blei, A. M. (2010). *The Persistent Decline of Urban Densities: Global and Historical Evidence of Sprawl*, available at http://www.lincolninst.edu/pubs/1834_The-Persistent-Decline-in-Urban-Densities (1 December 2023)
2. Angel, S., Parent, J., Civco, D. L., Blei, A. & Potere, D. (2011). The Dimensions of Global Urban Expansion: Estimates and Projections for all Countries, 2000-2050, *Progress in Planning, 75*(2), 2, 53–107. https://doi.org/10.1016/j.progress.2011.04.001
3. Arends-Kuenning, M., Baylis, K., & Garduño-Rivera, R. (2018). The effect of NAFTA on internal migration in Mexico: a regional economic analysis. *Applied Economics*, *51*(10), 1052–1068. https://doi.org/10.1080/00036846.2018.1524976
4. Black K., & Hyer K. (2020). Generational Distinctions on the Importance of Age-Friendly Community Features by Older Age Groups. *Journal of Applied Gerontology, 39*(9), 1025-1034. doi: 10.1177/0733464819847885.
5. Bogataj, D., Battini, D., Calzavara, M., & Persona, A. (2019a). The Ageing Workforce Challenge: Investments in Collaborative Robots or Contribution to Pension Schemes, From the Multi-Echelon Perspective, *International Journal of Production Economics, 210*, 97–106. https://doi.org/10.1016/j.ijpe.2018.12.016
6. Bogataj, D., Bogataj, M., & Drobne, S. (2019b). Interactions Between Flows of Human Resources in Functional Regions and Flows of Inventories in Dynamic Processes of Global Supply Chains, *International Journal of Production Economics, 209*, 215–225. https://doi.org/10.1016/j.ijpe.2017.10.018
7. Bogataj, D., Bolarín, F. C., Kavšek, M., & Rogelj, V. (2020a). Smart Silver Villages as Part of Social Infrastructure for Older Adults in Rural Areas, *IFAC PapersOnLine, 53*(2), 16914–16919. https://doi.org/10.1016/j.ifacol.2020.12.1233
8. Bogataj, D., Rogelj, V., Drobež, E., & Salaj, A. T. (2020b). Ambient Assisted Living in Lifetime Neighbourhoods, *IFAC PapersOnLine, 53*(2), 16896–16901. https://doi.org/10.1016/j.ifacol.2020.12.1226

9. Bogataj, D., Bogataj, M., Black, K., Hyer, K., & Kavšek, M. (2020c). Distinctions on the Importance of Age-Friendly Services by Old Age Group : A Comparative Study between the USA and the EU, *Innovation in aging, 4*(1), 32.
10. Bogataj, D., Campuzano-Bolarin, F., Moreno Nicolás, J. A., & Sethi, S. P. (2024). Gravity modelling of intercontinental supply chains considering a new port location. *Central European Journal of Operations Research*. https://doi.org/10.1007/s10100-024-00909-0
11. Bogataj, M., & Drobne, S. (2011). Decision Making in Competitivness of Regions, *Croatian Operational Research Review, 2*, 194–207.
12. Calzavara, M., Battini, D., Bogataj, D., Sgarbossa, F., & Zennaro, I. (2020). Ageing workforce management in manufacturing systems: state of the art and future research agenda. *International Journal of Production Research, 58*(3), 729-747. https://doi.org/10.1080/00207543.2019.1600759
13. Christian, W., Braden, W. (1966). Rural Migration and Gravity Model, *Rural Sociology, 31*(1), 73–80.
14. Conway, K. S., & Houtenville, A. (2003). Out with the Old, in with the Old. A Closer Look and Younger Versus Older Elderly Migration, *Social Science Quarterly, 84*, 309–328. https://doi.org/10.1111/1540-6237.8402006
15. Cunningham, C. R., & Engelhardt, G. V. (2008). Housing Capital-gains Taxation and Homeowner Mobility: Evidence from the Taxpayer Relief Act of 1997, *Journal of Urban Economics, 63*(3), 803–815. https://doi.org/10.1016/j.jue.2007.05.002
16. CSSI (2023). Community of Slovenian Social Institutions. Available at https://www.ssz-slo.si/splosno-o-posebnih-domovih/pregled-kapacitet-in-pokritost-institucionalnega-varstva-starejsih-in-posebnih-skupin-odraslih/ (15 August 2023)
17. Das, B., Singh, A., Panda, S. N., & Yasuda, H. (2015). Optimal Land and Water Resources Allocation Policies for Sustainable Irrigated Agriculture, *Land Use Policy, 42*, 527–537. https://doi.org/10.1016/j.landusepol.2014.09.012
18. De Jong, P. W., Adsera, A., & De Valk, H. A. G.(2020). The Role of Welfare in Locational Choice: Modelling Intra-European Migration Decision Across the Life-Course, Tijdshrift voor Economische en Sociale Geografie, 111(2), 163–181. https://doi.org/10.1111/tesg.12390
19. Drobne, S., Janež, P., Bogataj, M., & David, B. (2019). Taxation and revenues of central places as the regulator of available human resources in manufacturing - related to the bid rent curve. *IFAC-PapersOnLine, 52*(13), 2674-2679. https://doi.org/10.1016/j.ifacol.2019.11.611
20. Drobne, S., & Bogataj, M. (2022). Migration Flows through the Lens of Human Resource Ageing. *Business Systems Research Journal, 13*(3), 47-62. https://doi.org/10.2478/bsrj-2022-0024
21. Engelhardt, G. V. (2003). Nominal Loss Aversion, Housing Equity Constraints, and Housing Mobility: Evidence from the United States, *Journal of Urban Economics, 53*, 171–195. https://doi.org/10.1016/S0094-1190(02)00511-9
22. Espon. (2013). ATTREG - Attractiveness of European Regions and Cities for Residents and Visitors. https://www.espon.eu/programme/projects/espon-2013/applied-research/attreg-attractiveness-european-regions-and-cities
23. European Commission. (2023). 2024 Ageing Report. Underlying Assumptions and Projection Methodologies, Economy and Finance, Institutional Paper 257, available at https://economy-finance.ec.europa.eu/publications/2024-ageing-report-underlying-assumptions-and-projection-methodologies_en (6 January 2024)
24. Fang, C., Cui, X., Li, G., Bao, C., Wang, Z., Ma, H., Sun, S., Liu, H., Luo, K., & Ren, Y. (2019). Modeling Regional Sustainable Development Scenarios using the Urbanization and Eco-Environment Coupler: Case Study of Beijing-Tianjin-Hebei Urban Agglomeration, China, *Science of The Total Environment, 689*, 820–830. https://doi.org/10.1016/j.scitotenv.2019.06.430
25. Ferreira, F., Gyourko, J., & Tracy, J. (2010). Housing Busts and Household Mobility, *Journal of Urban Economics, 68*(1), 34–45. https://doi.org/10.1016/j.jue.2009.10.007

26. Groth, J., & Corijn, E. (2005). Reclaiming Urbanity: Indeterminate Spaces, Informal Actors and Urban Agenda Setting. *Urban Studies*, *42*(3), 503-526. https://doi.org/10.1080/00420980500035436
27. Gu, H. Y., Jie, Y. Y., & Lao, X. (2022). Health Service Disparity, Push-Pull Effect, and Elderly Migration in Ageing China, *Habitat International, 125*, 102581. https://doi.org/10.1016/j.habitatint.2022.102581
28. Hanhörster, H. (2009). Migration Dynamics in a Shrinking City, A Qualitative Investigation of Urban Internal Removals, 63(2), 196–197.
29. Hillmann, F. (2009). Migration Dynamics in a Shrinking City. A Qualitative Study of Intra-Urban Moves, Geographische Zeitschrift, 97(2–3) , 170–172.
30. Isard, W. (1954). Location Theory and Trade Theory: Short-Run Analysis, *Quarterly Journal of Economics, 68*(2), 305–320.
31. Johnston, R. J. (1970). Latent Migration Potential and Gravity Model - New-Zealand Study, *Geographical Analysis, 2*(4), 387–397.
32. Karpestam, P. (2018). Who Benefits from More Housing? A Panel Data Study on the Role of Housing in the Intermunicipal Migration of Different Age Cohorts in Sweden, *Review of Regional Studies, 48*(3), 401–425. https://doi.org/10.52324/001c.7994
33. Li, W., Li, H., Wang, S., & Feng, Z. (2022). Spatiotemporal Evolution of County-Level Land Use Structure in the Context of Urban Shrinkage: Evidence from Northeast China, *Land, 11*(1), 1709, https://doi.org/10.3390/land11101709
34. Li, W. C., Li, H., Hao, F. L., Feng, Z. Q., & Wang, S. J. (2024). A Simulation of the Spatial Expansion Process of Shrinking Cities Based on the Concept of Smart Shrinkage: A Case Study of the City of Baishan, Land, 13(2), 239. https://doi.org/10.3390/land13020239
35. Liang, X., Liu, X., Li, X., Chen, Y., Tian, H., & Yao, Y. (2018). Delineating Multi-Scenario Urban Growth Boundaries with a CA-based FLUS Model and Morphological Method, Landscape and Urban Planning, 177, 47–63. https://doi.org/10.1016/j.landurbplan.2018.04.016
36. Lin, H. Z. (2020). The Shrinking of Beijing and the Rising of Xiong'an: Optimize Population Migration in Terms of Transport Service, *Discrete Dynamics in Nature and Society, 2020*, 8282070. https://doi.org/10.1155/2020/8282070
37. MFRS. (2023). *Ministry of Finance of the Republic of Slovenia*, available at https://www.gov.si/en/state-authorities/ministries/ministry-of-finance/ (15 August 2023)
38. Oswalt, P., & Bundes, K. (2005). *Shrinking Cities: International Research*, Distributed Art Pub Incorporated: New York, NY, USA.
39. Rienow, A., Stenger, D., & Menz, G. (2014). Sprawling Cities and Shrinking Regions - Forecasting Urban Growth in the Ruhr for 2025 By Coupling Cells and Agents, *Erdkunde, 68*(2), 85–107.
40. SIR. (2023). Slovenian Infrastructure Agency, available at https://www.gov.si/en/state-authorities/bodies-within-ministries/slovenian-infrastructure-agency/ (15 August 2023)
41. SMARS (2023). Surveying and Mapping Authority of the Republic of Slovenia, available at https://www.gov.si/en/state-authorities/bodies-within-ministries/surveying-and-mapping-authority/ (15 August 2023)
42. Smith, M. H., & Smith, G.N. (2007). Homeownership in an Uncertain World with Substantial Transaction Costs, *Journal of Regional Science, 47*, 881–896. https://doi.org/10.1111/j.1467-9787.2007.00534.x
43. SORS. (2023). SiStat Database – SURS, Statistical Office of the Republic of Slovenia, available at https://pxweb.stat.si/SiStat/en (15 August 2023)
44. Wang, Z., Li, X., Mao, Y., Li, L., Wang, X., & Lin, Q. (2022a). Dynamic Simulation of Land Use Change and Assessment of Carbon Storage Based on Climate Change Scenarios at the City Level: A Case Study of Bortala, China, Ecological Indicators, 134, 108499. https://doi.org/10.1016/j.ecolind.2021.108499
45. Wang, J., Zhang, J., Xiong, N., Liang, B., Wang, Z., & Cressey, E. L. (2022b). Spatial and Temporal Variation, Simulation and Prediction of Land Use in Ecological Conservation Area of Western Beijing, Remote Sensing, 14(6), 1452–1472. https://doi.org/10.3390/rs14061452

46. Wolff, M., Haase, D., & Haase, A. (2018). Compact or spread? A quantitative spatial model of urban areas in Europe since 1990. *PLOS ONE, 13*(2), e0192326. https://doi.org/10.1371/journal.pone.0192326
47. World Health Organization (2007). Global age-friendly cities: A guide, available at http://www.who.int/ageing/publications/Global_age_friendly_cities_Guide_English.pdf (24 June 2023)
48. Zuo, Z., Li, Y., Guo, H., & Cheng, J. (2023). Spatial Evolution and Decomposition of Energy-Related CO_2 Emissions in China's Mining Industry: From the Perspective of Regional Heterogeneity, Environmental Science and Pollution Research, 30(45), 101599–101615. https://doi.org/10.1007/s11356-023-29244-8

About the authors

Samo Drobne, PhD, is an Associate Professor of Geodesy and Geoinformatics at the Faculty of Civil and Geodetic Engineering, University of Ljubljana, Slovenia. He teaches courses on statistics, geographical information systems (GIS), and spatial analyses in GIS. His main research areas include regional development and planning, spatial interaction models, functional regions, commuting, migration, spatial analysis in GIS and operational research in spatial systems. He is actively involved in several international and national research projects. The author be contacted at **samo.drobne@fgg.uni-lj.si**.

Marija Bogataj, PhD, is a professor of operational research and statistics who was promoted at the University of Ljubljana in 1995 and head of the CERRISK Research Group at the Institute INRISK, Slovenia. Her main research areas include management and control of the supply chains, including the supply network in long-term care, spatial interaction models, commuting, migration, and operational research in spatial systems. She is also an editor for three WoS-indexed journals. She is actively involved in several international and national research projects. The author can be contacted at **marija.bogataj@guest.arnes.si**.

Beyond Parametric Bounds: Exploring Regional Unemployment Patterns Using Semiparametric Spatial Autoregression

Andrea Furková, Peter Knížat
University of Economics in Bratislava, Slovak Republic

Abstract

Background: It is a well-known phenomenon that nonlinearities that are inherent in the relationship among economic variables negatively affect the commonly used estimators in the econometric models. The nonlinearities cause an instability of the estimated parameters that, in particular, are unable to capture a local relationship between the response and the covariate. **Objectives:** The main aim of the paper is the simultaneous consideration of spatial effects as well as nonlinearities through an advanced semiparametric spatial autoregressive econometric model. The paper seeks to contribute to empirical studies of regional science focused on the application of semiparametric spatial autoregressive econometric models. **Methods/Approach:** We outline an approach that can be used to correct nonlinearities by incorporating a semiparametric idea within the framework of econometric models. We use an expansion by penalised basis splines that are highly flexible and are able to capture local nonlinearities between variables. **Results:** In the empirical study, we fit different econometric models that attempt to explain the dynamics of the European Union's regional unemployment. **Conclusions:** The results show that regional unemployment exhibits significant spatial dependence, indicating interconnectedness among neighbouring regions and suggesting the adoption of a semiparametric spatial autoregressive model for improved modelling flexibility, surpassing traditional parametric approaches.

Keywords: regional unemployment; linear regression; semiparametric model; generalised additive model; spline regression; spatial autoregressive semiparametric model
JEL classification: C14, C21, C52
Paper type: Research article

Received: Feb 2, 2024
Accepted: Jun 9. 2024

Acknowledgements: This work was supported by the Grant Agency of the Slovak Republic VEGA 1/0047/23.

Citation: Furková, A. Knížat, P. (2024). Beyond Parametric Bounds: Exploring Regional Unemployment Patterns Using Semiparametric Spatial Autoregression. *Business Systems Research, 15*(2), 48-66.
DOI: doi.org/10.2478/bsrj-2024-0017

Introduction

The use of semiparametric models in regional science is not as common as traditional parametric models, but it has gained attention and traction in recent years. Regional science involves the study of spatial patterns, relationships, and dynamics within specific geographic areas, making it a field where spatial econometrics plays a crucial role. The choice between a semiparametric model and a parametric spatial autoregressive (SAR) model (for more details, see, e.g., Anselin and Rey, 2014; Lung-Fei, 2022 or Chi and Zhu, 2019) depends on the specific characteristics of the data and the underlying assumptions of the modelling approach. Both approaches have their advantages and disadvantages, and the decision should be made based on the context and goals of the analysis. In regions with complex spatial patterns, where simple parametric forms do not easily capture the relationships between variables, semiparametric models can offer advantages. These models allow for more flexibility in accommodating spatial heterogeneity (see, e.g., Anselin and Rey, 2014). Semiparametric models (see Basile and Mínguez, 2018; Perperoglou et al., 2019), particularly those incorporating splines or other flexible functions, are useful when dealing with nonlinear relationships. Regional data may exhibit nonlinear patterns that linear parametric models cannot adequately capture. Semiparametric models are well-suited for capturing local variations within regions. Regional science often involves studying spatially heterogeneous phenomena, and semiparametric models can adapt to these variations more effectively than rigid parametric models. Semiparametric models may be more robust to misspecification and outliers, which can be important in the presence of spatial dependencies and complex regional dynamics. On the other hand, parametric models often have the advantage of being more interpretable, with coefficients corresponding to specific parameters. However, as the understanding and acceptance of semiparametric models grow, researchers are finding ways to interpret and communicate results from these models effectively. While semiparametric models offer advantages, researchers should carefully consider the trade-offs, including model complexity, interpretability, and computational demands.

One of the motivational factors of this paper is to contribute to filling the gap of empirical studies in regional science on the application of semiparametric spatial autoregressive econometric models. In this paper, we deal with the problem of unemployment in the regions of the European Union (EU). The novelty of the study can be seen in the simultaneous consideration of spatial effects (spatial autocorrelation and spatial heterogeneity) as well as nonlinearities in the functional form through an advanced semiparametric spatial autoregressive econometric model. This paper aims to investigate the dynamics of regional unemployment through the application of advanced econometric models, with a primary focus on the semiparametric spline spatial autoregressive model. The overarching goal is to enhance our understanding of the spatially-dependent nature of unemployment patterns across different regions. Our specific objectives – hypotheses of this research include:

- *Hypothesis 1 (Spatial Dependence Hypothesis): Regional unemployment rates exhibit significant spatial dependence, indicating that the unemployment rates influence the unemployment status in one region in neighbouring regions.*
- *Hypothesis 2 (Nonlinear Relationships Hypothesis): The relationship between regional unemployment and its determinants is nonlinear, and a semiparametric spline spatial autoregressive model can more effectively capture these nonlinearities than traditional parametric models.*

We commence our study by reviewing a classical linear regression model, where the model's parameters are estimated through the method of ordinary least squares.

The study of linear models is now part of any standard textbook on the introduction to multivariate statistical analysis, e.g., the manuscript in Fahrmeir et al. (2021). To extend the linear model to a more flexible framework, i.e., in the case when the relationship of the response and covariates exhibits some nonlinearities, we introduce basis spline functions and use these to transpose individual covariates into the functional form, which is then regressed onto the response. A basis spline functions are piece-wise polynomials joined in breakpoints, also called knots, which can easily be extrapolated onto a real-valued variable. A thorough theoretical discussion and application of basis spline functions are discussed by Perperoglou et al. (2019).

In the context of the nonlinear spline regression, the estimated parameters correspond to each polynomial, with a number of polynomials (spline curves) defined by the user or through a cross-validation procedure. On the other hand, a generalised additive model (GAM) uses a local smoothing algorithm for the estimation of the regression function and hence belongs to a family of fully nonparametric models. The advantage of GAM is that it further relaxes the assumption of the linearity that parameters in the model would normally restrict. GAM was first proposed by Hastie and Tibshirani (see Hastie and Tibshirani, 1986; 1990), and its most recent theory, including the application in R software, is provided by Wood (2017) and Wood (2023). In Mínguez et al. (2022) authors introduce a new R package for the estimation of flexible semiparametric spatial autoregressive models, which makes it possible to control for spatial dependence simultaneously, nonlinearities in the functional form, and spatiotemporal heterogeneity.

It is a well-known phenomenon that economic data that are observed in specific locations are affected by observations from neighbouring locations, which is called a spill-over effect. The manuscripts Lung-Fei (2022) provide a comprehensive review of spatial regression models that are used for spatial observations in the context of econometrics. We outline a theoretical framework of spatial autoregressive models that incorporate spatial spillover effects. Similarly, we follow by extending the spatial parametric model into the more flexible nonparametric approach, which should capture the nonlinearities between the response and covariates and hence improve the functional form of the estimated model. The paper by Basile et al. (2014) demonstrates the estimation technique for the spatial semiparametric model, which is carried out by using a 2-step "control function" approach since the two-stage least squares method might lead to inconsistent estimates of the regression parameters. In Basile and Mínguez (2018), a critical review of parametric and semiparametric spatial econometric approaches can be found. The author focuses on the capability of each class of models to fit the main features of spatial data (such as strong and weak cross-sectional dependence, spatial heterogeneity, nonlinearities, and time persistence).

As we have already stated, the application of semiparametric models in regional science is not as common as traditional parametric models. We can find the use of this approach in the works of Wahyuni and Fajri (2020) or Mínguez et al. (2022). However, only a few empirical works apply the semiparametric spatial autoregressive econometric approach in connection with the modelling of regional economic problems. From this point of view, we believe that this paper might contribute to supplement empirical analyses of this nature.

The rest of the paper was structured as follows: the methodology section provides the main theoretical background, and the results section presents an overview of a study area, a description of the data, model specification, and main empirical results. The main concluding remarks are presented in Discussion and Conclusion sections. The paper closes with References.

Methodology

Let us assume that we observe a matrix **X** = (x_{ik}), where $i = 1, 2, \ldots, N$ refers to the sample unit that is observed for each covariate $k = 1, 2, \ldots, K$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be the fixed vectors in $\mathbb{R}^K$ and let $y_1, \ldots, y_N$ be dependent variables. In general, the functional relationship between the response vector y_i and the covariate $\mathbf{x}_i$ can be expressed as:

$$y_i = f(\mathbf{x}_i) + \varepsilon_i, \qquad i = 1, 2, \ldots, N \tag{1}$$

where $\varepsilon_1, \ldots, \varepsilon_N$ are independent and identically distributed (*i.i.d.*) random errors with mean zero and variance σ_ε^2, i.e., $\varepsilon_i \sim i.i.d.(0, \sigma_\varepsilon^2)$. The function $f(.)$ can be of the parametric or nonparametric form.

In the following subsections, we introduce parametric regression models that can be described by a finite number of estimated parameters. The estimated parameters determine the model's functional form. Subsequently, we outline nonparametric regression models that do not require a predetermined functional form but are constructed according to information derived from data.

Linear regression model and its extension to nonlinear spline regression

A linear Ordinary Least Squares (OLS) regression model can be expressed as (Fahrmeir et al., 2021):

$$y_i = \alpha + \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i, \qquad \varepsilon_i \sim i.i.d.(0, \sigma_\varepsilon^2) \tag{2}$$

where $\mathbf{x}_i$ represents a $1 \times K$ vector of covariates with associated parameters $\boldsymbol{\beta}$ contained in a $K \times 1$ vector and α is the intercept. The OLS method is used to estimate the parameters, which yields the following (ibid):

$$\boldsymbol{\beta} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} \tag{3}$$

where $\boldsymbol{\beta} = \left(\alpha, \beta_1, \beta_2, \ldots \beta_p\right)$, noting that the matrix **X** includes ones in its first column for the estimation of α and **y** denotes a vector of dependent variable. The matrix $\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)$ is of full rank in order to be invertible.

In many regression scenarios, the relationship between the response and covariates exhibit local nonlinearities, which implies that the parametric model can be too restrictive. In this case, the functional form of the Eq. (1) is mis-specified and its estimated values y_i lie far off the observed values y_i.

To capture the local nonlinearities between the response and covariates, a nonlinear regression model with basis spline functions, which are extrapolated onto each covariate, and hence replace matrix **X**, could be more suitable. The covariate matrix can be expressed in terms of basis spline expansion as follows (Perperoglou et al., 2019):

$$\mathbf{X} = \mathbf{B}_\delta^m(\tau_l)\mathbf{C}_\delta \tag{4}$$

where the spline curves $\mathbf{B}_{\delta}^{m}\left(\tau_{l}\right)$ are piecewise polynomials of order *m* that are merged at the break points, also called knots, τ_{l}, $l=1,2,\ldots,L-1$, where $\delta=1,2,\ldots,\Delta$, refers to the number of spline curves. $\mathbf{C}_{\delta}$ is a $\Delta\times K$ matrix of parameters that needs to be estimated. The reader interested in more theoretical details of basis spline functions can consult Perperoglou et al. (2019).

Within the context of the nonlinear spline regression model, the estimation of parameters in the matrix $\mathbf{C}_{\delta}$ is carried out through an OLS method that minimises the sum of squares errors, which yields the following (omitting τ_{l} and superscript *m* for simplification):

$$\mathbf{C}_{\delta}=\left(\mathbf{B}_{\delta}^{\mathrm{T}}\mathbf{B}_{\delta}\right)^{-1}\mathbf{B}_{\delta}^{\mathrm{T}}\mathbf{y} \tag{5}$$

The interpretation of estimated parameters $\mathbf{C}_{\delta}$ is more elaborative than in the classical linear regression since each $\mathbf{C}_{\delta}$ is linked to intervals of responses y_i and the covariate $\mathbf{x}_i$. Therefore, the model is able to capture the local nonlinearities between the response and covariates specific to these intervals.

Generalised additive regression model

The generalised additive model (GAM) is considered a nonparametric version of the nonlinear model, where the linear form $\boldsymbol{\beta}\mathbf{X}$ is replaced by a sum of unspecified functions $g\left(\mathbf{x}_{i}\right)$ that are estimated through a method of the local backfitting (smoothing) algorithm, first proposed by Hastie and Tibshirani (1986). The user can define various forms of smooth functions in *g*(.). We opt to use basis spline functions that have various advantages, refer to Perperoglou et al. (2019) for details.

GAM can be expressed as (Wood, 2017):

$$y_{i}=\alpha+\sum_{p=1}^{P}g\left(\mathbf{x}_{i}\right)+\varepsilon_{i},\qquad\qquad \varepsilon_{i}\sim i.i.d.\left(0,\sigma_{\varepsilon}^{2}\right) \tag{6}$$

The estimation method of the GAM is based on the minimisation of the cross-validation sum of squares (*CVSS*):

$$CVSS\left(k\right)=\frac{1}{N}\sum_{i=1}^{N}\left(y_{i}-\alpha-\sum_{p=1}^{P}g^{-i}\left(\mathbf{x}_{i}\right)\right)^{2} \tag{7}$$

where $g^{-i}\left(\mathbf{x}_{i}\right)$ is the basis spline function with k number of basis, having removed one observation $\left(x_{i},y_{i}\right)$ from the sample at each iteration. The estimation procedure of minimising *CVSS*(*k*) is a repetitive smoothing of the dependent variable y_i on $\mathbf{X}_i$, which is carried out through *a local backfitting* algorithm. The iterative procedure is described in details in Wood (2017), with an application in R software.

Spatial autoregressive regression model and its extension to nonlinear spline regression

In the socioeconomic problem, we usually observe data from regional economic activities that are known to be regionally correlated, i.e., an observation from the location (region) *i* is affected by observations from other locations *j*, where $i \neq j$, also called spatial spill-over effects. In general, a formal expression of the spatial correlation between different locations can be defined in terms of corresponding non-zero covariances (Anselin and Rey, 2014):

$$\operatorname{cov}\left[y_i y_j\right] = \mathrm{E}\left[y_i y_j\right] - \mathrm{E}\left[y_i\right]\mathrm{E}\left[y_j\right] \neq 0 \qquad i \neq j \tag{8}$$

where E refers to the expected value and y_i and y_j are observed values from regions *i* and *j*, respectively. So, any influences that spread from one location to nearby ones (spatial spill-over effects) should be taken into account when building the regression model. Traditional spatial econometric estimation framework is based on models with spatially autoregressive process, the models that explicitly allow for spatial dependence through spatially lagged variables. The type of spatial model can be determined using *LM* tests (see, e.g., Anselin and Rey, 2014). One of the well–known model from this class is SAR (Spatial Autoregressive) model, which assumes spatial spillover effects within the dependent variable *y*. We present this model in relation to our empirical analysis. The SAR model is formulated as follows (Anselin and Rey, 2014):

$$y_i = \rho \mathbf{W}\mathbf{y} + \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i, \qquad \varepsilon_i \sim i.i.d.\left(0, \sigma_\varepsilon^2\right) \tag{9}$$

where ρ is a spatial autoregressive parameter, **Wy** denotes a spatially lagged dependent variable and **W** is a $N \times N$ spatial weighting matrix. In this paper, queen contiguity spatial weighting matrix was used in all spatial econometric models and spatial statistics calculations. Due to possible problems with isolated units as well as with high variability of neighbouring regions resulting from other approaches, the queen contiguity form seemed to be suitable for determining spatial regional structures. In the case of spatial weights, for instance, based on a distance function, inverse or radial (see, e.g., Pavlovčič-Prešeren et al., 2019), there can be a problem with the bimodal distribution when some regions have very few neighbours. On the other hand, the other regions have very many neighbouring units. In the scientific and empirical literature, there are many other traditional definitions for the spatial structure among spatial locations (see, e.g., Lung-Fei, 2022 or Chi and Zhu, 2019).

Estimation of models with spatial autocorrelation and/or spatial heterogeneity requires special estimation methods and procedures. For instance, the estimation of spatial autoregressive models (e.g., SAR model) is affected by the presence of the spatially lagged variable **Wy** on the right-hand side of the regression equation, which causes problems with endogeneity. Therefore, OLS is not a suitable estimation method. The estimation of such models is based on familiar econometric estimation methods, but they must be modified with respect to spatial aspects: Maximum Likelihood (ML), Two-Stage Least Squares (2SLS) or Generalised Moment Method (GMM). A review of these estimation methods can be found in Anselin and Rey (2014) or Chi and Zhu (2019).

In general, spatial autoregressive models are sometimes unfeasible in the presence of model misspecification. Geniaux and Martinetti (2018) pointed out that it can often be problematic to disentangle between a real spatial autocorrelation and different sources of violation of *i.i.d.*, such as spatial heterogeneity through unobserved covariates and spatially varying relationships. Modelling spatial data requires flexible econometric tools that allow us to control spatial dependence, spatial heterogeneity, non-linearities and other possible model specification biases. To address this demand for flexibility, the adoption of the nonparametric structure or semiparametric structure of the spatial regression model is advisable. Similarly, using basis spline functions for covariates, the spatial autoregressive semiparametric model can be defined as (Basile et al., 2014):

$$y_i = \rho \mathbf{W}\mathbf{y} + \alpha + \mathbf{x}_i \boldsymbol{\beta} + g(\mathbf{x}_i) + \varepsilon_i, \qquad \varepsilon_i \sim i.i.d.\left(0, \sigma_\varepsilon^2\right) \tag{10}$$

where spline basis expansions of original covariates are defined in Eq. (4). Some covariates could enter Eq. (10) in the parametric form, which can be determined through preliminary statistical analysis.

The estimation of Eq. (10) can be carried out by using either a restricted maximum likelihood (REML) or a 2-step "control function" approach, refer to Basile et al. (2014) for theoretical details. The REML approach combines penalised regression spline (PS) methods (see, e.g., Perperoglou et al., 2019) with standard spatial autoregressive models such as SAR defined in Eq. (9), Spatial Error Model (SEM) or Spatial Durbin Model (SDM). An important advantage of such models is that they make it possible to capture local nonlinearities within the specification of spatial autoregressive terms, i.e., to capture spatial interaction effects and parametric and nonparametric relationships. In addition, a geoadditive term, i.e., a smooth function of the spatial coordinates can be included in Eq. (10) to capture a spatial trend effect (to capture spatially autocorrelated unobserved heterogeneity).

Results

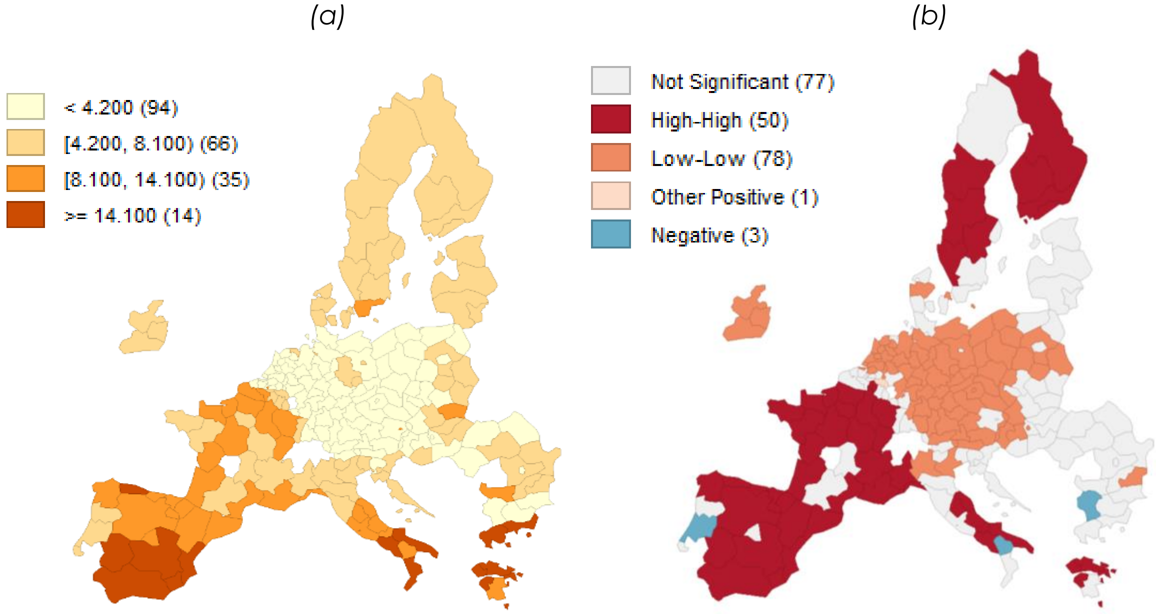
In this section, we apply the theoretical framework outlined in the previous section. We perceive the semiparametric SAR model defined by Eq. 10 to be highly useful for modelling cross-sectional spatial data considering nonlinearities, spatial dependence, and spatial heterogeneity. We empirically illustrate this model's performance in modelling the European unemployment problem.

Regional Unemployment Data

The paper uses data from the Eurostat regional statistical database (Eurostat, 2023). After excluding isolated observations (island regions) and missing data, the corrected database contains 209 European regions at the NUTS 2 level (NUTS—Nomenclature of territorial units for statistics). Figure 1 provides an overview of the study area. This figure shows a real spatial distribution (a) and local Geary cluster map (b) of Unemployment rates in 2019 across the EU regions.

The maps presented in Figure 1 already indicate disparities among the EU regions. In addition to regional disparities, we can also notice that regions are considerably clustered. The existence of strong positive spatial autocorrelation indicates the statistically significant value of global Moran's I statistic (0.683 with pseudo-p-value 0.001). The local Geary cluster map (see e.g., Chi and Zhu, 2019) provides more evidence about indicated unequal distribution and spatial clustering of the EU unemployment. Based on Figure 1 (b), we identify statistically significant locations – regions with positive spatial autocorrelation so-called hot spots and cold spots locations (50 high-high and 78 low-low locations). The high-high locations are mainly the regions of Spain and France. These regions are regions where high values of unemployment rates are clustered. Low-low values are mainly concentrated in the regions of Germany, Austria, the Netherlands, and some regions of Eastern Europe, such as the Czech Republic and Poland. This suggests that the geographic position of the region and the spatial regional spillovers probably affect the level of regional unemployment.

Figure 1
Spatial distribution of Unemployment rates in 2019 – natural breaks map (a) and Local Geary cluster map (b)

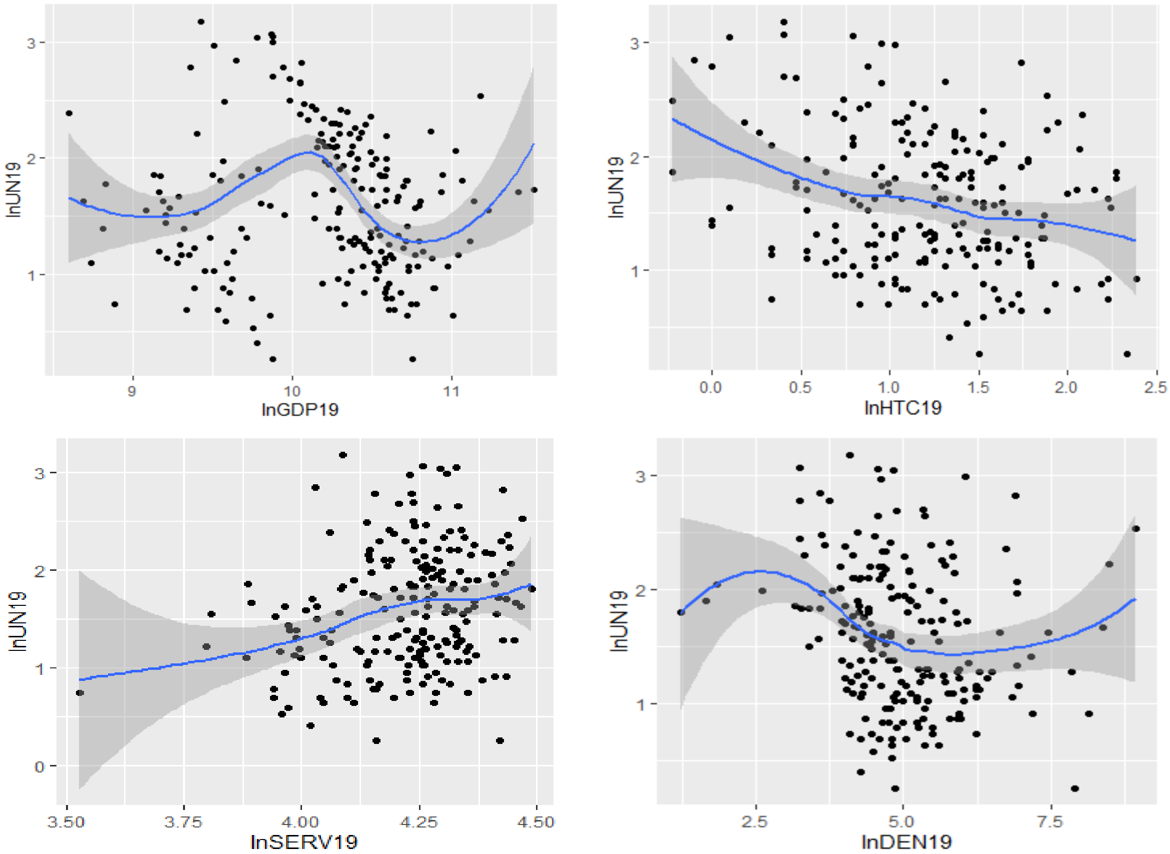


Source: Authors' work.
Note: The numbers in brackets indicate the number of regions in the category. Compared to the quartile map, the natural breaks criterion is better at grouping extreme observations. Interestingly, unlike quantile maps, the number of observations in each category can be highly unequal.

The initial empirical analysis will be based on a linear regression model that draws from the "regional competitiveness" theory (Formánek, 2019) explaining unemployment dynamics in terms of its key determining factors *GDP* – gross domestic product (Euro per inhabitant) and two convenient labour-force structure and competitiveness indicators: *HTC - employment in technology and knowledge-intensive sectors – high-technology sectors* (percentage of total employment), *SERV - employment in technology and knowledge-intensive sectors - services* (percentage of total employment). In addition, we also consider variable *DEN - population density* (persons per square kilometre) as a possible determinant of unemployment. All variables have logarithmically transformed forms and the observed period is 2019. Due to the skewed distribution of the dependent variable – Unemployment rates, we use log-transformation.

In the next step, we briefly examine the relationship between the response and each covariate. Figure 2 displays *one-to-one* relationships, with a fitted line using the *scatterplot* smoother and its corresponding 95% confidence intervals. It was created using the ggplot function in R.

Figure 2
Scatterplots of Unemployment vs four key determinant factors (GDP, HTC, SERV, DEN)



Source: Authors' work.

Figure 2 clearly shows that the *one-to-one* relationship between unemployment and all other key determinant factors is nonlinear. Short preliminary analyses indicate that an OLS regression might be far from sufficient to investigate the determinants of EU regional unemployment. First of all, we have seen, that spatial preliminary analysis reveals the problem of spatial autocorrelation and heterogeneity. On the other hand, the scatterplots of unemployment versus four key determinant factors, i.e. one-to-one relationships analysis, point to the problem of nonlinearity. It follows that the nonparametric regression could be a more flexible modelling of the effects of continuous covariates on the dependent variable since the classical linear model might not sufficiently capture local nonlinearities.

Econometric Models

The empirical part of the paper consists of the estimation of five econometric models to determine the factors affecting regional unemployment and to compare the performance of different specifications of the econometric models:

- **Model1**
  - non-spatial parametric (linear) model - OLS regression:

$$y_i = \alpha + \sum_{k=1}^{K} \beta_k x_{k,i} + \varepsilon_i, \quad i = 1, 2, \dots, N \qquad \varepsilon_i \sim i.i.d.\left(0, \sigma_\varepsilon^2\right) \tag{11}$$

- **Model2**
  - non-spatial nonlinear (non-parametric) model – spline regression:

$$y_i = \alpha + \sum_{\delta=1}^{\Delta} g_\delta\left(x_{\delta,i}\right) + \varepsilon_i, \quad i = 1,2,\dots,N \qquad \varepsilon_i \sim i.i.d.\left(0,\sigma_\varepsilon^2\right) \tag{12}$$

- **Model3**
  - SAR parametric (linear) model:

$$y_i = \rho\sum_{j=1}^{N} w_{ij} y_j + \alpha + \sum_{k=1}^{K} \beta_k x_{k,i} + \varepsilon_i, \quad i = 1,2,\dots,N \qquad \varepsilon_i \sim i.i.d.\left(0,\sigma_\varepsilon^2\right) \tag{13}$$

- **Model4**
  - semiparametric (nonlinear) SAR model without spatial trend:

$$y_i = \rho\sum_{j=1}^{N} w_{ij} y_j + \sum_{\delta=1}^{\Delta} g_\delta\left(x_{\delta,i}\right) + \varepsilon_i, \quad i = 1,2,\dots,N \qquad \varepsilon_i \sim i.i.d.\left(0,\sigma_\varepsilon^2\right) \tag{14}$$

- **Model5**
  - semiparametric (nonlinear) SAR model with spatial trend:

$$y_i = \rho\sum_{j=1}^{N} w_{ij} y_j + \sum_{\delta=1}^{\Delta} g_\delta\left(x_{\delta,i}\right) + \tilde{f}\left(s_{1i}, s_{2i}\right) + \varepsilon_i, \; i = 1,2,\dots,N \;\; \varepsilon_i \sim i.i.d.\left(0,\sigma_\varepsilon^2\right) \tag{15}$$

where y_i denotes the response variable and $x_{k,i}$ denotes the individual predictors, all defined in the section Regional Unemployment Data. Eq. 15 represents the spatial trend and denotes the spatial coordinates of ith region. We have already defined the other remaining terms in the section Methodology.

The spatial regression models defined in (13), (14) and (15) contain a spatial lag of the dependent variable. This means that the expected value of unemployment in the *i*th region is no longer influenced only by exogenous regional characteristics but also by the exogenous characteristics of all other regions through a spatial multiplier (for more details, see, e.g., Chi and Zhu, 2019)). The specifications of all spatial econometric models are based on the queen contiguity weights (matrix **W**) – these binary weights indicate whether regions share a boundary or not. As the last model, we introduce a semiparametric spatial model with a spatial trend (see Eq. 15) in order to control for unobserved spatial heterogeneity.

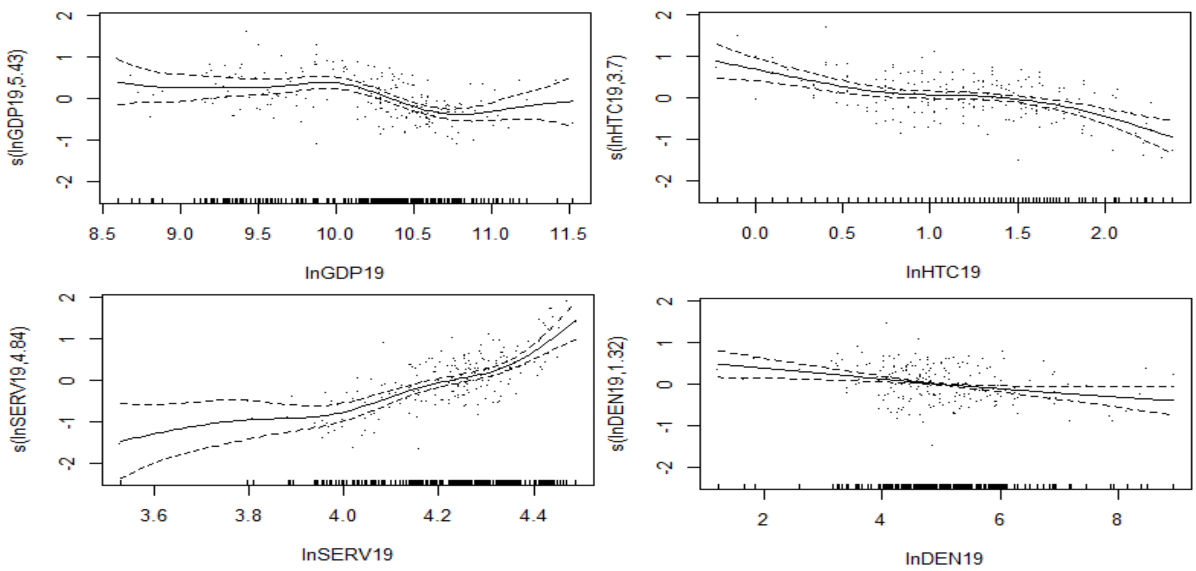
We estimated all models defined by (11) – (15) equations in the R package pspatreg (Mínguez et al., 2022). The non-parametric terms (either trends or covariates) were modeled using P-Splines. The estimation methods were maximum likelihood (ML) and restricted maximum likelihood (REML).

The performed analysis of one-to-one relationships pointed to the problem of nonlinearity. Nonparametric functions *g*(.) in models (12), (14) and (15) are formed by transposing a real-valued covariate by spline basis expansion. The spline functions have high flexibility and can handle data that changes in subintervals, which relates to local nonlinearities. In this context, it was crucial to determine which variables belong to the parametric or non-parametric component, to select the optimal knots and their location. We relied on the starting GAM model defined in Eq. 6, and we utilised adaptive knot selection methods used in the mgcv package in R (Wood, 2023), which automatically select knot locations based on data characteristics. The approach used for knot selection involves automatic smoothness selection using penalized likelihood methods. The resulting number of knots was 9 and was used in all non-parametric and semi-parametric models.

Empirical Results

Since the estimations of models (11) - (15) provide extensive estimation outputs, it is not possible to list them within the scope of this article. In this section, we present the most important outputs that allow us to evaluate the hypothesis stated in the Introduction. Other outputs are available at the request of the authors.Figure 3 shows plots of non-parametric covariates resulting from GAM model estimation.

Figure 3
Plots of terms of non-parametric covariates - GAM model



Note: Pointwise confidence intervals in dashed lines.
Source: Authors' work.

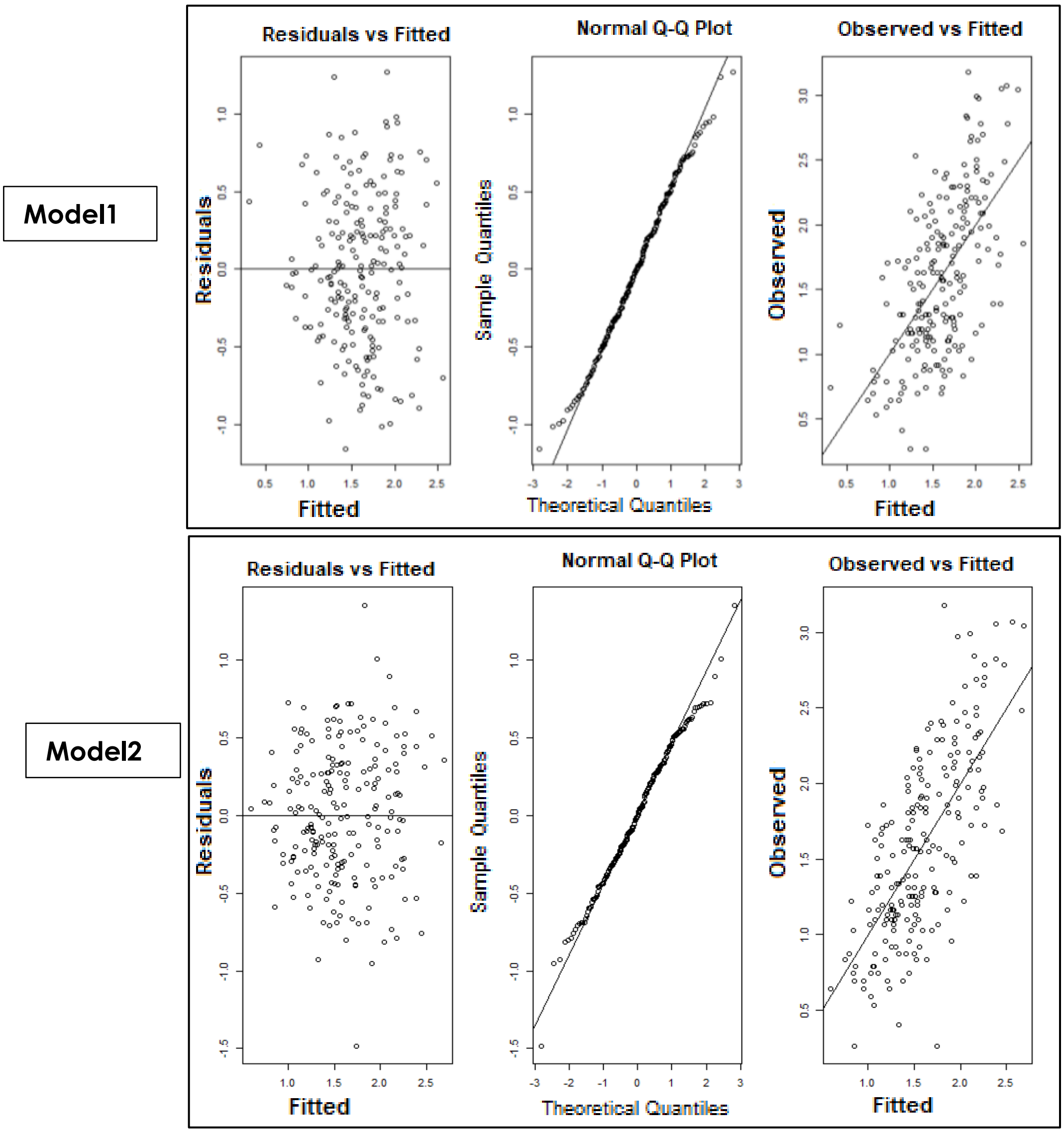
Figure 3 shows that both the left and right tail of confidence intervals, for all variables, are very wide, which could indicate a potential disturbance by extreme values in the estimation of classic econometric models, i.e. the assumption of normality of estimated residuals can be violated at its tails. Hence, it further emphasises that semiparametric models with spline functions could be more appropriate. From a statistical point of view, regarding the issue of identifying statistically significant determinants of EU regional unemployment, we can conclude that the results show a statistically significant influence of all selected factors, and the parameter estimates have the expected signs.

Also, the consideration of the spatial aspect in spatial SAR models suggests strong spatial spillover effects between regions. The statistical significance of the spatial autoregressive parameter and its high positive value (approx. 0.67) in both spatial models contribute to the confirmation of the hypothesis of spatial regional connectivity. In this context, it was necessary to calculate and verify the statistical significance of the average direct, average indirect and average total effect of all explanatory variables due to the correct interpretation of the model parameters. In addition, the assumption of non-linearity in SAR models (14) and (15), i.e., the situation that we consider a non-parametric smooth function for all covariates (except spatial lag variable), caused these effects to have a non-parametric character.

In the following figures, we show the specific results of individual econometric models, Model1 – Model5, which allows us to evaluate the stated hypothesis of the paper. Particularly, we look at whether the assumptions of each model are met, which determines the stability of the estimated parameters and their corresponding statistical tests of significance. Figures 4 – 6 display three plots: i) the residuals versus fitted values, ii) the Q-Q plot of the residuals, and iii) the observed values versus fitted values. If the first plot shows no pattern, it implies that residuals are independent and identically distributed; any other pattern could indicate a correlation in residuals or an unstable variance of residuals. The second plot assesses if the residuals come from a normal distribution, which is met if all values are close to the diagonal line. By the third plot, we are able to examine the predictive power of the model, i.e., the closer the spread of the observed versus fitted values to the diagonal line, the better the model's fit to the observed data (noting a risk of overfitting if the values are too close to the diagonal line).

Figure 4 shows a comparison between the estimate of ordinary least squares and its amended version, where we use spline functions in the estimator. Both models show distorted residuals, i.e., the first plot shows a fan-shaped pattern that indicates an unstable, nonconstant variance of the residuals. The second diagram shows that the assumption of normality is violated by extreme values, which is made clear by the deviation of the values at both ends. The third diagram shows that a model with spline functions (Model 2) performs slightly better in terms of predictive power.

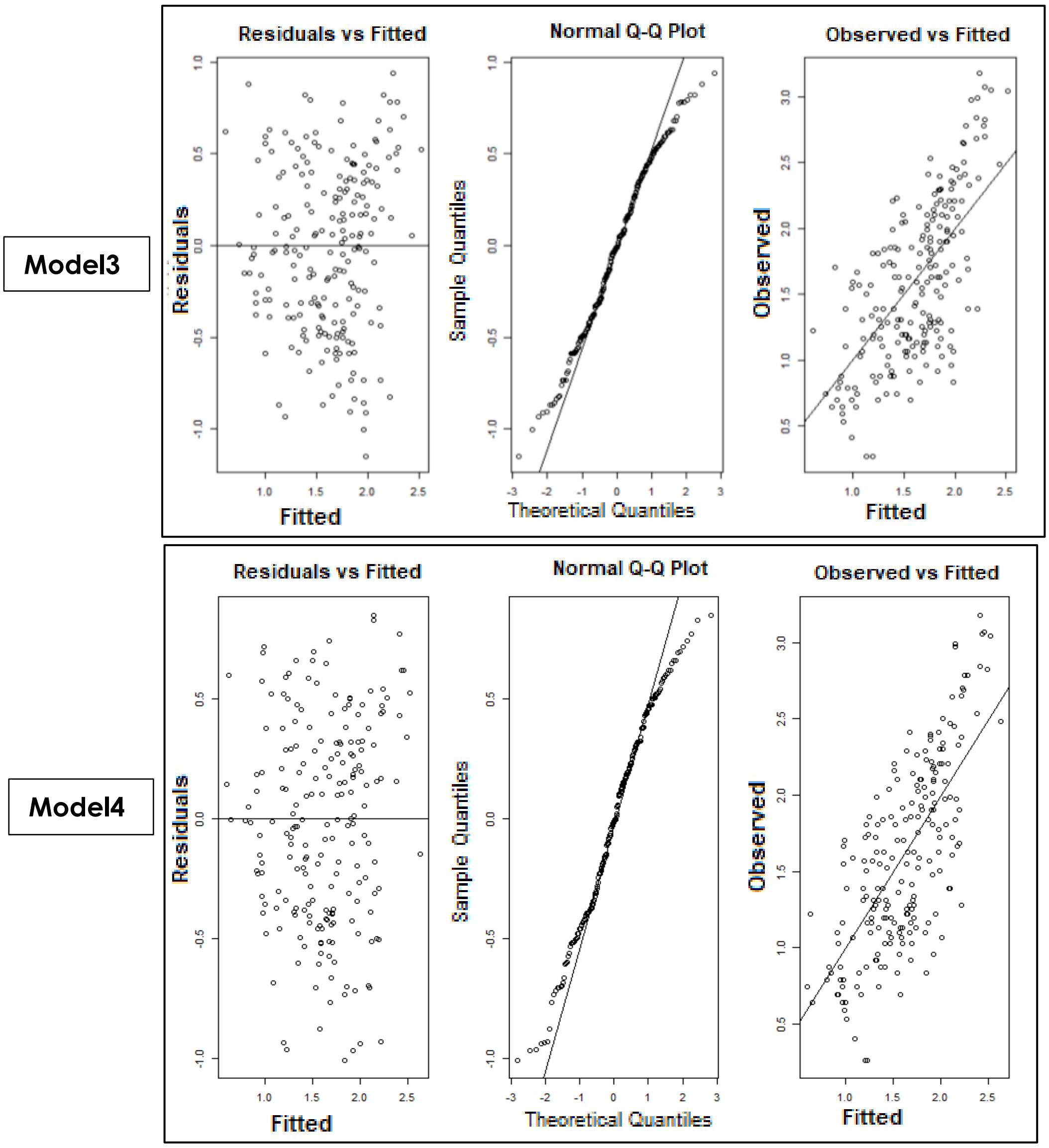
Figure 4
Residuals vs Fitted, Normal Q-Q plot and Observed vs Fitted – Model1 and Model2



Source: Authors' work.

Figure 5 compares the regression models with spatial effects that are expected to improve the model's stability since they correct the estimator for the spatial correlation across observations. Similarly, the first model is a classic spatial autoregressive (SAR) model, with the second one being its counterpart with spline functions in the estimator. Contrary to expectations, neither of the models with spatial effects (Model 3 and Model 4) significantly improves the estimation results in terms of the normality of the residuals and the predictive power of the model. The normality of residuals is still violated, as can be seen in the second plot, and the predicted power of the models, shown in the third plot (see Figure 5), is similar to regressions with no spatial effects. However, a spatial regression attains an improvement in the stability of the estimated models, i.e., the first plot of residuals versus fitted values shows a somehow random pattern, in both the classic and spline SAR models.

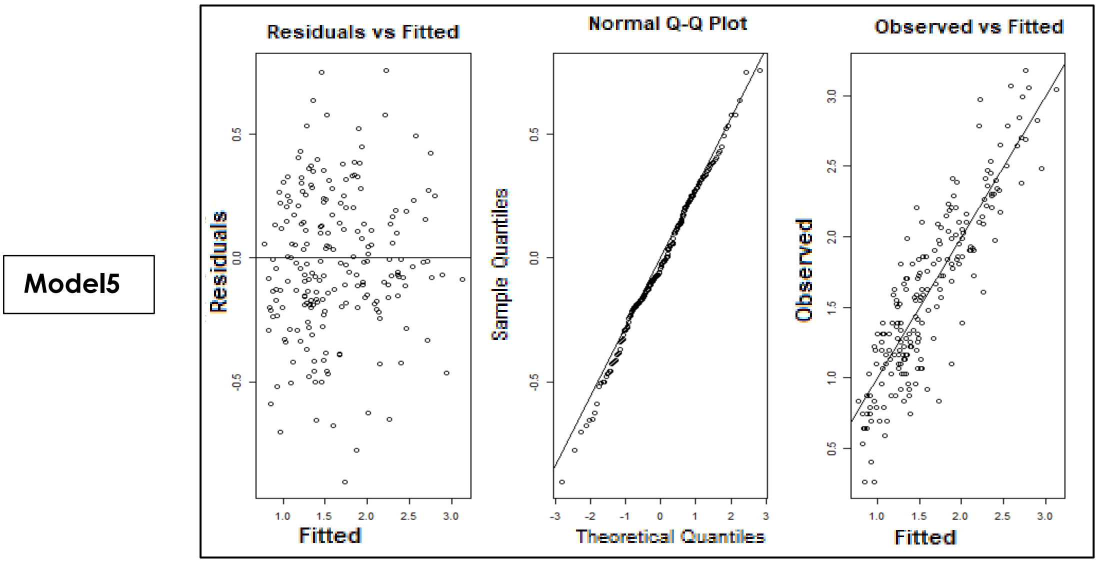
Figure 5
Residuals vs Fitted, Normal Q-Q plot and Observed vs Fitted – Model3 and Model4



Source: Authors' work.

Figure 6 displays the same matrix of the evaluation measures for the spatial regression model with spline functions, which also include a spatial trend in the estimator. We intentionally show the performance of this model on a standalone basis since it attains a significant improvement in all of the performance measures. Including the spatial trend in the model's estimator corrected the deviation from the normality at the tails of the Q-Q plot, i.e., the extreme values do not cause any disturbance, and we can safely keep them in the dataset (some analyses exclude extreme values). The other two plots also show a more accurate behaviour of the model when compared to its previous spatial counterparts.

Figure 6
Residuals vs Fitted, Normal Q-Q plot and Observed vs Fitted – Model5



Source: Authors' work.

Moreover, we further examine the analysis of variance between models by looking at different conventional statistics which are calculated as part of the models' estimation. This allows us to pick the best estimator.

Table1
Linear vs Nonlinear with splines (Model1 vs Model2)

	logLik(1)	rlogLik(2)	EDF(3)	AIC(4)
Linear	45.51	32.32	5.00	-81.02
Nonlinear with splines	49.83	45.90	15.46	68.74

Note: (1) Note: Log-Likelihood; (2) restricted Log-Likelihood; (3) Effective degrees of freedom; (4) Akaike information criterion
Source: Authors' work.

Table 2
SAR vs SAR with splines (Model3 vs Model4)

	logLik(1)	rlogLik(2)	EDF(3)	AIC(4)
SAR	109.31	94.254	6.000	-206.62
SAR with splines	111.14	104.366	13.118	-196.05

Note: (1) Note: Log-Likelihood; (2) restricted Log-Likelihood; (3) Effective degrees of freedom; (4) Akaike information criterion
Source: Authors' work.

Table 3
Nonlinear with spline vs SAR with splines (Model2 vs Model4)

	logLik(1)	rlogLik(2)	EDF(3)	AIC(4)
Nonlinear with splines	49.83	45.90	15.46	68.74
SAR with splines	111.14	104.366	13.118	-196.05

Note: (1) Note: Log-Likelihood; (2) restricted Log-Likelihood; (3) Effective degrees of freedom; (4) Akaike information criterion
Source: Authors' work.

Table 4
SAR with splines vs SAR with splines and spatial trend (Model4 vs Model5)

	logLik(1)	rlogLik(2)	EDF(3)	AIC(4)
SAR with splines	111.14	104.366	13.118	-196.05
SAR with splines and spatial trend	122.48	118.49	27.040	-190.88

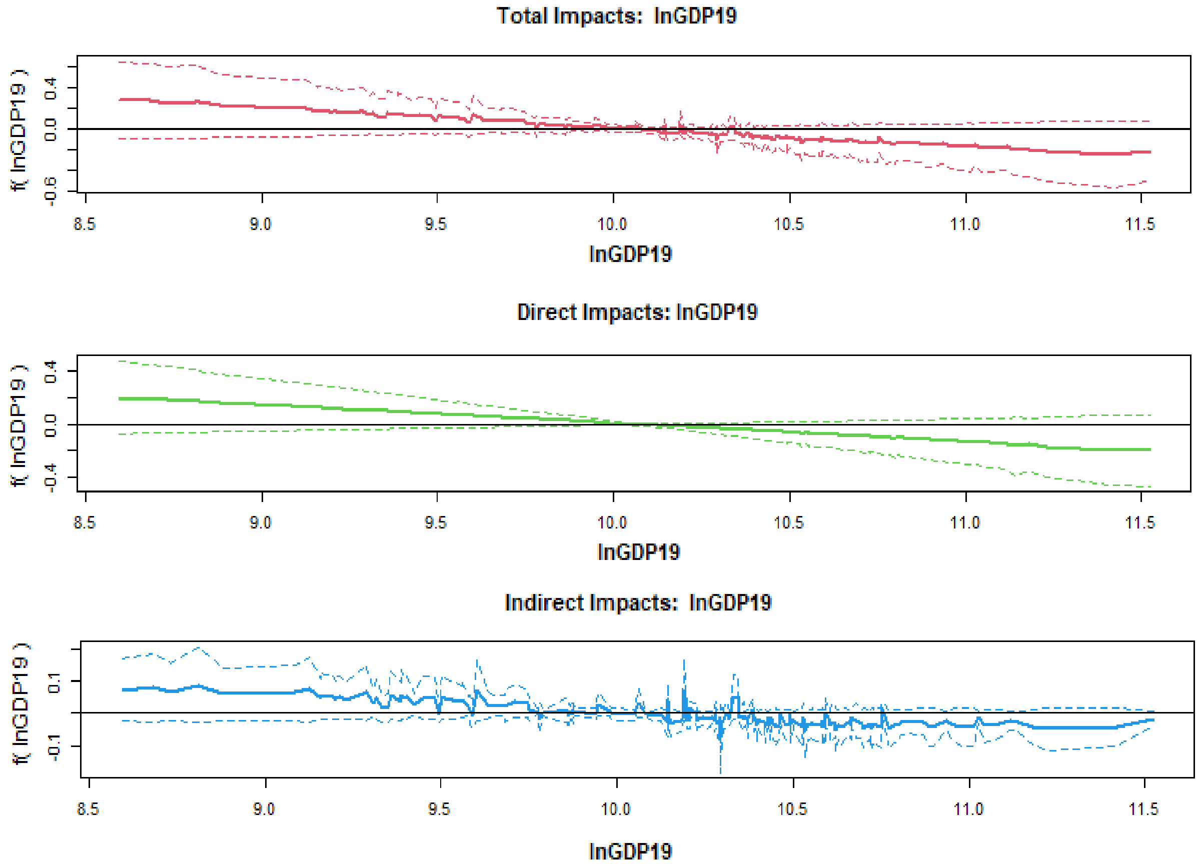
Note: (1) Note: Log-Likelihood; (2) restricted Log-Likelihood; (3) Effective degrees of freedom; (4) Akaike information criterion
Source: Authors' work.

From tables 1-4, we can re-state the conclusion that the SAR model with spline functions, which also includes the spatial trend, has the highest log-likelihood, restricted log-likelihood, and EDF numbers. The EDF measurement shows that all models are highly nonlinear. Regarding the information criteria, the AIC values are similar for the spatial models, and these values speak in their favour.

The semiparametric SAR model with a spatial trend (Model5) appears to be extremely useful for modelling spatial data with respect to nonlinearities, spatial dependence, and spatial unobserved heterogeneity when this heterogeneity is smoothly distributed over space. Figure 7 shows selected estimation results based on Model 5 as a brief preview of the results. Based on the estimation results of Model 5, we were able to calculate total, direct and indirect (or spillover) effects for all smooth (non-parametric) terms. Graphs of non-parametric covariate terms for the SAR model with spline and spatial trend (Model5) are presented in Figure 7, the results are presented only for the GDP variable.

Interpreting non-parametric impacts from a spatial semiparametric autoregressive model involves understanding the direct, indirect, and total effects of predictor variables on the response variable. The interpretation of these results compared to the results provided by the parametric SAR model is more complicated but probably provides very useful insights into the influence of predictors on the response variable. Figure 7 provides interesting information, e.g., regarding indirect effects. Indirect effects capture the impact of a predictor variable on the response variable through spatial dependencies, considering interactions with neighbouring regions. Positive indirect effects suggest that an increase in the predictor variable not only affects the response variable in the same region but also spills over to positively influence neighbouring regions. Conversely, negative indirect effects imply a negative spillover effect. In the case of the GDP variable for its different levels, we see that these impacts are different, and we notice that higher GDP values correspond to negative spillover effects. This means that a higher level of GDP in neighbouring regions contributes to reducing the level of unemployment in a particular region.

Figure 7
Plots of non-parametric direct, indirect and total impacts – GDP variable - Model5



Note: Pointwise confidence intervals in dashed lines.
Source: Authors' work.

In general, comparing direct, indirect, and total effects across different predictor variables helps prioritise their importance in influencing the response variable. Understanding the spatial dynamics can be very helpful for potential policy implications. Above all, regional policies and interventions can be targeted much more precisely.

Discussion

The empirical findings demonstrate that it is of the utmost importance to choose an appropriate theoretical framework for the econometric model, including its corresponding estimator. An incorrect model leads to weak estimated parameters, which are important when interpreted in the context of the economic impact. The model may also suffer from poor predictive power. The complex econometric models have a difficult structural form and might require more elaboration in their interpretation. However, as shown by the empirical analysis, they can lead to stable estimated parameters and improvements in the predictive power, which is crucial when using the economic interpretation of the estimated parameters to draw conclusions that can have implications for the decision of macroeconomic policies.

From the methodological perspective, we observed that the relationship between economic variables is usually subject to local nonlinearities that are not possible to be captured by the classic linear econometric models. The local nonlinear behaviour can be captured through the application of the spline function in the model's estimator. The spline functions are piece-wise polynomials that are fitted to the observed data within the specified periods – a number of spline functions determine a degree of smoothing – which directly models the local nonlinearities.

Conclusion

The main objective of the paper is to outline a theoretical framework of econometric models with different forms in their regressor function. We start with a classic linear regression which is extended to a more flexible nonlinear form by transforming its covariates into spline functions. The spline functions have the advantage that they can capture local nonlinearities that are usually present in the relationship between economic variables. In the follow-up models, we include spatial spillover effects that are common in the observations from different regions. We start with a classic spatial autoregressive (SAR) model, which is further extended to have spline functions as its covariates, with an additional version that includes a spatial trend in the estimator.

In the empirical analysis, we apply these models to the economic dataset, which contains 209 European regions, with the aim of explaining the dynamics of the unemployment rate through four key economic determining factors.

The preliminary analysis shows that all of the determining factors have a strong nonlinear relationship with the unemployment rate on a standalone basis, which indicates that a simple linear model might not be the best estimator. The findings essentially confirm the importance of the identified determinants, and, in addition, the spatial econometric model estimates also highlight the significant spatial interdependence in the context of regional unemployment in the EU. The results show that the models with spline functions are a better fit than their classic counterparts. However, the only model that corrects the instability of the estimated parameters, which is caused by the violation of normality in residuals, is the spatial regression model with spline functions that also contain a spatial trend in the regressor function.

We conclude that a more complex model can correct local nonlinearities that cause the distortion in the models' estimates. Even though these models might be more elaborate in terms of economic interpretation, they eliminate the instability in the estimated parameters that might lead to incorrect conclusions that are used for decision-making in economic policies.

Our research can be further expanded to include more variables, and it can be tested in different economic scenarios.

References

1. Anselin L., & Rey, S. J. (2014). Modern Spatial Econometrics in Practice. GeoDa Press LLC, Chicago.
2. Basile, R., Durbán, M., Mínguez, R., María Montero, J., & Mur, J. (2014). Modeling regional economic dynamics: Spatial dependence, spatial heterogeneity and nonlinearities. *Journal of Economic Dynamics and Control, 48*, 229-245. https://doi.org/10.1016/j.jedc.2014.06.011
3. Basile, R., & Mínguez, R. (2018). Advances in Spatial Econometrics: Parametric vs. Semiparametric Spatial Autoregressive Models, In: Commendatore, P., Kubin, I., Bougheas, S., Kirman, A., Kopel, M., Bischi, G. (eds) The Economy as a Complex Spatial System., pp. 81-106, Springer Proceedings in Complexity. Springer.
4. Chi, G., & Zhu, J. (2020). Spatial Regression Models for the Social Sciences. https://doi.org/10.4135/9781544302096

5. Eurostat. (2023). Regional statistics, available at https://ec.europa.eu/eurostat/web/regions/database (15 March 2023).
6. Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. D. (2021). Regression Models, Methods and Applications, Springer Berlin, Heidelberg.
7. Formánek, T. (2019). Spatial econometric analysis with applications to regional macroeconomic dynamics. Habilitation Thesis, University of Economics, Prague.
8. Geniaux, G., & Martinetti, D. (2018). A new method for dealing simultaneously with spatial autocorrelation and spatial heterogeneity in regression models. *Regional Science and Urban Economics, 72*, 74-85. https://doi.org/10.1016/j.regsciurbeco.2017.04.001
9. Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science, 1*(3). https://doi.org/10.1214/ss/1177013604
10. Hastie, T. J., Tibshirani, R. J. (1990). Generalized Additive Models, Chapman & Hall/CRC.
11. Lung-Fei, L. (2022). Spatial Econometrics: Spatial Autoregressive Models , World Scientific Publishing Company, p. 896.
12. Mínguez, R., Basile, R., & Durbán, M. (2022). An introduction to pspatreg: A new R package for semiparametric spatial autoregressive analysis. *REGION,* 9(2), R1-R15. https://doi.org/10.18335/region.v9i2.450
13. Pavlovčič-Prešeren, P., Stopar, B., & Sterle, O. (2019). Application of different radial basis function networks in the illegal waste dump-surface modelling. *Central European Journal of Operations Research, 27*(3), 783-795. https://doi.org/10.1007/s10100-018-0586-z
14. Perperoglou, A., Sauerbrei, W., Abrahamowicz, M., & Schmid, M. (2019). A review of spline function procedures in R. *BMC Medical Research Methodology, 19*(1). https://doi.org/10.1186/s12874-019-0666-3
15. Wahyuni, S. A., Ratnawati, R., Indriyani, I., & Fajri, M. (2020). Spline Regression Analysis to Modelling The Open Unemployment Rate in Sulawesi. *Natural Science: Journal of Science and Technology,* 9(2). https://doi.org/10.22487/25411969.2020.v9.i2.15202
16. Wood, S. N. (2017). Generalized Additive Models: An Introduction with R, Second Edition (2nd ed.). Chapman and Hall/CRC.
17. Wood, S. (2023). Mixed GAM Computation Vehicle with Automatic Smoothness Estimation, available at https://cran.r-project.org/web/packages/mgcv/mgcv.pdf (15 January 2024).

About the authors

Andrea Furková works as an Associate professor at the University of Economics in Bratislava. Her research interests are spatial econometrics and multi-criteria optimization. She participated as co-researcher, deputy head and head of several completed VEGA grant (the Grant Agency of the Slovak Republic) projects and COST (European Cooperation in Science and Technology) projects. The author can be contacted at **andrea.furkova@euba.sk**.

Peter Knížat is an external PhD student at the University of Economics in Bratislava. His research interests are spatial econometrics and functional data analysis. He works full-time at the Statistical Office of the Slovak Republic as a statistician, where he is responsible for proposing a statistical methodology for big data analysis. The author can be contacted at email: **peter.knizat@euba.sk**.

Sentiment and Stock Characteristics: Comprehensive Study of Individual Investor Influence on Returns, Volatility, and Trading Volumes

Aleš Kresta, Jialei Xiong, Bahate Maidiya
VSB – Technical University of Ostrava, Faculty of Economics, Czech Republic

Abstract

Background: Traditional asset pricing models face challenges from financial anomalies, prompting exploration through behavioural finance theory. This study analyses the nuanced relationship between individual investor sentiment and key stock market variables. **Objectives:** To assess the impact of individual investor sentiment on stock returns, volatilities, and trading volumes using the American Association of Individual Investors (AAII) sentiment index. **Methods/Approach:** Using regression models, we examine the relationship between individual investor sentiment and various stock characteristics across 480 components of the Standard & Poor's 500 index. **Results:** We find a positive relationship between the AAII sentiment index and stock returns and a negative relationship with volatility and trading volume. **Conclusions:** Our study contributes to understanding the intricate role of individual investor sentiment in financial markets.

Keywords: investor sentiment; stock characteristics; behavioural finance; AAII sentiment index

JEL classification: G12, G14, G41
Paper type: Research article

Received: Jan 23, 2024
Accepted: Jul 26, 2024

Acknowledgements: The authors express gratitude for the financial support to the Czech Science Foundation (GACR) under project no. 23-06280S, VSB – Technical University of Ostrava under project no. SP2024/047, and the European Union under the REFRESH Research Excellence for Region Sustainability and High Tech Industries project no. CZ.10.03.01/00/22_003/0000048.

Citation: Kresta, A., Xiong, J., & Maidiya, B. (2024). Sentiment and Stock Characteristics: Comprehensive Study of Individual Investor Influence on Returns, Volatility, and Trading Volumes. *Business Systems Research, 15*(2), 67-82.
DOI: doi.org/10.2478/bsrj-2024-0018

Introduction

For many years, traditional asset pricing models have dominated the assessment of risk-return trade-offs. However, the discovery of financial anomalies suggests that the efficient market hypothesis (Fama, 1970) may be challenged from the perspective of behavioural finance theory. Since the efficient market hypothesis (EMH) does not take into account the presence of investors' idiosyncratic behaviour (Haritha & Rishad, 2020), relying solely on EMH for asset pricing may lead to distortions. Investor sentiment, on the other hand, is popular for the empirical support it provides to asset pricing, emphasizing the impact of human biases on market behaviour. Because traditional theories do not account for the impact of abnormal investor behaviour on market outcomes, behavioural finance incorporates psychological perspectives into the description of financial markets so that we can gain a better understanding of why markets may deviate from the predictions of traditional theories such as EMH.

Baker and Wurgler (2007, p. 129) broadly defined investor sentiment as 'a belief about future cash flows and investment risks that is not justified by the facts at hand.' Both researchers and practitioners are interested in measuring market sentiment, reflecting the overall sentiment of market participants or their subgroups. González-Sánchez and Morales de Vega (2021) identified three main approaches to constructing investor sentiment indices: aggregation of market variables, investor surveys, or utilizing information from the media. Each of them has its own advantages and disadvantages. The potential drawback of constructing a sentiment index through the aggregation of market variables is that it may include unrelated information. Investor surveys, while widely used, suffer from low observation rates, usually monthly or with a lower frequency, and reliability issues when nonresponse rates are high (Sun et al., 2016). The third, rapidly evolving approach involves explaining the return on assets through textual analysis of news, but there is no clear evidence of its explanatory capacity (González-Sánchez & Morales de Vega, 2021).

In this paper, we focus on the second approach – measuring investor sentiment through investor surveys. Unlike other approaches, investor surveys provide a direct measurement of investor sentiment, as they involve directly asking and observing the sentiment among investors. Notable indexes for measuring investor sentiment in the US market include the monthly University of Michigan Consumer Sentiment Index, the weekly American Association of Individual Investors (AAII) sentiment survey, and the daily Investor Intelligence and Daily Sentiment Index. We concentrate our research on the AAII sentiment survey, which has a high number of survey participants and a long history of data since its inception in 1987.

The objective of this study is to explore the relationship between individual investor sentiments and characteristics of the stock market, such as stock returns, volatility, and trading volumes. The research question is whether there is any relationship between the AAII sentiment survey index and the characteristics of the stocks in the periods following the publication of the sentiment index data. According to the efficient market hypothesis, all relevant information should already be priced in, and the sentiment should have no predictive power for future returns, which is our null hypothesis.

The research addressed by this study also examines what should be used for prediction – whether the absolute value of the sentiment index or its change from the previous value. We hypothesize that changes in market sentiment are better predictors of future characteristics. For example, if sentiment improves from bearish to neutral, individual investors might start buying stocks and increase their bid and ask prices. On the contrary, if sentiment worsens from bullish to neutral, individual investors could start selling stocks. In both cases, the sentiment value is the same (neutral), but the actions taken by individual investors are different. Therefore, changes in the sentiment index are likely more significant than absolute values.

Our results have important implications for investors. We find a positive relationship between sentiment and future returns and a negative relationship with future volatility, suggesting that sentiment could be a useful indicator in developing investment or trading strategies. The findings contribute to understanding the role of individual investor sentiment in financial markets and its implications for investment strategies. However, it is important to approach these results with caution. Although our findings indicate a relationship between sentiment and stock returns, the sentiment variable used in our study does not capture the full spectrum of influencing factors. Therefore, relying solely on sentiment indicators for investment decisions may not consistently yield high returns, and investors should consider sentiment as one of many factors in their decision-making process.

Our study differs from previous studies, which primarily concentrated on a single time series, usually the market index, see (Fisher & Statman, 2006; Y.-H. Wang et al., 2006; Kurov, 2008; Jacobs, 2015; Białkowski et al., 2023). We focus on a more robust dataset comprising the component stocks of the market index. Specifically, we use components of the Standard & Poor's 500 index. This approach allows for a comprehensive analysis that considers the dynamics and interactions within a broader range of securities, providing more robust results.

The remainder of the paper is structured as follows. In the next section, we provide a short review of the literature. Then we introduce the data and methods applied. In the next sections, we present the results and their discussion. The last section presents the conclusion of the paper.

Literature Review

The selection of stock returns, volatilities, and trading volumes as dependent variables in this study is based on their fundamental importance in financial market analysis. Stock returns are a primary measure of a stock's performance and are crucial for investors, as they directly relate to the gains or losses experienced. Volatility, on the other hand, serves as a key proxy of risk, reflecting the degree of uncertainty, with higher volatility indicating greater risk. Trading volumes provide an important measure of market activity and liquidity, and higher trading volumes typically indicate greater market interest and ease of transacting without significantly affecting stock prices. Together, these variables are crucial factors for investors, as they directly impact investment decisions (Hawaldar & Rahiman, 2019; Veld & Veld-Merkoulova, 2008).

These stock characteristics are also interrelated. A fundamental principle in finance is that investors require, or expect, higher returns for undertaking higher risks (represented by volatility). Research has also examined the relationship between trading volume and returns. Chen et al. (2001) and Naik and Sethy (2022) found a positive correlation between trading volumes and stock price changes, with trading volume contributing to the return process. Naik and Sethy (2022) also highlighted the asymmetric effect of stock returns on trading volume and the positive volume-volatility relationship.

However, stock returns can also be explained by other factors than trading volume and volatility. Traditional pricing models, such as the Fama-French five-factor model (Fama & French, 2015, 2017), explain the stock returns based on factors related to market return, company size, book-to-market ratio, profitability, and investment style. Macroeconomic indicators can also serve as predictors; for instance, Hjalmarsson (2010) identified the short interest rate and term spread as robust predictors in developed markets. In addition to these fundamental and macroeconomic factors, technical analysis provides another approach to understanding and predicting stock returns. Technical analysts believe that all relevant information is already reflected in stock prices and that price movements follow certain patterns that can be exploited. However, there is a broad academic debate on the applicability of technical analysis, as reviewed by Park and Irwin (2007).

Traditional asset pricing models face growing challenges in their explanatory power and research in behaviour finance theory has further demonstrated the role of investor sentiment in driving the stock markets. Johnson and Tversky (1983) suggested that people make risky decisions based on their sentiment state. Compared with traditional asset pricing models, one of the main arguments of behaviour finance is that imperfectly rational traders (known as noise traders) generate deviations from fundamental values (Uygur & Taş, 2014) and that these deviations can significantly affect investor behaviour and market prices (Daniel et al., 2002).

Focusing on investor sentiment is particularly important because this factor can lead to market anomalies that traditional models fail to explain. In recent years, research has increasingly explored the role of sentiment in returns, volatility, and trading volumes. Although the empirical results of Lee et al. (2002) suggest that sentiment is priced in systematic risk, the excess returns are still positively correlated with changes in sentiment. Additionally, the extent of bullish (bearish) sentiment changes leads to downward (upward) corrections in volatility and higher (lower) future excess returns. Uygur and Taş (2012) demonstrated that investor sentiment has a significant positive effect on the conditional volatility of the stock market during periods of high sentiment, whereas investor sentiment has a negative effect during periods of low sentiment.

Audrino et al. (2020) investigated the effect of sentiment and attention indicators on daily stock market volatility and showed that sentiment and attention variables have significant predictive power for future volatility and that the addition of the sentiment variable leads to a further decrease in the mean square prediction error, especially on days with high volatility. By distilling the sentiment of the news text, Zhang et al. (2016) found that changes in sentiment, especially those with negative views, affect volatility and volume.

Studies that investigate trading volume also show a relationship between investor sentiment and trading volume. From the results of So and Lei (2011), we can understand that the increase in the volatility index (VIX) is associated with an increase in trading volume, especially during periods when the VIX is high. This suggests that the higher the VIX level, the greater the change in trading volume. This is further confirmed by Lai et al. (2014), who found a positive correlation between investor sentiment and abnormal trading volumes. However, Kim and Ryu (2021) added a nuanced perspective by noting that the term structure of the impact of sentiment on trading volume is downward sloping, suggesting that instances of sentiment-induced trading anomalies are relatively short-living.

However, empirical studies that use the AAII sentiment survey as a sentiment measure are scarce. Previous studies involving data from the AAII sentiment survey have shown that sentiment-driven investors often trade based on data from the AAII sentiment survey (Chau et al., 2016). The AAII sentiment index not only affects the stock price (Bouteska, 2019) but also significantly affects both the stock return and volatility (Sayim et al., 2013). Additionally, the AAII sentiment index plays a crucial role in the performance of initial public offerings (Ibrahim & Benli, 2022).

Methodology

The sentiment index used in this study is obtained from the Investors Intelligence Survey (American Association of Individual Investors, 2024), which is conducted by the American Association of Individual Investors (AAII). The sentiment survey collects data from individual investors on their current market outlooks and investment decisions. Weekly, participants in the sentiment survey receive an email with a straightforward question: 'Do *you feel the direction of the stock market over the next six months will be up (bullish), no change (neutral), or down (bearish)?*' Participants can only submit one vote and their responses are used to calculate the indices, representing the percentages of bullish, bearish, and neutral market outlooks. By bullish outlook, we refer to the expectation of the participants that the stock market will grow in value. On the other hand, the bearish outlook represents the opinion of a future decline in the value of the stock market. Neutral perspectives are neither bullish nor bearish.

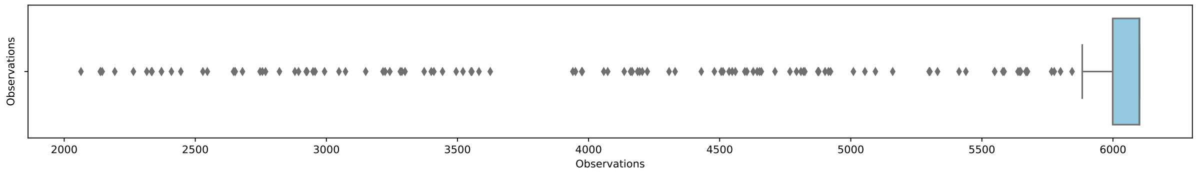
The sentiment index (AAII sentiment index) is then calculated as the spread between the bullish and bearish percentages of votes, ranging from –100% to 100%. Additionally, in our research, we experiment with an alternative construction of the sentiment index (the ratio of bullish to bearish) and their week-by-week differences or differences between the index and its moving average, as we believe that the change in sentiment can be more important than its absolute value.

It is crucial to note that this index is based on the opinions and investment decisions of individual investors and may not always align with broader market trends or sentiments. However, the survey offers valuable information on individual investors' perspectives and can help to understand the general outlook of the market.

We assume that the individual investors are usually in a long position (Visaltanachoti et al., 2007), that is, they hold the stocks. As investor sentiment measures the beliefs of the market participant, we believe that the same market participants adjust their supply and demand for the stock accordingly. When they expect the market to soar, they increase the prices for which they are willing to buy and sell (ask and bid prices) as they consider the stocks to be undervalued at the current prices. They also increase the bid price, i.e., the price for which they are willing to buy because they speculate on the price increase. However, since there is positive market sentiment, fewer market participants are willing to sell, which decreases the trading volume and volatility. At the same time, we expect that as the sentiment in the market is negative or worsens, individual investors want to sell the stocks quickly, which both decreases the prices and increases the volatility and the trading volumes at the same time. Thus, we expect a positive relationship between sentiment and future returns and a negative relationship between sentiment and future return volatility and trading volume.

Furthermore, for the stocks, our dataset comprises the components of the S&P 500 index (Wikipedia, 2024), and we obtained the daily adjusted close prices and volumes from the Yahoo Finance website (Yahoo, 2024). All data were collected for the period from January 1, 2000, to December 31, 2023. We chose this period as the trade-off between the length of the data, as the longer length provides more robust results because all market phases (bullish, neutral, bearish) are present in the data, and data availability because the composition of the index is changing during the time and some companies become no longer publicly trading and the data are not available. At the same time, since the newcomers to the index do have shorter price histories, we excluded stocks with time series lengths of less than 2,000 observations, resulting in a database reduction from 502 stocks to 480 stocks. From these, for 357 stocks, we used the full history of 6,101 daily observations that covered the last 24 years, while for the rest the time series length was shorter; see Figure 1.

Figure 1
Boxplot of the Lengths of Time Series Applied



Source: Authors' work

As the sentiment data are weekly, we need to recalculate the daily data to weekly. The sentiment data are published each Thursday, and we assume that these data can be utilized to forecast the next week's characteristics; thus, we align the sentiment indexes published on Thursday with the subsequent Thursday-to-Thursday characteristics.

We employ several regression models, each estimated via the Ordinary Least Squares (OLS) method, incorporating a Newey-West heteroskedasticity and autocorrelation consistent covariance matrix (Newey & West, 1987). The general regression equation is expressed as follows,

$$E(y_{i,t}) = \alpha_i + \beta_i \cdot x_t, \tag{1}$$

were, *y* represents the chosen dependent variable, *i* is the index that specifies the stock, *t* is the time, α_i and β_i are regression coefficients for stock *i*, intercept and slope respectively, and x_t is the sentiment at time *t*.

In our research, we examine four dependent variables (*Y*): return, risk premium, volatility, and trading volume. We also assume different specifications of the independent variable (*X*): the value of the sentiment index, week-to-week differences, and the differences between the index value and its five-week moving averages for the sentiment index calculated as both the spread and the ratio of bullish and bearish percentages of votes. For all these sentiment index series, the null hypothesis of a unit root is rejected by the Augmented Dickey-Fuller unit root test at 0.01 significance level. In each model, sentiment indices are used as an independent variable, while stock returns, trading volumes, and volatilities are used as a dependent variable.

The first characteristic analysed is the one-week return calculated as a percentage change of the adjusted close prices *p* from Thursday to next Thursday,

$$r_{i,t} = \frac{p_{i,t}}{p_{i,t-1}} - 1. \tag{2}$$

According to the CAPM model, the returns can be divided into the risk-free rate and the risk premium. Thus, we also focus on one-week risk premiums, calculated as one-week returns minus the risk-free rates obtained from French (2024). The third dependent variable under consideration is the volatility of the returns. As it is not directly observable in the market, we estimate it ex-post from the returns using the generalized autoregressive conditional heteroskedasticity (GARCH) model (Bollerslev, 1986), specifically, we assume GARCH(1,1) specification:

$$r_{i,t} = \mu_i + \sigma_{i,t} \cdot \epsilon_{i,t}, \tag{3}$$
$$\sigma_{i,t}^2 = \omega_i + a_i \cdot \sigma_{i,t-1}^2 + b_i \cdot \epsilon_{t-j}^2, \tag{4}$$

where μ_i is the mean return of the *i*th stock, $\sigma_{i.t}$ is the standard deviation (volatility) for the *i*th stock modelled by the GARCH model and $\epsilon_{i,t}$ is a white noise. Parameters ω_i, a_i and b_i need to be estimated. Positive variance is ensured if $\omega_i > 0$, $a_i \geq 0$, and $b_i \geq 0$, and the model is stationary if $a_i + b_i < 1$. The fourth dependent variable considered is the volume in US dollars traded in one week from Thursday to Thursday.

In all characteristics, we align the newly announced values of the sentiment indices at time *t*, with the characteristics of the following week, that is, return, premium, volatility, and trading volume in the period from *t* to *t*+1.

Results

In this section, we present the results of the estimated regression models (1), which are carried out for the 480 component stocks of the S&P 500 index. The reported results include the number of stocks for which the estimated parameters are considered statistically significant at significance levels of 10%, 5% and 1% by means of the t-test and box plots of the parameter values.

First, we focus on returns, where the dependent variable in Equation (1) is the weekly return. Table 1 illustrates the number of stocks for which the estimated parameters are statistically significant. As can be seen, the sentiment index calculated as the spread is a better predictor than the index calculated as the ratio. In fact, when using the spread, the slope is statistically significant in the case of 242 stocks out of 480, that is, for around half of the stocks, compared to only 107 in the case of ratio. When we transform the independent variable to differences, either week-to-week or value-to-average, the number of statistically significant parameters increases. We can observe that for 390 stocks from 480, the difference between the index calculated as spread and its previous month's average is statistically significant in predicting the future one-week return.

Figure 2 shows the box plots of the parameters for all stocks. As can be seen, for all dependent variables except the ratio, both the intercept and the slope are positive in most of the stocks. We can conclude that there is a positive relationship between sentiment and subsequent one-week returns. Therefore, investors can use sentiment as a predictor of the return next week.

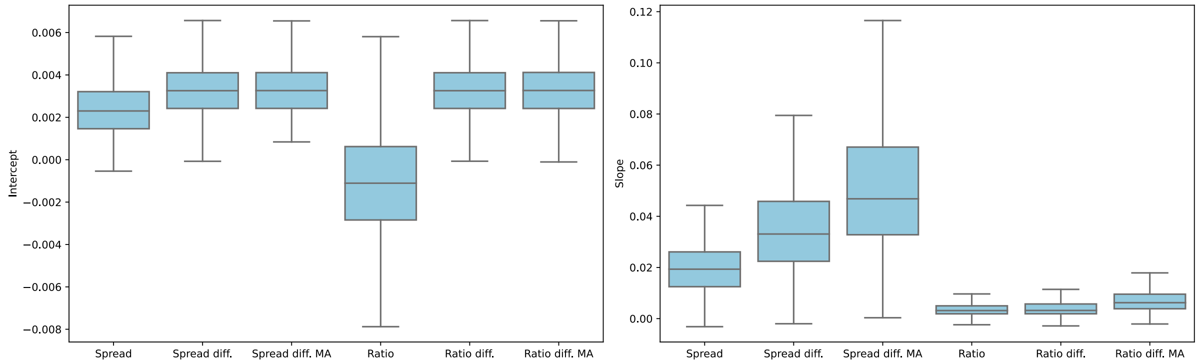
Considering the best model (difference of spread to its MA), the median values can be interpreted as follows. The expected value of the weekly return from Thursday to Thursday is 0.33% (intercept) plus 4% (slope) for every 100 bps of the difference between the sentiment index value and its average in the previous four weeks. Alternatively, we can annualize the returns for better comparability. Then, the expected value of the next-week return is 16.96% p.a. plus 2.44% p.a. for every 1 bps of the difference between the sentiment value and its average in the previous four weeks.

Table 1
Quantities of Statistically Significant Parameters of Regression of Return on Sentiment

Independent variable	10% significance level		5% significance level		1% significance level	
	Intercept	Slope	Intercept	Slope	Intercept	Slope
Spread	254	375	175	339	63	242
Spread diff.	404	443	343	420	202	349
Spread diff. MA	406	447	344	439	202	390
Ratio	48	296	27	223	3	107
Ratio diff.	403	289	342	225	201	129
Ratio diff. MA	403	369	342	333	202	225

Source: Authors' work

Figure 2
Parameter Values of Return Regressions



Note: Outliers positioned significantly far from the median have been excluded for the sake of clarity. Source: Authors' work

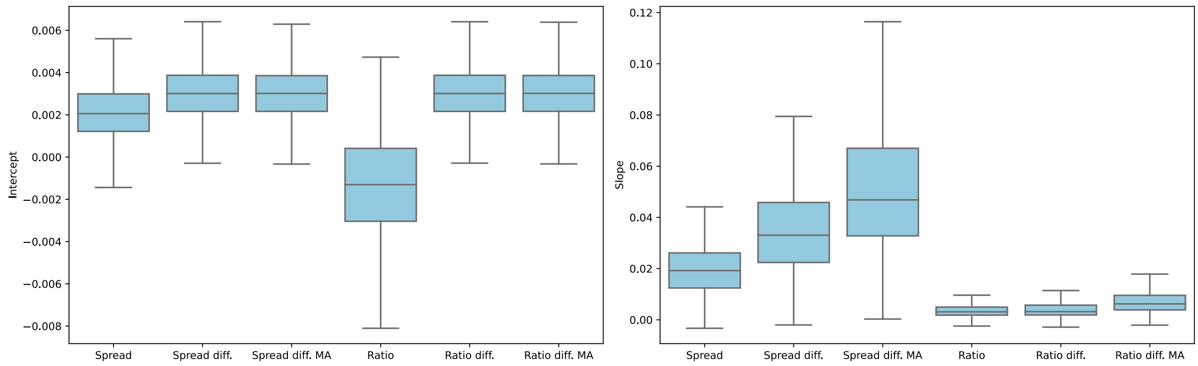
The same conclusions can be drawn when analysing the premiums, that is when we subtract the risk-free rate from the returns; see Table 2 and Figure 3. The median value of the intercept is 15.66% p.a. and the median value of the slope is 2.44% p.a. for every 1 bps of the difference between the sentiment value and its average in the previous four weeks. We can see that the slope has not changed while the intercept has changed by 1.3 % p.a., roughly equalling the risk-free return during the analysed period.

Table 2
Quantities of Statistically Significant Parameters of Regression of Premium on Sentiment

Independent variable	10% significance level		5% significance level		1% significance level	
	Intercept	Slope	Intercept	Slope	Intercept	Slope
Spread	206	373	142	338	43	238
Spread diff.	371	443	300	420	160	348
Spread diff. MA	374	447	302	439	161	390
Ratio	54	286	27	216	6	103
Ratio diff.	369	289	300	225	160	129
Ratio diff. MA	370	369	300	333	160	225

Source: Authors' work

Figure 3
Parameter Values of Premium Regressions



Note: Outliers positioned significantly far from the median have been excluded for the sake of clarity. Source: Authors' work

Third, we turn our attention to the association between sentiment and volatility. Given that volatility is not directly observable in the market, we opt for the GARCH(1,1) model for ex-post estimation. For each stock, we apply the GARCH(1,1) model; see equations (3) and (4), extracting weekly estimated volatilities, which are then utilized as the dependent variable in regression (1).

The results are presented in Table 3 and Figure 4. The intercept consistently exhibits statistical significance and a positive value for all stocks in each model. On the other hand, the slope is statistically significant only for the spread and ratio, i.e., the volatility depends on the value of the sentiment index and not on its change. The sentiment indexes calculated as both the spread and the ratio show similar results. There exists a statistically significant linear relationship between the sentiment index value and the next week's return volatility for 377 (347, respectively) stocks out of 480. Even when factoring in the potential for Type I errors (4.8 at 1%), we confirm a significant relationship between sentiment and volatilities. The predominant direction of this relationship is negative, as seen in Figure 4, which is observed in more than 75% of stocks, indicating that a higher sentiment index value is associated with lower volatility. However, a smaller fraction of stocks (less than 25%) exhibits a positive relationship, where higher sentiment aligns with higher volatility.

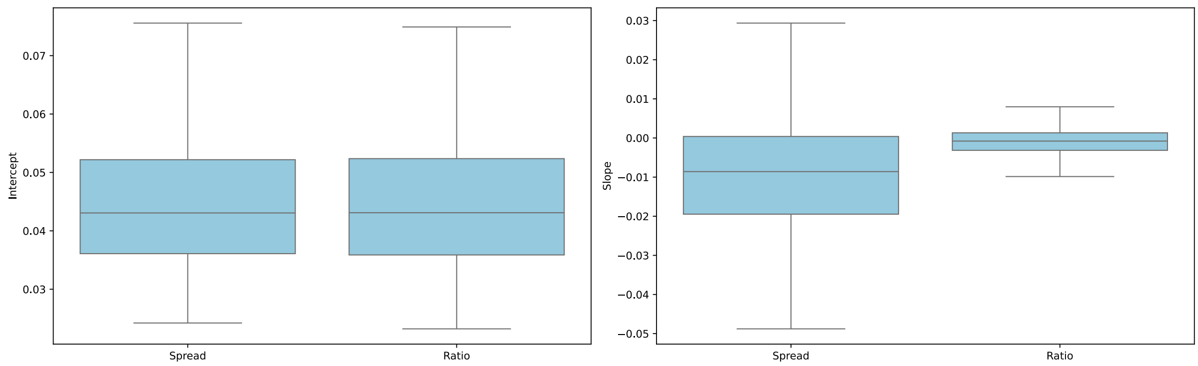
Concerning the median values, we can interpret the intercept and slope for the spread sentiment index as follows. The expected value of the next week's standard deviation of the return is 4.3% minus 0.86% for every 100 bps in the index value or, when annualized, 31% p.a. minus 0.62% p.a. for every 10 bps in the sentiment index value.

Table 3
Quantities of Statistically Significant Parameters of Regression of Volatility on Sentiment

Independent variable	10% significance level		5% significance level		1% significance level	
	Intercept	Slope	Intercept	Slope	Intercept	Slope
Spread	480	413	480	404	480	377
Spread diff.	480	4	480	1	480	1
Spread diff. MA	480	63	480	27	480	3
Ratio	480	412	480	387	480	347
Ratio diff.	480	1	480	0	480	1
Ratio diff. MA	480	10	480	2	480	0

Source: Authors' work

Figure 4
Parameter Values of Volatility Regressions



Note: Outliers positioned significantly far from the median have been excluded for the sake of clarity. Source: Authors' work

Fourth, our focus shifts to the trading volume, the dependent variable in regression equation (1) being the trading volume in one week in US dollars. The results, illustrated in Table 4 and Figure 5, closely resemble those of the volatility regression, that is, the intercept consistently proves statistically significant and positive for all regressions, while the slope generally exhibits statistical significance and negativity only for the values of the sentiment index and not their differences. Specifically, we affirm the statistically significant relationship between the AAII sentiment index and trading volume in 378 (383, respectively) stocks out of the 480 considered at a 1% significance level. Even when factoring in the potential for Type I errors (4.8 at 1%), we confirm a significant relationship between sentiment and volatilities.

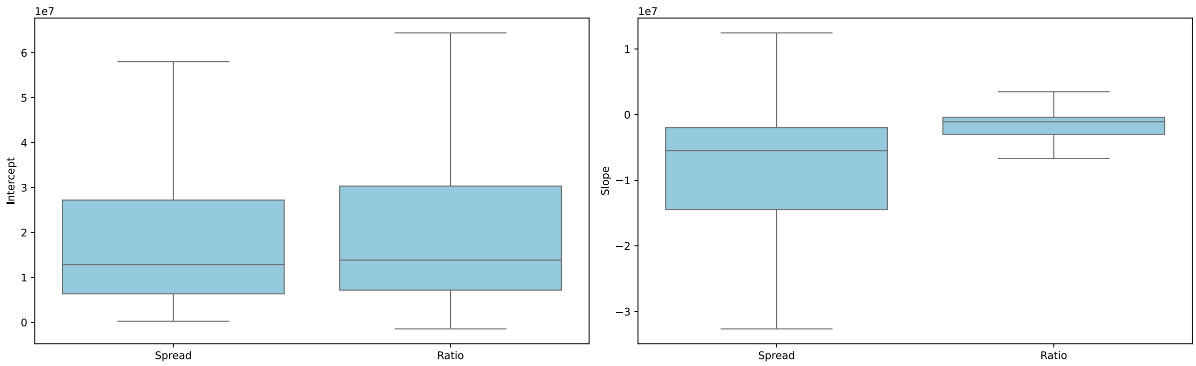
Concerning the median values, we can interpret the intercept and slope for the spread sentiment index as follows: the expected value of the next week's trading volume is $12,866,011 minus $55,129 for every 1 bps in the sentiment index value. In the sentiment case of the ratio index, the expected value of the next week's trading volume is $13,852,933 minus $11,002 for every 1 bps in the sentiment index value.

Table 4
Quantities of Statistically Significant Parameters of Regression of Volume on Sentiment

Independent variable	10% significance level		5% significance level		1% significance level	
	Intercept	Slope	Intercept	Slope	Intercept	Slope
Spread	480	419	480	412	480	**378**
Spread diff.	480	23	480	14	480	**2**
Spread diff. MA	480	117	480	65	480	**20**
Ratio	479	420	479	410	479	**383**
Ratio diff.	480	8	480	3	480	**1**
Ratio diff. MA	**480**	**49**	**480**	**32**	**480**	**9**

Source: Authors' work

Figure 5
Parameter Values of Volume Regressions



Note: Outliers positioned significantly far from the median have been excluded for the sake of clarity. Source: Authors' work

Discussion

Our investigation uncovers evidence of a discernible linear relationship between individual investors' sentiments and returns, volatilities, and trading volumes. Specifically, our findings indicate a general positive relationship between sentiment changes and stock returns. Furthermore, we found general negative associations between sentiment and volatility and sentiment and trading volume. However, these relationships are not valid for all analysed stocks, as approximately 25% of stocks show the opposite relationship for volatility and trading volume and are not always statistically significant. In general, statistically significant relationships are discovered for 390 (returns), 377 (volatility), and 378 (trading volume) stocks out of 480 at a 1% significance level. It is important to account for Type I errors, which would cause 4.8 false relationships out of 480 at a 1% significance level. However, the quantities of statistically significant relationships are still relatively high.

When comparing independent variables, we confirm that returns and premiums are influenced by the change in the sentiment index more than by the value of the index itself. However, volatility and trading volume depend on the sentiment index's value and are independent of changes in its values. We also found that the original construction of the sentiment index as the spread between the bullish and bearish percentages of votes performs better than its ratio as an alternative specification.

Our findings in terms of the influence of sentiment on returns contradict the existing body of research, which emphasises a predominantly negative relationship between returns and sentiment indices. Works such as (Baker & Wurgler, 2006; Baker et al., 2012; Jiang et al., 2019; Białkowski et al., 2023; Aissia & Neffati, 2023) have consistently reported this trend. Additionally, a meta-analysis conducted by (Gric et al., 2023) supports this observation, suggesting that the true effect is negative, although in some specifications it is not significant, and in the majority of specifications, researchers tend to report this effect as being much stronger than it actually is.

The explanation for our findings can be found in (Wang et al., 2022), who found that the relationship differs based on the market regime. In bull regimes, optimistic (pessimistic) shifts in investor sentiment increase (decrease) stock returns, whereas in bear regimes, optimistic (pessimistic) shifts decrease (increase) stock returns. Our period under investigation, i.e., the years 2000-2023, although containing recent crises and bear periods such as a burst of a dot-com bubble, global financial crisis, COVID-19 pandemic, and the Russian invasion of Ukraine, also contains a strong bull market in the period from 2009 to 2023. The positive relationship can, therefore, be caused by this long bull market period present in our data. In the study of Haritha and Rishad (2020), it was also discovered that when investors' sentiment is positive, their return expectations tend to be positive as well. This positive sentiment may prompt investors to capitalise on the situation for speculative activities, encouraging increased investment, which increases the prices.

Our study confirms the negative relationship between sentiment and volatility. The sentiment level exhibits a significant relationship with volatility, revealing a pronounced negative correlation trend. More than 75% of the stocks in our study demonstrated a connection between higher sentiment and lower volatility. The minority of stocks (less than 25%) that show a positive relationship between sentiment and volatility can be attributed to specific market conditions or idiosyncratic factors that affect these particular securities. Interestingly, our findings contrast with previous research by Brown (1999), who reported a positive correlation between sentiment and volatility. However, there also exists a related body of literature that yields findings consistent with ours. For instance, Sayim et al. (2013) observed that an unforeseen rise in the emotional rationality component among individual investors has a notably adverse effect on industry volatility, particularly in the US automotive and financial sectors.

The realm of sentiment's impact on trading volume remains largely unexplored in the existing literature, presenting a gap in the understanding of market dynamics. In particular, most studies investigating trading volume have traditionally focused on proxies for investor attention, exemplified by the use of indicators like Google searches

(Joseph et al., 2011). Our findings reveal a compelling negative relationship between individual investor sentiment and trading volume. This suggests a distinctive pattern of investor response, characterised by increased reactions to negative sentiment. As the market sentiment is positive, fewer individual investors are willing to sell stocks, reducing the trading volume. On the other hand, when the market sentiment is negative, individual investors sell the stocks, which increases the trading volumes.

However, it is essential to acknowledge several limitations of our study. First, we assume the one-directional causality is from the sentiment to stock characteristics; however, the causality can be bidirectional. Second, we use the sentiment index as the only independent variable, neglecting other factors that can influence stock characteristics. There are certainly other independent variables that can be used to predict future stock characteristics. Furthermore, our study focuses on a single period, from 2000 to 2023. The choice of a different period could influence the results. Future research efforts could address these limitations, providing a more comprehensive understanding of the intricate dynamics at play in financial markets.

Conclusion

In summary, our study provides comprehensive information on the intricate relationship between individual investor sentiment and various characteristics of the stock market, including returns, volatility, and trading volumes. Through rigorous analysis, we have discovered several key findings that contribute to our understanding of market dynamics.

First, our results reveal a positive relationship between the change in individual investor sentiment and future stock returns. This suggests that sentiment can serve as a valuable predictor of market performance, with higher changes in sentiment levels generally being associated with higher future returns. This finding underscores the importance of considering investor sentiment in investment decision-making processes.

Second, we find a negative association between sentiment and both the volatility and the trading volume. Specifically, higher levels of sentiment tend to coincide with lower levels of volatility and trading volume. For institutional and individual investors, understanding the impact of sentiment on trading volume can provide more appropriate strategies for trade execution, and understanding the relationship between sentiment and market volatility can be helpful in developing risk management tools and trading strategies that are more profitable and less risky.

However, it is important to approach these results with caution. Although our findings indicate a relationship between sentiment and stock returns, the sentiment variable used in our study does not capture the full spectrum of influencing factors. Thus, relying solely on sentiment indicators for investment decisions may not consistently yield an improved performance and investors should consider sentiment as one of many factors in their decision-making process.

References

1. Aissia, D. B., & Neffati, N. (2023). Investor sentiment metrics and stock market returns: A study of the causality relationship using VAR models. *American Journal of Finance and Accounting*, *7*(2), 90–124. https://doi.org/10.1504/AJFA.2023.134710
2. American Association of Individual Investors. (2024). *Sentiment Survey Past Results | AAII*. Sentiment Survey Historical Data. https://www.aaii.com/sentimentsurvey/sent_results

3. Audrino, F., Sigrist, F., & Ballinari, D. (2020). The impact of sentiment and attention measures on stock market volatility. *International Journal of Forecasting*, *36*(2), 334–357. https://doi.org/10.1016/j.ijforecast.2019.05.010
4. Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, *61*(4), 1645–1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x
5. Baker, M., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, *21*(2), 129–152. https://doi.org/10.1257/jep.21.2.129
6. Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, *104*(2), 272–287. https://doi.org/10.1016/j.jfineco.2011.11.002
7. Białkowski, J., Wagner, M., & Wei, X. (2023). Differences between NZ and U.S. individual investor sentiment: More noise or more information? *New Zealand Economic Papers*, *0*(0), 1–13. https://doi.org/10.1080/00779954.2023.2235751
8. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1
9. Bouteska, A. (2019). The effect of investor sentiment on market reactions to financial earnings restatements: Lessons from the United States. *Journal of Behavioral and Experimental Finance*, *24*, 100241. https://doi.org/10.1016/j.jbef.2019.100241
10. Brown, G. W. (1999). Volatility, Sentiment, and Noise Traders. *Financial Analysts Journal*, *55*(2), 82–90. https://doi.org/10.2469/faj.v55.n2.2263
11. Chau, F., Deesomsak, R., & Koutmos, D. (2016). Does investor sentiment really matter? *International Review of Financial Analysis*, *48*, 221–232. https://doi.org/10.1016/j.irfa.2016.10.003
12. Chen, G., Firth, M., & Rui, O. M. (2001). The Dynamic Relation Between Stock Returns, Trading Volume, and Volatility. *Financial Review*, *36*(3), 153–174. https://doi.org/10.1111/j.1540-6288.2001.tb00024.x
13. Daniel, K., Hirshleifer, D., & Teoh, S. H. (2002). Investor psychology in capital markets: Evidence and policy implications. *Journal of Monetary Economics*, *49*(1), 139–209. https://doi.org/10.1016/S0304-3932(01)00091-5
14. Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.
15. Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010
16. Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics*, *123*(3), 441–463. https://doi.org/10.1016/j.jfineco.2016.11.004
17. Fisher, K. L., & Statman, M. (2006). Market Timing in Regressions and Reality. *Journal of Financial Research*, *29*(3), 293–304. https://doi.org/10.1111/j.1475-6803.2006.00179.x
18. French, K. R. (2024). *Kenneth R. French*. Kenneth R. French - Data Library. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
19. González-Sánchez, M., & Morales de Vega, M. E. (2021). Influence of Bloomberg's Investor Sentiment Index: Evidence from European Union Financial Sector. *Mathematics*, *9*(4), Article 4. https://doi.org/10.3390/math9040297
20. Gric, Z., Bajzík, J., & Badura, O. (2023). Does sentiment affect stock returns? A meta-analysis across survey-based measures. *International Review of Financial Analysis*, *89*, 102773. https://doi.org/10.1016/j.irfa.2023.102773
21. Haritha, P. H., & Rishad, A. (2020). An empirical examination of investor sentiment and stock market volatility: Evidence from India. *Financial Innovation*, *6*(1), 1–15. https://doi.org/10.1186/s40854-020-00198-x
22. Hawaldar, I. T., & Rahiman, D. H. U. (2019). *Investors Perception Towards Stock Market: An Exploratory Approach*. *8*(12).

23. Hjalmarsson, E. (2010). Predicting Global Stock Returns. *Journal of Financial and Quantitative Analysis*, *45*(1), 49–80. https://doi.org/10.1017/S0022109009990469
24. Ibrahim, F. A., & Benli, V. F. (2022). Impact of investors sentiment on IPO performance: Evidence from NASDAQ and NYSE. *Journal of Business Economics and Finance*, *11*(1), Article 1. https://doi.org/10.17261/Pressacademia.2022.1548
25. Jacobs, H. (2015). What explains the dynamics of 100 anomalies? *Journal of Banking & Finance*, *57*, 65–85. https://doi.org/10.1016/j.jbankfin.2015.03.006
26. Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, *132*(1), 126–149. https://doi.org/10.1016/j.jfineco.2018.10.001
27. Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, *45*(1), 20–31. https://doi.org/10.1037/0022-3514.45.1.20
28. Joseph, K., Babajide Wintoki, M., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, *27*(4), 1116–1127. https://doi.org/10.1016/j.ijforecast.2010.11.001
29. Kim, K., & Ryu, D. (2021). Term structure of sentiment effect on investor trading behavior. *Finance Research Letters*, *43*, 102005. https://doi.org/10.1016/j.frl.2021.102005
30. Kurov, A. (2008). Investor Sentiment, Trading Behavior and Informational Efficiency in Index Futures Markets. *Financial Review*, *43*(1), 107–127. https://doi.org/10.1111/j.1540-6288.2007.00188.x
31. Lai, J., Zheng, T., Yi, H., & Dong, D. (2014). Investor sentiment affects the trading volume: An evidence from network forums text analysis. *2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 666–670. https://doi.org/10.1109/FSKD.2014.6980914
32. Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking & Finance*, *26*(12), 2277–2299. https://doi.org/10.1016/S0378-4266(01)00202-3
33. Naik, P. K., & Sethy, T. K. (2022). Stock Returns, Trading Volumes and Market Volatility: A Study on the Indian Stock Market. *The Indian Economic Journal*, *70*(3), 406–416. https://doi.org/10.1177/00194662221104753
34. Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*(3), 703–708. https://doi.org/10.2307/1913610
35. Park, C.-H., & Irwin, S. H. (2007). What do we know about the profitability of technical analysis? *Journal of Economic Surveys*, *21*(4), 786–826.
36. Sayim, M., Morris, P. D., & Rahman, H. (2013). The effect of US individual investor sentiment on industry-specific stock returns and volatility. *Review of Behavioural Finance*, *5*(1), 58–76. https://doi.org/10.1108/RBF-01-2013-0006
37. So, S. M. S., & Lei, V. U. T. (2011). *Investor Sentiment and Trading Volume* (SSRN Scholarly Paper 2007696). https://doi.org/10.2139/ssrn.2007696
38. Sun, L., Najand, M., & Shen, J. (2016). Stock return predictability and investor sentiment: A high-frequency perspective. *Journal of Banking & Finance*, *73*, 147–164. https://doi.org/10.1016/j.jbankfin.2016.09.010
39. Uygur, U., & Tas, O. (2012). Modeling the effects of investor sentiment and conditional volatility in international stock markets. *Journal of Applied Finance and Banking*, *2*(5), 239–260.
40. Uygur, U., & Taş, O. (2014). The impacts of investor sentiment on returns and conditional volatility of international stock markets. *Quality & Quantity*, *48*(3), 1165–1179. https://doi.org/10.1007/s11135-013-9827-3

41. Veld, C., & Veld-Merkoulova, Y. V. (2008). The risk perceptions of individual investors. *Journal of Economic Psychology*, *29*(2), 226–252. https://doi.org/10.1016/j.joep.2007.07.001
42. Visaltanachoti, N., Lu, L., & (Robin) Luo, H. (2007). Holding periods, illiquidity and disposition effect in the Chinese stock markets. *Applied Financial Economics*, *17*(15), 1265–1274. https://doi.org/10.1080/09603100600905053
43. Wang, W., Su, C., & Duxbury, D. (2022). The conditional impact of investor sentiment in global stock markets: A two-channel examination. *Journal of Banking & Finance*, *138*, 106458. https://doi.org/10.1016/j.jbankfin.2022.106458
44. Wang, Y.-H., Keswani, A., & Taylor, S. J. (2006). The relationships between sentiment, returns and volatility. *International Journal of Forecasting*, *22*(1), 109–123. https://doi.org/10.1016/j.ijforecast.2005.04.019
45. Wikipedia. (2024). List of S&P 500 companies. In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=List_of_S%26P_500_companies&oldid=1193421148
46. Yahoo. (2024). *Yahoo Finance—Stock Market Live, Quotes, Business & Finance News*. https://finance.yahoo.com/
47. Zhang, J. L., Härdle, W. K., Chen, C. Y., & Bommes, E. (2016). Distillation of News Flow Into Analysis of Stock Reactions. *Journal of Business & Economic Statistics*, *34*(4), 547–563. https://doi.org/10.1080/07350015.2015.1110525

About the authors

Aleš Kresta serves as an Associate Professor in the Department of Finance at the Faculty of Economics, VSB—Technical University of Ostrava, where he earned his Ph.D. in Finance. His primary research interests encompass risk estimation, backtesting, portfolio optimisation, financial time series modelling, and soft computing within the realm of quantitative finance. Actively involved in numerous scientific projects, he can be contacted at **ales.kresta@vsb.cz**.

Jialei Xiong is a Ph.D. student in the Department of Finance at the Faculty of Economics, VSB – Technical University of Ostrava. Her research and dissertation focus on issues related to portfolio optimisation, specifically emphasising sentiment measures. The author can be contacted at **jialei.xiong@vsb.cz**.

Bahate Maidiya, is a PhD student in the Department of Finance at the Faculty of Economics, VSB – Technical University of Ostrava. She directs her research toward biases in investment decision-making. The author can be contacted at **bahate.maidiya@vsb.cz**.

Age Management Practices and Benefits in Organisation: An Evaluation of the Effect of Economic Sector, Organisation Size, and Family Business Status

Terezie Krestová, Aleš Kresta, Lucie Bestová
VSB – Technical University of Ostrava, Faculty of Economics, Czech Republic

Abstract

Background: The ageing of the population impacts all areas and is also a challenge for human resources management. To solve this issue, age management practices are used in organisations. To get the best potential out of everyone, this concept should not only involve older employees but should be focused on all of them. Choosing the right practice to get the desired results is a task for managers. **Objectives:** The paper aims to determine whether the sector of the economy, the size of the organisation, or the family business status plays a role in determining age management practices and observed benefits. **Methods/Approach:** Using a questionnaire survey, the most commonly used practices and the observed benefits were identified. Using the chi-square test, differences in chosen categories were confirmed. **Results:** It can be stated that choosing age management practices is influenced by the economic sector, and the size of the organisation influences observed benefits. **Conclusions:** The results can guide organisations on which practices to choose and what benefits to expect from implementing age management practices.

Keywords: age management; age management practices; age management benefits; economic sector; size of the organisation; family business

JEL classification: J11, J14, L20, M51
Paper type: Research article

Received: Jan 26, 2024
Accepted: Jun 11, 2024

Acknowledgements: This paper was supported by the Student Grant Competition of the Faculty of Economics, VSB-Technical University of Ostrava, project registration number SP2023/025: Family Business: An Analysis of Factors Influencing Employees' Attitudes in Choosing an Employer and the European Union under the REFRESH Research Excellence For Region Sustainability and High Tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition. All support is greatly appreciated.

Citation: Krestová,T., Kresta, A., & Bestová, L. (2024). Age Management Practices and Benefits in Organisation: An Evaluation of the Effect of Economic Sector, Organisation Size, and Family Business Status. *Business Systems Research, 15*(2), 83-99.
DOI: doi.org/10.2478/bsrj-2024-0019

Introduction

Demographic ageing is the current issue (Eurostat, 2022; Bureau of Labour Statistics, 2021) and is evident in many areas, not least in the management of organisations. As the population ages, so will the workforce. While young people tend to enter the labour market later due to longer education, older individuals have increased their participation in the labour market, contributing to a greying workforce (Sousa et al., 2020). Current demographic trends in the EU put long-term economic prosperity and competitiveness at risk and contribute to labour shortages in the EU (Europa, 2023). For organisations, this represents changes in the age structure of employees, and this fact must be considered in the strategic management of organisations, as well as reflected in human resource management (HRM) (Salminen et al., 2017). Changes are not only caused by the ageing of the population but also by different attitudes to the workplace of different generations and affect both the labour market and the social policy of the state.

Ageing employees bring new challenges to human resource management (HRM). These include, for example, the ageing of skilled employees, a decrease in the number of graduates who enter the labour market, problems of the middle generation (the so-called sandwich generation) who take care of children as well as elderly parents, and more (Marcus et al., 2020). On the other hand, changes in attitudes are linked to different expectations in the work regime, motivation, career development, and others (Kinger et al., 2023). Despite all these aspects, it is recommended that the age diversity of the employees at work be ensured. Scheuer and Loughlin (2021) highlight the potential competitive advantages, while Li et al. (2020), Ali et al. (2019), and OECD (2020) highlight the positive impact on organisational performance and results. Li et al. (2020) further suggest that age diversity can enhance human and social capital, with the effects being amplified by functional diversity and age-inclusive management.

Organisations must adopt practices to prolong the working careers of employees and also to harness the potential of all employee groups in the company to achieve employment satisfaction. An approach that incorporates diverse employee requirements across different age groups to optimise their contributions to the organisation is termed age management and is part of diversity management. Many companies have incorporated this approach into their human resources management in recent years. However, for many organisations, the question remains about which practices should be chosen to be most effective. The suggestion of selection by measuring the Work Ability Index is addressed in Ilmarinen and Ilmarinen (2015) or Hlatká et al. (2021). The disadvantage is that a specialist should develop this procedure. It mainly focuses on measuring the Work Ability Index. In the Czech Republic, this method is not widely available and can only be used on a paid basis. Krestová et al. (2021a) or Franek et al. (2023) suggest the use of multi-criteria decision-making methods, but these are more difficult to implement on their own. The results proposed in this article can serve as a guide for managers on which methods are appropriate to use in terms of their economic sector and firm size.

The paper aims to determine whether the sector of the economy, the size of the organisation, or the family business status plays a role in choosing age management practices and observed benefits.

The remainder of the paper is organised as follows: The next section presents a literature review focusing on age management, its benefits, and practices. The third section discusses the methods applied. The fourth section presents the results. The fifth and sixth sections provide the discussion and conclusion, respectively.

Literature review

Human resource management is indisputably a key activity in an organisation today. Human capital is a competitive advantage and a driver of corporate performance. However, more than in the past, HR managers must face a rapidly changing environment and take into account that there are four generations working side by side with different habits and expectations (OECD, 2020). Creating healthy working conditions for employees of all ages to maintain the performance level and competitiveness of organisations will be an especially important challenge (Urbancová et al., 2020). How to take into account the requirements of employees of all age groups and thus exploit their potential is the subject of age management.

Age management

Age management belongs to the concept of diversity management, which is a managerial approach that promotes diversity in the workplace as a means of greater efficiency (Richter, 2014; Aaltio, 2017; Inegbedion et al., 2020; Li et al., 2020; Sousa et al., 2020). In addition to age, organisations deal with racial, gender, sexual, and other types of diversity. Creating diverse teams and respecting personal and cultural diversity leads to both success and increased organisational performance (Howard et al., 2017; Lee et al., 2020).

"Age management requires taking the employee's age and age-related factors into account in daily work management, work planning, and work organisation; therefore, everyone, regardless of age, can achieve personal and organisational goals healthily and safely" (Ilmarinen, 2005, p.120).

Age management is commonly associated with older employees (50+), and initial studies focused primarily on this group (Ilmarinen et al., 1997). However, this perspective is not entirely accurate. To extend an individual's career, the approach should not be limited to those classified as older workers (Covarrubias Venegas, 2019). Additionally, it is inappropriate to prioritise a specific employee category exclusively. Age management strategies should include all groups of workers. The organisational objective and HRM policies should strive to leverage the strengths of employees irrespective of their age (Wiktorowicz et al., 2022). The concept of age management is related to the concept of work ability (Ilmarinen et al., 1997; Gould et al., 2008) and can be measured by the Work Ability Index (WAI). Work ability is defined as the balance between personal resources and work demands. Personal resources include, for example, health and functional capacity, education, knowledge, skills, and motivation. The demands of work include, for example, the job description, its complexity, its organisation, the work environment, and the management style. Heilmann (2017) summarised that age management could be understood as a collection of best HRM practices aiming to sustain and enhance employees' work ability while helping organisations reach their goals.

Salminen et al. (2016) noted that sometimes age management is viewed and analysed from an organisational perspective, while other approaches see it from an employee perspective or a combination of these perspectives. Heilmann (2017) identified four different perspectives: social, managerial, organisational and individual worker perspectives. The implementation of age management will affect all of these perspectives. In general, the authors see the contribution of age management at the following three levels: individual, organisational and social.

At the ***individual level***, it is evident that age management improves the efficient utilisation of capabilities, providing an opportunity for prolonged participation in the workforce and adaptation to evolving educational demands in later life stages (Fabisiak et al., 2012). To extend an individual's working life, age management practices cannot be applied only to older workers. It is necessary to care for their work ability in the long term, which consists of changing attitudes toward the planning of their future, lifelong learning, and health (Nilsson, 2020). The goal is to achieve satisfaction with a reasonable quality of professional and personal life. Furthermore, age management not only contributes to individual well-being but also establishes optimal conditions for workers to utilise their abilities fully, thus prolonging employment (Košir et al., 2016). This comprehensive approach ensures the integration of age management strategies that improve both the individual and organisational dimensions of the workforce.

From an ***organisational point of view***, the age management problem can be dissected through the lens of human resource management. The engagement of managers in managing age becomes prominent during scenarios of restructuring, organisational changes, technological advancements, or the departure of experienced and highly qualified personnel. Although the frequent motivation behind the introduction of age management in an organisation is the departure of experienced and highly qualified ageing employees, its introduction has many benefits for the organisation as a whole (Hlatká et al., 2021; Boehm et al., 2021). Beyond responsive actions, age management can be strategically considered for cost control, meeting customer expectations, or retaining qualified staff. Corporate practices include raising awareness among managers and employees, implementing established practices in recruitment, training, and development, and facilitating age optimisation. Initiatives include the introduction of lifelong learning programmes, health and safety practices, and flexible employment arrangements (Fabisiak et al., 2012). The adoption of age management practices results in socially responsible behaviour of employers, positively influencing the retention and development of employees and building the image of the organisation (Covarrubias Venegas, 2019). Meanwhile, as Blomé et al. (2020) noted, organisations that apply age management are those from which other actors can benefit.

From society's perspective, age management is important. In light of demographic changes, age management has become a macroeconomic issue of labour market policies, extending to the national level (Fabisiak et al., 2012). The public interest and the contribution of ageing workers to economic and social development are combined. Pension and social policy instruments must be used to maximise these benefits.

Age management practices and benefits

Implementing age management is also a strategic issue. Age management must be fully supported by top management, and its goals should be reflected in the strategic goals of the organisation (Ilmarinen, 2005). Hlatká et al. (2021) and Urbancová et al. (2020) see the biggest advantages of this concept in increasing HR quality and stability, cost reductions, greater satisfaction of workers, sharing of experience workers' skills with young employees, and increasing market competitiveness. Based on the research of Krestová et al. (2021b) or Krestová et al. (2023), the benefits of age management may be seen in these areas:

- Retaining Key Staff
- Reducing Employee Turnover
- Increasing Employee Motivation
- Improving Organisational Culture and Positively Impacting Employee Values
- Developing Employee Education, Knowledge, and Skills
- Enhancing the Company's Reputation
- Improving Employee Performance

- Enhancing Employee Health
- Acquiring New Talents

Specific activities related to age management were described in Blomé et al. (2020), Urbancová (2017), Čiutiene et al. (2015), Fabisiak et al. (2012) and served as sources for the survey (Krestová et al., 2021b; Krestová et al., 2023). The practices are aimed at both workers themselves, their health and psychological well-being, and their workplace. How Farr-Wharton et al. (2023) stated that it is the organisation of work and the work environment, rather than chronological age per se, that influences workers' well-being and, therefore, the length of their working life. Sousa et al. (2020) see potentially four areas in which age management practices should focus: development, maintenance, utilisation, and accommodative practices. Based on these articles and surveys, specific practices used within age management were defined. They are:

- Adaptation of work programmes to fit different age groups
- Reassignment of staff to more suitable positions
- Care for employee health, promotion of physical fitness, and healthy eating
- Support for personal and career development through skills training at every stage of an employee's career
- Development of health and safety practices
- Workforce planning for age diversity and promotion of a positive age policy
- Creating Tools to Support Intergenerational Learning
- Motivational Programmes According to the Needs of Different Generations
- Tailored Further Training for Older Workers
- Use of Mentoring and Reverse Mentoring
- Special Forms of Recruitment for Different Age Groups

The applicability of age management practices varies across organisations, and not all associated benefits are universally observable. In implementing age management within an organisation, as proposed by Krestová et al. (2021a) or Franek et al. (2023), the organisation's size or orientation must be considered when selecting suitable practices for introduction.

This raises the research question of whether or not the economic sector, the size of the organisation, or whether or not the organisation is a family business have an impact on the choice of practices applied or the observed benefits. These 3 factors were chosen because not all practices are usable in all economic sectors, and the size of the organisation may, in turn, influence how many employees are affected by the practices. The impact of the family business was examined because it is different from non-family businesses in many ways. There is a noticeable combination of two systems of values, expectations, and roles - the family and the business (Baron et al., 2021; Woodfield et al., 2021) also mention specifics in the organisational culture of family businesses, which provides a relaxed atmosphere that contributes to higher employee satisfaction and well-being, and also allows work and family life to be reconciled. On the other hand, Tabor et al. (2018) state that the conditions for learning and growth are often only available to family members and then limited to non-family employees. A frequent issue in families is that of succession. Passing on experience and power to successor generations can be a challenge that age management practices can help.

Therefore, it is useful to see whether the applied practices and the observed benefits differ depending on the factors selected, and this is the subject of our research.

Methodology

The methods used in this paper were a questionnaire survey and statistical analysis. The research focused on the use of age management in organisations was carried out in March 2023. The goal was to find out if organisations know this concept, and if so, if age management is or is not used, how long it has been applied, what would force organisations to address this issue, which specific practices are used, and which benefits are observed. A questionnaire on age management applications consisting of 20 questions in total was created and distributed to organisations. Questions were both open and closed. The respondents could select the practices applied and the benefits observed from the options offered. These options were defined based on available literature (Urbancová, 2017; Blomé et al., 2020; Čiutiené et al., 2015; Fabisiak et al., 2012) and our previous research (Krestová et al., 2021b). The respondents were also able to add their responses. As part of the identification questions, organisations filled out which sector of the economy they belong to, how many employees they have, and if they are a family business.

Czech organisations across all sectors were contacted. From a purchased database of organisations (Databaze firem CR, 2019) containing 3,217 family business contacts and 3,294 non-family business contacts, 650 organisations from each group (approximately 20%) were randomly selected and contacted. These organisations were selected without considering the size and economic sector, as these characteristics were not known beforehand. The CAWI method was used to reach the organisations. In total, 1,300 organisations were contacted, and 192 completed the questionnaire. Therefore, the return rate was 14.8%. After verifying the data, three questionnaires were identified as filled only partially and were removed from the data set. From the responses, the most commonly applied practices (observed benefits, respectively) were identified, and it was analysed whether they differ depending on the economic sector, the size of the organisation, or the family business status.

The calculation scheme is shown in Table 1. The columns correspond to the organisation characteristics under study: whether the organisation is primary, secondary, or tertiary in the case of the economic factor; whether it is small (less than 50 employees), medium (50-249 employees), or large (more than 250 employees) in the case of the size of the organisation; and whether it is a family business or not in the case of the family status.

The rows always represent a binary variable identifying whether the organisation selected the specific practice as applied or the benefit as observed. The data in the tables represent the number of organisations from a given category (column) that apply the given practice or observe the given benefit (row). The sums of rows and columns are called marginal totals.

The sums of the rows represent the number of organisations that apply or do not apply the practice or observe the benefit. The sums of the columns represent the total number of organisations in each category. Moreover, the frequencies as the ratios of the number of organisations in the category applying the practice (observing the benefit) to the total number of organisations in the category were also calculated.

Table 1
Frequencies of Age Management Practices

Status	Category 1	Category 2	Marginal Total
Applies	A	B	A+B
Does not apply	C	D	C+D
Marginal total	A+C	B+D	A+B+C+D
Frequencies	A/ (A+C)	B/ (B+D)	(A+B)/(A+B+C+D)

Source: Authors' work

The underlying question is whether the application of practice or the observation of the benefit depends on the organisation's category. If there is no difference between the categories, the observed quantities should correspond to the expected frequencies given by the marginal totals. On the contrary, if there is an association between the organisation category and the relative number of organisations selecting the practice or benefit, the observed quantities would differ from the expected ones. This is tested by the chi-square test of the independence of variables in a contingency table with Yates's (1934) correction for continuity. For each practice (benefit, respectively), the null hypothesis is as follows:

- H_0: *There is no significant association between the application of the practice (observing the benefit, respectively) and the organisation category.*

Versus the alternative hypothesis:

- H_A: *There is a significant association between the practice (observing the benefit, respectively) and the organisation category.*

When the null hypothesis is rejected in favour of the alternative hypothesis, it is found that the attitude to applying the given practice (observing benefit, respectively) depends on the organisation category. However, since a two-sided test was performed, the alternative hypothesis does not indicate the direction of the effect. To find the direction of the effect, the calculated frequencies in the contingency tables were compared to each other.

A total of 189 organisations participated in the survey, of which 6 were from the primary sector, 129 were from the secondary sector, and 54 were from the tertiary sector. With respect to the requirements of the chi-square test, the primary sector was omitted due to the low number of responses. The analysis was carried out using only 183 organisations from the secondary and tertiary sectors. From these, in terms of organisation size, 69 respondents belong to small organisations, 76 respondents belong to medium organisations, and 38 organisations belong to large organisations. In terms of family business status, 97 organisations are family businesses, and 86 organisations are not.

Results

First, the most frequently identified applied age management practices and the observed benefits are presented (Table 2). Then, findings on whether the applied practices and observed benefits differ depending on the economic sector, size of the organisation, and family business status are disclosed.

Table 2
Usage of Age Management Practices

Age management practices	Frequency	%
Adaptation of Work Programmes to Fit Different Age Groups	110	58,2%
Reassignment of Staff to the More Suitable Position	101	53,4%
Care for Employee Health, Promotion of Physical Fitness, and Healthy Eating	94	49,7%
Support for Personal and Career Development through Skills Training at Every Stage of an Employee's Career	91	48,1%
Development of Health and Safety Practices	62	32,8%
Workforce Planning for Age Diversity and Promotion of a Positive Age Policy	52	27,5%
Creating Tools to Support Intergenerational Learning	29	15,3%
Motivational Programmes According to the Needs of Different Generations	27	14,3%
Tailored Further Training for Older Workers	19	10,1%
Use of Mentoring and Reverse Mentoring	11	5,8%
Special Forms of Recruitment for Different Age Groups	7	3,7%

Source: Authors' work

Next, it is examined whether the economic sector, size of the organisation, and family business status have an impact. Statistical significance was determined at significance levels of 10%, 5% and 1%.

To make the process clear, it is illustrated in the case of the *Adaptation of work programmes to fit different age groups*' practices and economic sector types, see Table 3. From the contingency table, the marginal totals can be observed, i.e. from the total of 183 organisations: 110 apply the practice, 73 do not apply the practice, 129 are from the secondary sector, and 54 are from the tertiary sector. Applying the chi-square test, a p-value of 0.046 was obtained (statistics 4.00, critical value 3.84 at 5% significance level, degrees of freedom 1) and thus the null hypothesis was rejected in favour of the alternative hypothesis at a 5% significance level, resulting in the finding that there is the dependence of practice application on economic type of organisation. The observed probability of applying the given practice conditioned on the economic sector is given by the frequencies. It can be summarised that applying *Adaptation of Work Programmes to Fit Different Age Groups* practice is more common for organisations in the tertiary sector and less common in the secondary sector. In this way, the process continues with other practices. For the p-values of the chi-square test, see Table I in the Appendix. In the case of the two least common practices, i.e. *Use of Mentoring and Reverse Mentoring*, *Special Forms of Recruitment for Different Age Groups*, in the contingency table, there are some values which are lower than 5 and thus, in this case, the rule of thumb is not satisfied. These results must be taken with caution.

Table 3
Illustration of Calculation for Adaptation of work programmes to fit different age groups' practice

Status	Secondary	Tertiary	Marginal Total
Applies	71	39	110
Does not apply	58	15	73
Marginal total	129	54	183
Frequencies	55,04%	72,22%	60,11%

Source: Authors' work

The results show that factors of economic sector and size influence the application of age management practices more often than family business status. Detailed results are presented in Table 4.

Table 4
Influence of Selected Factors on the Choice of Practices

Age Management Practices	Factors		
	Economic Sector	Organisational Size	Family Bus. Status
Adaptation of Work Programmes to Fit Different Age Groups	More Often Applied in III, Less in II**	Not Significant	Not Significant
Reassignment of Staff to the More Suitable Position	More Often Applied in II, Less in III**	Not Significant	Not Significant
Care for Employee Health, Promotion of Physical Fitness, and Healthy Eating	More Often Applied in III, Less in II***	Increases with the Organisation Size***	Not Significant

Support for Personal and Career Development through Skills Training at Every Stage of an Empl. Career	More Often Applied in III, Less in II**	Increases with the Organisation Size***	Not Significant
Development of Health and Safety Practices	Not Significant	Increases with the Organisation Size***	Not Significant
Workforce Planning for Age Diversity and Promotion of a Positive Age Policy	Not Significant	Not Significant	Not Significant
Creating Tools to Support Intergenerational Learning	More Often Applied in III, Less in II*	Increases with the Organisation Size***	Less Applied in Family Businesses, More in Non-family**
Motivational Programmes According to the Needs of Different Generations	Not Significant	Decreases with the Organisation Size***	Not Significant
Tailored Further Training for Older Workers	Not Significant	Not Significant	Not Significant
Use of Mentoring and Reverse Mentoring	More Often Applied in III, Less in II ***	Increases with the Organisation Size **	Not Significant
Special Forms of Recruitment for Different Age Groups	Not Significant	Not Significant	Not Significant

Note: II is the secondary, III is the tertiary sector of the economy
* 10% level of significance, ** 5% level of significance, *** 1% level of significance
Source: Authors' work

The next focus was on the benefits observed by organisations that apply age management. Table 5 shows their order and frequency.

Table 5
The level of usage of observed benefits

Age Management Benefits	Frequency	%
Retaining Key Staff	151	79,9%
Reducing Employee Turnover	113	59,8%
Increasing Employee Motivation	105	55,6%
Improving Organisational Culture and Positively Impacting Employee Values	72	38,1%
Developing Employee Education, Knowledge, and Skills	76	40,2%
Improving Employee Performance	71	37,6%
Enhancing the Company's Reputation	66	34,9%
Enhancing Employee Health	62	32,8%
Acquiring New Talents	40	21,2%

Source: Authors' work

Subsequently, it is examined whether the economic sector, the size of the organisation, and the family business status have an impact on the observed results. Statistical significance was determined at significance levels of 10%, 5% and 1%. To make the process clear, it is illustrated in the case of *Retaining Key Staff* benefits and organisation size, see Table 6. From the contingency table, marginal totals can be observed, i.e. from the total of 183 organisations: 69 are small (less than 50 employees), 76 are medium (50-249 employees), and 38 are large (more than 250 employees), 151 observe the benefit, and 32 do not. Applying the chi-square test described in the previous section, a p-value of 0.985 was obtained (statistics 0.31, critical value 4.61 at 10% significance level, degrees of freedom 2). Thus, the null hypothesis cannot be rejected in favour of the alternative hypothesis, resulting in the finding that there is no dependence on observing the benefit of *Retaining Key Staff* on organisation size. In this way, the process continues with other benefits. For the p-values of the chi-square test, see Table II in the Appendix.

Table 6
Illustration of Calculation for Retaining Key Staff Benefit

Status	Small	Medium	Large	Marginal Total
Observes	57	63	31	151
Does not observe	12	13	7	32
Marginal total	69	76	38	183
Frequencies	82.61%	89.90 %	81.58%	82.51%

Source: Authors' work

The results show that the size of the organisation influences the observed benefits more often than family business status and economic sector. Detailed results are presented in Table 7.

Table 7
Influence of Selected Factors on the Observed Benefits

Observed Benefits	Factors		
	Economic Sector	Organisational Size	Family Bus. Status
Retaining Key Staff	Not Significant	Not Significant	Not Significant
Reducing Employee Turnover	Not Significant	Not Significant	Not Significant
Increasing Employee Motivation	More Often Observed in III, Less II**	Increases with Organisation Size*	Not Significant
Improving Org. Culture and Positively Impacting Employee Values	More Often Observed in III, Less II*	Increases with Organisation Size***	Not Significant
Developing Employee Education, Knowledge, and Skills	More Often Observed in III, Less II ***	Increases with Organisation Size***	Less Observed in Family Bus., More in Non-family ***
Improving Employee Performance	Not Significant	Not Significant	Not Significant
Enhancing the Company's Reputation	Not Significant	Increases with Organisation Size***	Less Observed in Family Bus., More in Non-family **
Enhancing Employee Health	Not Significant	Increases with Organisation Size**	Not Significant
Acquiring New Talents	Not Significant	Increases with Organisation Size***	Less Observed in Family Bus., More in Non-family **

Note: II is the secondary, III is the tertiary sector of the economy
* 10% level of significance, ** 5% level of significance, *** 1% level of significance
Source: Authors' work

Discussion

As is visible in Table 4, *Reassignment of staff to more suitable positions* is used more frequently in the secondary economic sector and less in the tertiary sector. On the contrary, *Adaptation of work programmes to fit different age groups*, *Care for employee health, promotion of physical fitness and healthy eating*, *Support for personal and career development through skills training at every stage of an employee's career*, *Creating tools to support intergenerational learning* and *Use of mentoring and reverse mentoring* are more often applied in the tertiary economic sector and less in secondary.

With the increasing size of the organisation, they significantly increase using these five practices: *Care for employee health, promotion of physical fitness and healthy eating*, *Support for personal and career development through skills training at every stage of an employee's career*, *Development of health and safety practices*, *Creating tools to support intergenerational learning* and *Use of mentoring and reverse mentoring*. On the other hand, applying *Motivational programmes according to the needs of different generations* decreases with the size of the organisation.

As far as the relationship between being a family business or not and the frequency of application of certain practices is concerned, the effect is very small. In family business organisations, only *Creating tools to support intergenerational learning* is less often applied than in non-family businesses. Other practices and their frequency of application are not affected by this factor.

Focussing on the benefits, the two most frequently observed benefits, which are *Retaining key staff* and *Reducing employee turnover*, and *Improving Employee Performance* are not influenced by any of the factors examined. However, all other observed benefits are more or less influenced by the size of the organisation. The benefits are observed more frequently as the size of the organisation increases.

As far as the economic sector is concerned, its influence can only be seen in three benefits. These are more common in the tertiary sector and less in the secondary sector. They are *Increasing employee motivation, Improving Organisational Culture and Positively Impacting Employee Values* and *Developing employee education, knowledge, and skills*. Concerning the family business status, it can be argued that the benefits of *Developing employee education, knowledge, and skills*, *Enhancing the company's reputation* and *Acquiring new talents* are less frequently observed in the family business and more often in non-family.

If organisations decide to implement age management, these observed benefits from other organisations can help set expectations for the implementation of age management practices.

We can summarise that the economic sector of the organisation most strongly influences the choice of age management practices, and the observed benefits vary mostly depending on the size of the organisation. In the literature review, no prior research was found that examines the application of specific age management practices in relation to these or other factors. However, the partial results can be compared with those of Sousa et al. (2020), who found that workers in manufacturing firms, which corresponds to the secondary sector, will benefit from practices aimed at accommodating practices such as flexible work arrangements (e.g. compressed workweek, exemption from night shifts). This may correspond to practices *Adaptation work programmes to fit different age groups,* and *Reassignment of employees to a more suitable position*. However, Urbancová et al. (2017) concluded that the economic sector and the size of the organisation do not influence whether organisations apply the age management strategy or not.

In the context of further research, it would be appropriate to focus in detail on individual factors and their impact on the implementation of specific age management practices. The reasons for their influence can also be examined. It is possible to look at the links between applied practices and observed benefits both in general and depending on sector, firm size, or family business status. The authors now focus on the analysis of whether using a combination of different practices increases the likelihood of observing different benefits.

A limiting factor may be that HR managers will not be able to define which practices produce which benefits. Suggested practices are often applied not only in the context of age management but are used in general to contribute to greater employee satisfaction in organisations. The second limitation arises from the fact that only organisations operating in the Czech Republic were contacted, which may limit the generalisation of the results. However, it is assumed that many companies in the Czech Republic belong to international concerns and, at least within the European Union, the practices will not differ much. The last limiting factor is the lack of respondents from the primary sector. Given the nature of work in this sector, it seems more appropriate to approach these organisations differently.

Conclusion

Human resources are an inherent competitive advantage for organisations today. However, the changing needs of the emerging generation, the need to extend an individual's active working career, and the demands placed on the workplace are challenging today's managers. Age management can help manage an age-diverse workforce.

However, selecting the appropriate practices and expectations of their application can be challenging. This paper aimed to find out whether the practices applied and the benefits observed vary depending on the economic sector, the size of the organisation, and whether it is a family business.

Based on a proprietary survey conducted in March 2023, the most commonly used age management practices and the observed benefits were identified. Subsequently, a chi-square test was used to determine if the results differed based on the selected categories.

It can be concluded that the influence of individual factors has been identified, but most of the time, it is not significant or occurs in less common practices. The economic sector, in which the organisation operates most commonly, influences the applied practices, while the size of the organisation primarily influences the observed benefits. These effects are often rather minor, so it cannot be assumed that the practices will not have an effect if they are implemented in the organisation.

This knowledge can help an organisation that decides to implement age management to select the most appropriate practices for the organisation. The suggested benefits can also serve to define better the objectives that the organisation can expect to achieve with the implementation of age management.

References

1. Ali, M., & French, E. (2019). Age diversity management and organisational outcomes: The role of diversity perspectives. Human Resource Management Journal, 29(2), 287-307. https://doi.org/10.1111/1748-8583.12225
2. Aaltio, I. (2017). Ageing, Gender and Leadership: A Study Based on Tolkien's The Lord of the Rings. In I. Aaltio, A. J. Mills, & J. H. Mills (Eds.), Ageing, Organisations and Management (pp. 65-85). Springer International Publishing. https://doi.org/10.1007/978-3-319-58813-1_4

3. Baron, J. & Lachenauer, R. (2021). Harvard Business Review Family Business Handbook: How to Build and Sustain a Successful, Enduring Enterprise. Harvard Business Review Press, Watertown
4. Blomé, M.W., Borell, J., Håkansson, C. & Nilsson, K. (2020). Attitudes toward elderly workers and perceptions of integrated age management practices, International Journal of Occupational Safety and Ergonomics, 26(1), 112–120
5. Boehm, S. A., Schröder, H., Bal, M., & Wang, M. (2021). Age-Related Human Resource Management Policies and Practices: Antecedents, Outcomes, and Conceptualizations. Work, Aging and Retirement, 7(4), 257-272. https://doi.org/10.1093/workar/waab024
6. Bureau of Labor Statistics. (2021). Number of people 75 and older in the labor force is expected to grow 96.5 percent by 2030' https://www.bls.gov/opub/ted/2021/ number-of-people-75-and-older-in-the-labor-force-is-expected-to-grow-96-5-ercent-by-2030.htm
7. Čiutienė, R., & Railaitė, R. (2015). Challenges of managing an ageing workforce. Engineering Economics, 26(4), 391-397. https://doi.org/10.5755/j01.ee.26.4.7081
8. Databaze firem CR (2019). Databáze firem ČR. http://www.databazefiremcr.cz
9. Eurostat. (2022). Population structure and ageing. https://ec.europa.eu/eurostat/ statisticsexplained/index.php?title=Population_structure_and_ageing
10. Europa. (2023). Demographic change in Europe. https://europa.eu/ eurobarometer/surveys/detail/3112
11. Fabisiak, J., & Prokurat, S. (2012). Age Management as a Tool for the Demographic Decline in the 21st Century: An Overview of its Characteristics. Journal of Entrepreneurship, Management and Innovation, 8(4), 83-96. https://doi.org/10.7341/2012846
12. Farr-Wharton, B., Bentley, T., Onnis, L. -ann, Caponecchia, C., Neto, A. D. A., O'Neill, S., & Andrew, C. (2023). Older Worker-Orientated Human Resource Practices, Wellbeing and Leave Intentions: A Conservation of Resources Approach for Ageing Workforces. International Journal of Environmental Research and Public Health, 20(3). https://doi.org/10.3390/ijerph20032725
13. Finkelstein, L. M., Truxillo, D. M., Fraccaroli, F., & Kanfer, R. (Eds.). (2015). Facing the Challenges of a Multi-Age Workforce. Routledge. https://doi.org/10.4324/9780203776322
14. Franek, J., & Krestová, T. (2023). Selection of appropriate age management measures using multi-criteria decision-making methods with interrelationships. International Journal of Multicriteria Decision Making, 9(4), 351-378. https://doi.org/10.1504/IJMCDM.2023.134925
15. Gould, R., Ilmarinen, J., Järvisalo, J. & Koskinen, S. (2008) Dimensions of Work Ability: Results of the Health 2000 Survey, Finnish Centre of Pensions, The Social Insurance Institution, Finnish Institute of Occupational Health, Helsinki.
16. Heilmann, P. (2017). Age management in organizations, perspective of middle aged employees, in Aaltio I., Mills A.J. and Mills, J.H. (Eds.): Ageing, Organizations and Management: Constructive Discourses and Critical Perspectives, pp.141–154, Springer International Publishing, Cham.
17. Hlatká, M., Stopka, O., Bartuška, L., Stopková, M., Yordanova, D. N., Gross, P., & Sádlo, P. (2021). Draft Methodology of the Age Management Implementation in Human Resource Management in a Transport Company. Journal of Risk and Financial Management, 14(4). https://doi.org/10.3390/jrfm14040183
18. Howard, T.L., Peterson, E.A. & Ulferts, G.W. (2017). A note on the value of diversity in organisations, International Journal of Management and Decision Making, 16(3), 187–195.
19. Ilmarinen, J. & Ilmarinen, V. (2015). Work Ability and Aging. in Finkelstein L. M., Truxillo D. M., Fraccaroli, F. and Kanfer, R. (eds). Facing the Challenges of a Multi-Age Workforce: A Use-Inspired Approach. SIOP Organizational Frontiers Series. New York: Routledge, pp. 134-156.

20. Ilmarinen, J. (2005). Towards a Longer Worklife, Finnish Institute of Occupational Health, Helsinki.
21. Ilmarinen, J., Tuomi K. & Klockars, M. (1997). Changes in the work ability of active employees over an 11-year period, Scandinavian Journal of Work, Environment & Health, 23(1), 49–57
22. Inegbedion, H., Sunday, E., Asaleye, A., Lawal, A., & Adebanji, A. (2020). Managing Diversity for Organizational Efficiency. SAGE Open, 10(1). https://doi.org/10.1177/2158244019900173
23. Košir, S. & Šoba, V.A. (2016). Theoretical and some practical perspectives on age diversity and comparative age management, International Journal of Work Innovation, 1(4), 309–327.
24. Krajňáková, E., & Vojtovič, S. (2017). Struggles of Older Workers at the Labour Market. Economics & Sociology, 10(1), 319-333. https://doi.org/10.14254/2071-789X.2017/10-1/23
25. Krestová, T. & Kresta, A. (2021a) Selection of suitable measures for the application of age management goals in a selected organization, Proceedings of the 14th International Conference on Strategic Management and its Support by Information Systems 2021. SMSIS 2021, pp.153–162.
26. Krestová, T. & Kresta, A. (2021b) Application of age management as a way to eliminate the risk associated with the demographic development of the population, Proceeding of the 15th International Conference Liberec Economic Forum. LEF 2021, pp.375–385.
27. Krestová, T., & Demlová. T. (2023). Applying Age Management in HRM: Evidence from Czech Companies. In: Proceedings of TheIIER international conference San Francisco, USA: 14th – 15th June, 2023, San Francisco, USA. Bhubaneswar: Institute for Technology and Research, 2023. s. 6-10.
28. Kinger, N., & Kumar, S. (2023). Generational differences in work values in the workplace. Folia Oeconomica Stetinensia, 23(2), 204-221. https://doi.org/10.2478/foli-2023-0027
29. Lee, H. W., & Kim, E. (2020). Workforce diversity and firm performance: Relational coordination as a mediator and structural empowerment and multisource feedback as moderators. Human Resource Management, 59(1), 5-23. https://doi.org/10.1002/hrm.21970
30. Li, Y., Gong, Y., Burmeister, A., Wang, M., Alterman, V., Alonso, A., & Robinson, S. (2021). Leveraging age diversity for organizational performance: An intellectual capital perspective. Journal of Applied Psychology, 106(1), 71-91. https://doi.org/10.1037/apl0000497
31. Marcus, J., Rudolph, C.W. and Zacher. H. (2020) An ecological systems framework for work and ageing, in Stone, D.L., Dulebohn, J.H. and Lukaszewski, K.M. (Eds.): Diversity and Inclusion in Organizations, Information Age Publishing, Charlotte.
32. Nilsson, K. (2020). A sustainable working life for all ages – The swAge-model. Applied Ergonomics, 86. https://doi.org/10.1016/j.apergo.2020.103082
33. Promoting an Age-Inclusive Workforce. (2020). OECD. https://doi.org/10.1787/59752153-en
34. Richter, D. (2014) 'Demographic change and innovation: the ongoing challenge from the diversity of the labor force', Management revue – Socio-economic Studies, 25(3), 166–184
35. Salminen, H. & von Bonsdorff, M. (2017) Late career and retirement in the context of changing careers, in Aaltio I., Mills A.J. and Mills, J.H. (Eds): Ageing, Organizations and Management: Constructive Discourses and Critical Perspectives, pp.89–112, Springer International Publishing, Cham.
36. Salminen, H., von Bonsdorff, M., Koponen, S. & Miettinen, M. (2016) Perceived development opportunities and reward satisfaction as antecedents of nurses' job withdrawal intentions, International Journal of Work Innovation, 1(4), 330–352.
37. Scheuer, C. -L., & Loughlin, C. (2021). Seizing the benefits of age diversity: could empowering leadership be the answer? Leadership & Organization Development Journal, 42(3), 495-515. https://doi.org/10.1108/LODJ-12-2019-0516

38. Sousa, I. C., Ramos, S., & Carvalho, H. (2019). Age-Diversity Practices and Retirement Preferences Among Older Workers: A Moderated Mediation Model of Work Engagement and Work Ability. Frontiers in Psychology, 10. https://doi.org/10.3389/fpsyg.2019.01937
39. Tabor, W., Chrisman, J. J., Madison, K., & Vardaman, J. M. (2018). Nonfamily Members in Family Firms: A Review and Future Research Agenda. Family Business Review, 31(1), 54-79. https://doi.org/10.1177/0894486517734683
40. Urbancová, H., Vnoučková, L., Linhart, Z., Ježková Petrů, G., Zuzák, R., Holečková, L., & Prostějovská, Z. (2020). Impact of Age Management on Sustainability in Czech Organisations. Sustainability, 12(3). https://doi.org/10.3390/su12031064
41. Urbancová, H. (2017). AGE management v organizacích: praktické využití a přínosy. Wolters Kluwer.
42. Wiktorowicz, J., Warwas, I., Turek, D., & Kuchciak, I. (2022). Does generativity matter? A meta-analysis on individual work outcomes. European Journal of Ageing, 19(4), 977-995. https://doi.org/10.1007/s10433-022-00727-w
43. Woodfield, P., Rasmussen, E., & Chandhok, Y. (2021). Family businesses and employment relations: Review and suggestions for future research. New Zealand Journal of Employment Relations, 46(1), 68-87. https://doi.org/10.24135/nzjer.v46i1.51
44. Yates, F. (1934) Contingency Tables Involving Small Numbers and the Chi-Squared Test. Supplement to the Journal of the Royal Statistical Society, 1, 217-235.

About the authors

Terezie Krestová is an Assistant Professor at the Department of Management of the Faculty of Economics, VSB – Technical University of Ostrava. Her scientific interest is in the field of human resources management; she specifically deals with the application of age management in organisations. She has participated in many student research projects at the VSB-Technical University of Ostrava focused on management. She can be reached at **terezie.krestova@vsb.cz**.

Aleš Kresta is an Associate Professor in the Department of Finance at the Faculty of Economics, VSB—Technical University of Ostrava, where he earned his Ph.D. in Finance. His primary research interests encompass risk estimation, backtesting, portfolio optimisation, financial time series modelling, and soft computing within the realm of quantitative finance. Actively involved in numerous scientific projects, he can be contacted at **ales.kresta@vsb.cz.**

Lucie Bestová is a doctoral student in the Department of Management of the Faculty of Economics, VSB Technical University of Ostrava. She focuses on the competencies of graduates. As a PhD student, she actively publishes in both scientific conferences and journals. She can be contacted at **lucie.bestova@vsb.cz**.

Appendix

Table I
p-values of the Chi-square Test for Age Management Practices

Age Management Practices	Factors		
	Economic Sector	Organisational Size	Family Bus. Status
Adaptation of Work Programmes to Fit Different Age Groups	0,0455	0,4975	0,7180
Reassignment of Staff to the More Suitable Position	0,0399	0,5906	0,5585
Care for Employee Health, Promotion of Physical Fitness, and Healthy Eating	0,0045	0,0000	0,1146
Support for Personal and Career Develop. through Skills Training at Every Stage of an Empl. Career	0,0312	0,0001	0,1607
Development of Health and Safety Practices	0,9441	0,0182	0,1721
Workforce Planning for Age Diversity and Promotion of a Positive Age Policy	0,6779	0,8886	0,1821
Creating Tools to Support Intergenerational Learning	0,0801	0,0002	0,0172
Motivational Programmes According to the Needs of Different Generations	0,2470	0,0039	0,2403
Tailored Further Training for Older Workers	1,0000	0,6418	1,0000
Use of Mentoring and Reverse Mentoring	0,0037	0,0493	0,1464
Special Forms of Recruitment for Different Age Groups	0,7135	0,9042	0,8709

Source: Authors' work

Table II
p-values of the Chi-square Test for Age Management Benefits

Age Management Benefits	Factors		
	Economic Sector	Organisational Size	Family Bus. Status
Retaining Key Staff	0,6875	0,9846	0,1771
Reducing Employee Turnover	0,6999	0,3344	0,2721
Increasing Employee Motivation	0,0327	0,0633	0,3446
Improving Org. Culture and Positively Impacting Employee Values	0,0813	0,0021	0,4193
Developing Employee Education, Knowledge, and Skills	0,0028	0,0003	0,0033
Improving Employee Performance	0,2378	0,1245	1,0000
Enhancing the Company's Reputation	0,1743	0,0192	0,0455
Enhancing Employee Health	0,1499	0,0611	0,1721
Acquiring New Talents	0,2902	0,0000	0,0410

Source: Authors' work

School-to-Work Transition in the Youth Labor Market in Central and Eastern Europe: A Cluster Analysis Approach

Tomislav Korotaj
University of Zagreb, Faculty of Economics & Business, Croatia
James Ming Chen
Michigan State University, College of Law, United States of America
Nataša Kurnoga
University of Zagreb, Faculty of Economics & Business, Croatia

Abstract

Background: This study analyzes education, training, and the youth labor market in central and eastern Europe. **Objectives:** This study aims to evaluate similarities and differences in youth labor markets among eleven central and eastern European countries from 2008 to 2021. It specifically examines three aspects: wage ratios, early departure from education or training, and the share of the population not in employment, education, or training. **Methods/Approach:** This study applies hierarchical clustering and multidimensional scaling to panel data. The complete-link method organizes countries into clusters. This study combines three-dimensional Cartesian projections and two-dimensional projections based on multidimensional scaling with dendrograms and heatmaps, to graphically illustrate the "school-to-work" transition across this region. **Results:** Clustering highlights the Visegrád countries, the Baltics, and the Balkans as zones with internally homogeneous yet externally heterogeneous challenges for the youth generation. As the outliers in each of these regions, Poland, Estonia, and Bulgaria support clustering solutions that deviate from conventional understandings of central and eastern Europe. **Conclusions:** Historical and geographical ties continue to define this region's youth labor markets across political and economic dimensions. Clustering analysis identifies triumphs and struggles in policymaking in some of the poorest and most politically challenging member-states of the European Union.

Keywords: hierarchical cluster analysis; complete-link method; time series; youth population; wage ratio; NEET; early departures from education; central and eastern Europe

JEL classification: C38, E24, I21, I28, J4
Paper type: Research article
Received: Apr 11, 2024
Accepted: Aug 23, 2024

Citation: Korotaj, T, Chen, J., M., & Kurnoga, N. (2024). School-to-Work Transition in the Youth Labor Market in Central and Eastern Europe: A Cluster Analysis Approach. *Business Systems Research, 15*(2), 100-139.
DOI: doi.org/10.2478/bsrj-2024-0020

Introduction

Today's youth population grapples with considerable societal pressures that exact varying expectations from different groups. Graduating from school marks a significant early milestone. Yet the journey does not end there. Some young people advance to higher levels of the educational system. Others directly enter the labour market. Another cohort finds itself in a challenging position: early dropouts from compulsory schooling. Education, an indispensable tool for enhancing the quality of life, may prove elusive for young people who lack motivation, family support, or equal opportunities. Navigating these challenges demands a determined collective response and a societal commitment to providing proper guidance and support to vulnerable young people.

International organizations (such as the International Labour Organization (ILO), Organization for Economic Co-Operation and Development (OECD)) and statistical offices (such as Eurostat) regularly report on the obstacles encountered by young Europeans. In 2021, the Council of the European Union (EU) set two crucial thresholds for educational progress (Council of the EU, 2021): (a) reducing the rate of early departure from education or training (EDET) to less than 9 percent and (b) increasing the share of tertiary-educated individuals aged 25-34 to at least 45 percent. Both targets apply across the entire EU, with a deadline of 2030. These targets build on the earlier Europe 2020 strategy (which covered the period 2010-2020) in which the thresholds (European Commission, 2010) were set at 10 percent for EDET and 40 percent for 30-to-34-year-olds with tertiary education.

This paper extends previous analysis of youth labor markets and the school-to-work transition in central and eastern Europe (Korotaj, Chen, and Kurnoga, 2023). In addition to emphasizing developments over the past decade, this paper shifts its focus beyond tertiary education and the cohort of youth wholly excluded from training, education, and employment. This new study examines the ratio of minimum to average wages and early departures from education or training alongside labor market failures. These three variables – the Kaitz ratio of minimum to average wages, the EDET factor revealing early departures from education or training, and the NEET factor identifying young people not participating in education, employment, or training – present a comprehensive summary of inputs and outcomes in the youth labor market, without presuming or relying on any specific causal mechanism connecting the variables.

This paper's motivation lies in uncovering the relative successes and failures of eastern and central European countries in addressing their youth labor markets. Numerous studies of youth labor markets have elaborated the relationship between educational shortcomings and eventual unemployment. Although causal relationships among educational inputs and labor market outcomes are latent in the data, this paper's reliance on *unsupervised* learning produces results that differ from the testable hypotheses of multiple regression and other forms of *supervised* learning (Valkenborg et al., 2023; Watson, 2023). "Unsupervised learning involves the use of training datasets without dependent variables. In this scenario, the goal is to explore the data and draw conclusions about, for instance, hidden patterns or structures that are present in the data" (Burzykowski et al., 2023, p. 733). "Unlike supervised learning, unsupervised learning methods cannot be directly applied to a regression or a classification problem …." (El Bouchefry & de Souza, 2020, p. 227). "[C]lustering algorithms," generally speaking, "do not test against the null hypothesis" (Watson, 2023, p. 27).

By "using datasets without clear notice of the dependent (response) variable," unsupervised learning directs a "machine or computer [to] learn patterns from the data without referring to any specific response" (Valkenborg et al., 2023, p. 877). "Unsupervised learning aims to explore the data structure and generate a hypothesis rather than to test any hypothesis by statistical methods or to construct prediction or classification models on the basis of a set of conditions and a specified response" (Valkenborg et al., 2023, p. 877). Ultimately, "unsupervised learning is ontologically fundamental," or at least "more so, at any rate, than supervised or reinforcement learning algorithms, which necessarily presume some a priori division of the world into endogenous inputs … and exogenous outputs" (Watson, 2023, p. 29).

Consistent with the "common" use of "unsupervised learning … to find hidden patterns or groupings in the data" (El Bouchefry & de Souza, 2020, p. 228), the clustering analysis in this paper aims to identify patterns in policymaking and policy success or failure among the nations of central and eastern Europe. Indeed, as a review of the relevant literature will demonstrate, clustering analysis has figured prominently in multiple comprehensive examinations of youth labor markets. Although clustering can serve as a form of preprocessing preceding null hypothesis significance testing through regression (Sakellariou et al., 2012), clustering as the primary form of unsupervised machine learning can stand as its own exclusive research method in a scientific study. One such study relying exclusively on clustering without naming, confirming, or invalidating a formal research hypothesis, has been published by this journal (Krpan et al., 2023).

The research hypothesis of this article relies upon unsupervised machine learning to achieve a taxonomy or ontology of countries according to mathematical relationships latent in economic data. In the sense of the ancient Greek word, ἀποκάλυψις (apokálupsis), as a "revelation" or "disclosure," clustering and manifold learning can produce mathematically informed, principled interpretations drawn directly from data. Specifically, this article evaluates similarities and differences in youth labor markets from 2008 through 2021 among eleven countries in central and eastern Europe. The three variables indicating the health of those labor markets, as young people complete the school-to-work transition, emphasize the ratio of minimum to society-wide average wages, the rate of early departures from education or training, and the total rate at which young people fail entirely to achieve employment, education, or training. Through the application of hierarchical clustering and multidimensional scaling to country-by-country time series capturing those three labor market variables, this article seeks to compare the performance of those eleven countries.

Conventional understandings of central and eastern Europe, informed by history that spans the Roman Empire through the world wars and intraregional conflicts of the twentieth and early twenty-first centuries, classify this distinctive and volatile part of Europe according to three geopolitical subregions: the Visegrád Group, the Baltic states, and the Balkan peninsula. "[A] page of history is worth a volume of logic" (Supreme Court of the United States, 1921, p. 349): From Diocletian's decision to split the Roman Empire to the Molotov-Ribbentrop Pact that precipitated World War II and the more recent politics of the European Union, these three subregions have played distinctive roles in parts of Europe beyond the Carolingian Empire and the "Inner Six" signatories of the Treaty of Rome (Bokros, 2013; Ghica, 2008; Nič, 2016). The identification of distinct clusters among central and eastern European nations then reveals relationships among wage ratios, education, and labor market outcomes. Those relationships follow distinct patterns in different subregions across central and eastern Europe. This paper ultimately illuminates relative successes and struggles among some of the poorest and most politically challenging member-states of the European Union.

Based upon results from clustering and manifold learning, this article tests whether this informal taxonomy can withstand rigorous mathematical evaluation. The extent to which clusters reported by unsupervised learning reinforce or depart from conventional understandings delivers an ontology that can guide the formulation of labor policies throughout this part of Europe, particular for countries that over- or underperform their immediate neighbors and historical counterparts.

After reviewing the relevant literature, this paper will provide a concise description of its data and research methodology. This study conducts hierarchical clustering analysis of youth labor market time series with the complete-link algorithm. Results will be presented through three-dimensional projections and their corresponding two-dimensional manifolds, dendrograms, and heatmaps, as well as verbal descriptions of plausible cluster solutions. The discussion section reviews the reasons that support possible groupings of central and eastern European countries. Concluding thoughts highlight the limitations of our research and suggest areas for further investigation.

Literature Review

In 1996, Hungarian law raised the age at which students could stop attending school from 16 to 18 years (Adamecz, 2023). Hungary adopted this measure to combat early departures from primary and secondary school. Despite increasing the required years of schooling, the reform did not reduce dropouts or improve employment outcomes for 20- and 25-year-olds. In schooling systems that force failing students to repeat a grade, obligatory education should aim to keep students in school until they obtain a secondary school diploma. Different regression discontinuity design models estimated the effects of extended compulsory schooling in Hungary.

Focusing on convergence among EU countries, Cuestas, Monfort, and Ordóñez (2021) comprehensively evaluated the Europe 2020 strategy. Their study unveils the presence of convergence clubs (Bernard & Durlauf, 1995, 1996) within the framework, albeit with varying paces of progress. Employing the Philips-Sul approach to analyze crucial educational variables emphasized in Europe 2020, the research identified three clubs for EDET and six for tertiary education.

These findings highlight a lack of convergence among EU countries. Notably, central and eastern European countries are dispersed across multiple clubs. Croatia stands out as the best EDET performer without club convergence. These results suggest that EU countries will probably fall short of achieving educational convergence within Europe 2020's anticipated timeframe.

Other sources have criticized Europe 2020. One multivariate analysis of non-stationary time series compared the EDET rates in Czechia and EU-28 countries between 2005 and 2018 (Blatná, 2020). Czechia performs significantly better than the EU-28. A mismatch in the integration order of the EDET indicator`s time series prevented the discovery of a meaningful statistical relationship between Czechia and the EU-28 countries. Nevertheless, the results for Czechia showed that higher social benefits and increased opportunities for well-paying jobs can potentially overcome incentives to leave education early.

An evaluation of the Scottish School Leavers Surveys (SSLS) delivered one of the first assessments of the NEET concept (Furlong, 2006). NEET describes persons who have "no employment, education, or training." The author investigated the reasons that young people leave school early and ultimately lack employment, education, and training. Many young people could not find a suitable job or course, reflected indecision about career choices, or lacked additional qualifications or skills for employment. Heterodox analysis of the extremely diverse backgrounds of young people is essential to the identification of subgroups among youth and the prescription of specific policies for each subgroup. A "one size fits all" approach promises little success in reducing either dropouts or unemployment. Moreover, forcing employment or training in any job against personal inclinations might do more harm than good.

Factors influencing the NEET population in Italy and Spain between 2007 and 2017 emerged in a time-varying correlation model that analyzed changing patterns in NEET and EDET rates (De Luca et al., 2020). Spain exhibited the highest EDET rates among EU countries, while Italy faced the highest NEET. Gender-specific patterns in Italy revealed that the vast majority of women progressed from EDET to NEET status.

A similar but less severe gender relationship was observed in Spain. Regression analysis confirmed a statistically significant influence of the EDET indicator on the NEET indicator in Italy but not in Spain. NEET rates depend on business cycles, unemployment, and the amount of time spent in education. Monitoring early departures from school can mitigate the NEET problem through remedial measures that address EDET. Policymakers can, therefore, tackle two challenges with a single intervention.

A multilevel logit model has investigated educational and socioeconomic reasons for early school departures in the EU (Lavrijsen & Nicaise, 2015). Data for persons aged 20 to 30 in 2009 showed a significant influence on parental background and financial situation. The children of poorly educated families and families with material deprivation experienced higher EDET rates. Vocational education has proved a suitable mechanism for keeping young people in school longer (at least until they finish high school). The results recommend educational reforms that should be implemented alongside improved social policies. Everything starts with equal opportunities, and a more equal society is a prerequisite for positive changes in education.

Panel analysis of the development of young people in Germany ten years after they left an apprenticeship without completing it revealed differences in work experience and wages of that cohort relative to peers who completed their apprenticeships (Patzina & Wydra-Somaggio, 2020). The timing of a departure from an apprenticeship is significant. Wage growth is more pronounced among individuals who drop out later (more than two years of training) than those who drop out in the middle (between one and two years of training) or early (up to one year of training) stages of training. Therefore, preventing early dropout must be a priority. Dialogue between employers, training companies, and political institutions is necessary for the success of apprenticeships.

Different reasons motivated dropouts to return to education in Spain (Portela Pruaño et al., 2022). In a qualitative study based on data from an education center in Ceuta, most respondents stated that the primary motivation for returning to education was inactivity, acquiring qualifications, and increasing their chances of finding a suitable job. The majority said they would leave the program if offered a job. The support of family, friends, and teachers enabled young people to return to and stay in education. Communal influence and a supportive environment can encourage young people to improve themselves through training. In cooperation with the government, training centers should provide clear financial barriers to education, training, and self-improvement.

EDET and NEET rates converged in EU regions from 2003 to 2015 (Rambla & Scandurra, 2021). This study combined data at the NUTS (Nomenclature des Unités Territoriales Statistiques) 1 and NUTS 2 levels. Geographic differences related to EDET decreased.

The picture is more complicated for NEET. Regional GDP appears to accelerate convergence among wealthier regions, but it also reinforces slower development for economically disadvantaged regions. No significant convergence effects were observed in post-socialist EU countries. Between 2007 and 2016, the NEET gap decreased in post-socialist countries as regional GDP increased. Still, the mitigation of regional disparities was not as significant as the EU's average reduction in the NEET rate.

In their analysis of the impact of minimum wage increases on youth employment in various EU NUTS 2 regions, Vodă, Bercu, and Sebestova (2021) used panel data analysis for the period 2008-2014. They found a negative link between relative minimum wage as measured by the Kaitz index, which is the ratio of the minimum to the average wage (Brown et al., 1982; Kaitz, 1970), and youth employment. The employment of young individuals aged 15 through 24 decreased as a result of a higher minimum wage.

Minimum wages affect youth employment in the Visegrád Group countries (Fialová & Mysíková, 2020). According to this analysis of panel data from 2003 to 2016, youth employment declined in Hungary from 2008 to 2011 and in Czechia from 2003 to 2007. However, the changes in relative minimum wage had an overall non-negative impact across all the observed countries.

A closer examination of Albania and North Macedonia uncovered not only persistent youth unemployment but also a subpopulation of young people who are resistant or unresponsive to education (Mehmetaj & Zulfiu Alili, 2020). These social pathologies appear peculiarly characteristic of these countries in ways that differ from the rest of the Balkan sub-region, to say nothing of the remainder of central and eastern Europe.

Multiple works have applied clustering analysis to the school-to-work transition in central and eastern Europe. One application of hierarchical clustering identified distinct school-to-work transition subregimes within this post-socialist region (Dingeldey & Buttler, 2023). The variables in this study described the broader economic backdrop, legal protections for employment, education and vocational training, and labor market policies across 28 European countries and Israel.

Hierarchical clustering has also illuminated the relationship between tertiary and labor market outcomes across the European Union (Krpan et al., 2023). This study distinguished sharply between tertiary education as a labor market input and a range of indicators for labor market outputs. The authors found an imperfect match between countries that attained the highest shares of adults with tertiary education, on one hand, and countries that realized the best outcomes, such as the highest average employment and income benefits arising from tertiary education. Limitations on data availability constrained the ability to draw stronger conclusions about the impact of tertiary education on labor market outcomes.

One study has combined clustering analysis with conventional linear regression of educational, labor-related, and developmental factors across Europe. Tudor et al. (2023) made broader use of variables such as gross domestic product and societal investment in education. These authors found three distinct clusters of EU countries with low, average, and high levels of educational investment. Although these authors did find that a higher rate of educational dropouts diminished compensation, hours worked, and productivity, the impact of education on wages and productivity does differ according to country-wide levels of investment.

Like Dingeldey and Buttler (2023), Krpan et al. (2023), and Tudor et al. (2023), this study contributes to the academic understanding of the similarities and differences in the experiences of young people as they navigate schooling, training, and employment. This study shares Dingeldey and Buttler's (2023) focus on central and eastern Europe. Unlike Krpan, Gardijan Kedžo, and Žmuk (2023), who confined their clustering analysis to exactly two years (2012 and 2021), this study has accounted for the effects of changes in policy and labor conditions across 14 years (2008 through 2021 inclusive).

At the methodological level, this study aligns more closely with Dingeldey and Buttler (2023) and with Krpan, Gardijan Kedžo, and Žmuk (2023) in emphasizing clustering and unsupervised learning as standalone methods of economic analysis that allow data to reveal insights without the stipulation of a formal research hypothesis. Tudor et al. (2023) did deploy clustering as a step in preprocessing *en route* to a more explicit exercise in causal inference. Whereas Tudor et al. consciously conducted linear regression, studies relying exclusively on unsupervised learning (including this one) are more cogently understood as indirect exercises in classification. This article extends the use of unsupervised learning as an indirect classification technique, routinely applied to images and computer vision (Olaode et al., 2014; Schmarje et al., 2021), to economics and other social sciences. Clustering analyses strive to identify differences in policies and performance across jurisdictional boundaries.

In other words, sorting the countries of central and eastern Europe into a cogent geopolitical ontology according to Euclidean distances in standardized time series representing the Kaitz wage ratio, the EDET rate, and the NEET rate is this study's research objective. Ontological or taxonomic outcomes identify possible differences in socioeconomic conditions and/or public policy among these "newer" members of the European Union. Studies relying exclusively on unsupervised machine learning – whether in economics, cognate social sciences, or seemingly remote domains such as medicine or computer vision – demonstrate that the identification of differences in economic performance and/or political choices is a legitimate and worthy research objective, even in the absence of answers to questions more conventionally addressed through generalized linear regression methods or other forms of supervised learning.

Like their counterparts in natural science and visual representations of physical reality, economic and political patterns can be detected by clustering. In light of this broader region's shared experience with socialism and (in some instances) Soviet domination after the Second World War, the discovery of mathematically meaningful distinctions in performance and policy adds to the academic understanding of youth labor markets across the region. This is especially true if those distinctions reinforce or modify the conventional, qualitative classification of central and eastern Europe into three distinct subregions known as the Visegrád Group, the Baltic states, and the Balkan peninsula.

This study's taxonomy of the post-socialist member-states of the European Union also differs from the conclusions of Dingeldey and Buttler (2023) in key respects. Although those authors place "[a]ll post-communist states [within] one cluster with the exception of Hungary" (*ibid.*, p. 167), their "results do not conform to the popular grouping" that recognizes the "Baltic states and the Visegrád Group … as homogeneous clusters" (*ibid.*, p. 168).

By contrast, the analysis in this study finds that conventional geopolitical boundaries do define (or come close to defining) cogent clusters within central and eastern Europe. Dingeldey and Buttler ultimately treat Poland and Czechia, Latvia, and Bulgaria as representative countries that serve as proxies for the three large subregions of central and eastern Europe: the Visegrád Group, the Baltics, and the Balkans "as a southern group of … countries with less developed economies" (*ibid*., p. 169). This study aspires to inform policymaking in the European Union by facilitating fruitful comparisons across national boundaries within the EU (Krpan et al., 2023, p. 209).

Methodology

Analytical Blueprint

A comprehensive analytical framework to youth labor markets has been applied, aiming to uncover distinct patterns based on wage ratios, early departures from education or training, and rates of nonparticipation in employment, education, or training, or NEET. The complete-link algorithm and Euclidean distances provide a robust mechanism for grouping data points based on their maximum distances. Heatmaps and dendrograms will facilitate the interpretation of clusters within the data.

Data were drawn from Eurostat databases and covered an observation period from 2008 to 2021. Standardization of the data ensures comparability and the accurate representation of social desirability in all variables. Three-dimensional (3D) projections, alongside dendrograms and heatmaps, provide a dynamic visualization of changes in labor markets. These results provide a nuanced understanding of progress and setbacks in youth labor markets across central and eastern Europe.

Data

This study's sample comprises 11 countries: Bulgaria (BG), Croatia (HR), Czechia (CZ), Estonia (EE), Hungary (HU), Latvia (LV), Lithuania (LT), Poland (PL), Romania (RO), Slovakia (SK) and Slovenia (SI). This study investigates the youth labor market across 14 years, from 2008 to 2021 inclusive. Data preprocessing and cluster analysis, as conducted in Python, aligned all three variables in the study in the same direction by reversing the NEET and EDET rate, and expressing them as "$100 - NEET$" and "$100 - EDET$", respectively. This transformation expresses rising values in each variable as socially beneficial. Specifically, higher values for all three indicators are preferable.

NEET is reported as the percentage of young individuals engaged in some form of employment, education or training. Likewise, EDET is reported as the percentage of young people with at least a high school diploma. It represents the group that stays in compulsory education or training for the full duration. Key indicators for this research are briefly explained below.

The Kaitz index, which is the ratio of minimum to average wages, is a crucial indicator of income inequality and poverty. "The Kaitz index has the advantage of summarizing a great deal of information about the minimum wage law in a single variable" (Brown et al., 1982, p. 500). The Kaitz index is often used as a relative measure of the minimum wage within national labor markets (Arpaia et al., 2017; Askenazy, 2003; Lenhart, 2017; O'Higgins & Moscariello, 2017; Williams & Mills, 1998).

The NEET rate refers to the percentage of individuals aged 25-34 who are not engaged in employment, education or training. It can be further divided into NEET rate for youth who are unemployed and NEET rate for youth who are inactive. Eurostat provides this metric for various age ranges, including 15-17, 15-19, 15-24, 15-29, 15-34, 18-24, 20-24, 20-34, and 25-29. This study used an inverted version of the NEET rate so that higher values indicate a positive social impact.

The EDET rate indicates young individuals aged 18-24 who have left compulsory school or training early. It is a group of youth which have completed primary school (lower secondary school according to Eurostat), but dropped out from secondary school (upper secondary school according to Eurostat) or were not involved in any kind of further education or training. As with NEET, this study used the inverted version of the EDET rate so that higher values reflect positive social impact.

Additional Methodological Considerations

This overview of this study's dataset warrants a few additional methodological observations. This dataset contains 14 annual observations for 11 countries. Consequently, n = 154. Among the three variables in the design matrix – the Kaitz wage index and modified versions of the EDET and NEET rates – NEET may be regarded as an implied target variable. By contrast with the Kaitz ratio as an indicator of background market conditions and EDET as a gauge of education and other labor market inputs, NEET measures labor market outcomes. If all three variables are treated equally, the cardinality of the dataset is quite low: p = 3. Treating NEET as the implied target variable reduces p even further to a value of 2.

Studies with deeper data, such as Tudor et al. (2023) and Vasilescu, Stănilă, Popescu, Militaru, and Marin (2024), have successfully deployed clustering analysis as a preprocessing step in anticipation of conventional linear regression. This study's low values *of n and p* motivate a simpler approach, one that relies exclusively on clustering. Unsupervised learning, unaccompanied by linear regression or any other apparatus for supervised learning, can still provide useful insights where both n and p are low. A dataset so small that it might struggle to support properly supervised learning, especially according to best practices that would split the data into training and test set instances and thereby reduce N for training purposes even further (Müller & Guido, 2017, pp. 17-18), may still offer useful information through clustering.

Unlike many other exercises in clustering and other applications of machine learning, this study does not suffer from a surfeit of variables or the curse of dimensionality (Evangelista et al., 2006; Marimont & Shapiro, 1979). The opposite is true: Scarcity of information forecloses resort to familiar techniques for reducing dimensionality. In order to preserve the richness of the information in the dataset, this study bypasses principal component analysis (PCA) and proceeds directly to cluster analysis. Although PCA can help simplify data, it can also lead to information loss (Jolliffe, 2002; Geiger & Kubin, 2012). Instead, this study engages in a direct evaluation of hierarchical clustering results and highlights the impact of political and economic history on the composition of clusters.

Clustering and Visualization Methods

The complete-linkage method calculates the distance between two objects, a and b, in two different clusters, A and B, according to this formula (Bezdek, 2021):

$$\boldsymbol{\delta_{CL}(A,B) = \max_{i\epsilon A, j\epsilon B}\{d_{ij}\}} \qquad (1)$$

Complete linkage uses the maximum distance between any single data point in the first cluster (A) and any single data point in the second cluster (B) to determine the distance between two clusters. This method tends to produce more compact clusters with greater similarity within clusters, making it suitable for identifying tightly knit groups within a dataset.

Hierarchical clustering can also use average linkage (Nielsen, 2016, p. 225; Xu et al., 2021, pp. 6010-6011):

$$\boldsymbol{\delta_{AL}(A,B) = \frac{1}{|A||B|}\sum_{i\in A}\sum_{j\in B}\{d_{ij}\}} \qquad (2)$$

The average-linkage method is embedded by default in heatmaps produced by the Seaborn graphical package for Python. Average linkage, therefore, provides an alternative approach to the complete-linkage method of hierarchical clustering.

Euclidean distance measures the shortest path between two data points in Euclidean space (Hair, 2018), serving as a standard metric for assessing the similarity between data points. Within an *N*-dimensional space, the distance between two objects, a and b, is computed based on the square root of the sum of the squared differences between corresponding coordinates of the two data points in Euclidean space:

$$\boldsymbol{d_{ij} = \|a - b\|_2 = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}} \qquad (3)$$

This research uses Euclidean distance for hierarchical clustering, primarily due to its ability to produce closely knit cluster solutions, its efficiency in handling clusters of varying sizes, and its computational speed. At the same time, this approach is limited by the irreversible nature of cluster agglomeration, reduced suitability for large datasets, and a pronounced vulnerability to outliers. The selection of Euclidean distance is recommended for lower-dimensional datasets, where the "curse of dimensionality" is diminished (Aggarwal, 2001; Domingos, 2012; Kriegel et al., 2009). This study's focus on a three-dimensional dataset encompassing three economic variables further substantiates the choice of the Euclidean metric.

In a three-dimensional Euclidean space, a Cartesian projection specifies a unique point for every triplet of numbers (*x*, *y*, *z*). This mapping allows for the precise positioning and visualization of data within a three-dimensional framework. 3D projections visually represent multidimensional data along three-dimensional Cartesian coordinates. They indicate the spatial distribution of clusters, revealing patterns and possible outliers. Moreover, 3D plotting makes complex relationships more visually accessible. Interpretation relies upon the identification of groupings and Euclidean distances between data points in three-dimensional space. This approach, therefore, enables a visual understanding of multidimensional data. The weaknesses of this approach include potential overlap and occlusion, which may make some data points or relationships less visible.

Two-dimensional representations of these same spaces accompany the 3D plots. Dimensionality reduction through multidimensional scaling (MDS) enables the visualization of the data in this study in two dimensions (Cox & Cox, 2008; Hout et al., 2013). A 2D projection based on multidimensional scaling replaces all underlying variables. It reduces them to two arbitrary dimensions, which in turn can be understood and visualized as the *x*- and *y*-axes of a conventional Cartesian projection. Although the application of MDS results in the loss of data and the application of a numerical scale not directly linked to the underlying data, MDS does display relationships in a more readily interpreted 2D format.

Dendrograms visually represent the arrangement of clusters formed by hierarchical clustering. They indicate the sequence of cluster mergers and the distance at which clusters merge. The tree-like diagram reveals cluster hierarchy and similarity. Dendrograms often inform decisions about possible and final cluster solutions. Using dendrograms in cluster analysis offers deep insights into relationships and hierarchies within clusters, improving the understanding of data grouping. However, the complexity of analyzing large datasets through dendrograms can be challenging as numerous branches can complicate the determination of a cluster solution.

Heatmaps display data through variations in coloring. They utilize color intensity to signify the magnitude of values and illustrate the level of observation similarity within a dataset. Interpretation involves assessing color gradients to identify patterns, correlations, or anomalies. Darker colors typically represent lower values or intensities, while brighter colors reflect higher values or intensities. Changes in heatmap colors reflect standardized Euclidean distances between observations. The advantages of heatmaps include a straightforward visual comparison of large data matrices. Nevertheless, challenges may arise in distinguishing subtle shades, potentially complicating the analysis of closely related values.

These graphical tools collectively enhance the understanding of clustering results. Each tool offers unique insights and interpretive considerations. The next section reports clustering results for eleven central and eastern European countries from 2008 to 2021.

Results

Descriptive analysis

The following tables, 1 through 3, present descriptive statistics for the three variables underlying this analysis – the Kaitz index (*wages*), (one hundred minus) the rate of early departures from education or training (*edet*), and (one hundred minus) the rate of young persons not in employment, education, or training (*neet*). These means, standard deviations (SD), and five-quantile statistics (minimum and maximum values and the 25th, 50th, and 75th percentiles) address all years covered from 2008 to 2021.

Table 1

Descriptive statistics for *wages*, the Kaitz index, by country

	Mean	SD	Min	25%	50%	75%	Max
Bulgaria	40.536	2.635	35.500	39.125	40.800	43.100	43.300
Czechia	35.636	3.265	31.600	32.925	34.700	37.925	41.400
Estonia	37.093	2.925	33.000	34.975	36.800	38.450	42.600
Croatia	40.671	3.233	37.000	37.925	40.200	42.250	46.600
Latvia	42.643	2.405	37.400	41.600	42.500	44.550	46.400
Lithuania	45.393	2.900	40.200	42.850	46.400	46.875	50.600
Hungary	41.593	2.284	38.000	39.225	42.450	43.300	44.500
Poland	45.000	3.161	39.100	42.525	45.400	46.250	50.500
Romania	40.429	6.490	31.300	35.025	39.500	47.300	48.400
Slovenia	51.171	3.323	43.400	51.400	51.950	52.725	55.200
Slovakia	37.850	3.193	33.600	35.775	36.500	39.400	44.400

Note: SD = Standard Deviation; Min = Minimum, 25% = First Quartile; 50% = Median; 75% = Third Quartile; Max = Maximum; *Source:* Authors' work

Table 2

Descriptive statistics for *edet*, or (one hundred minus) the EDET rate; by country

	Mean	SD	Min	25%	50%	75%	Max
Bulgaria	86.907	0.903	85.200	86.300	87.250	87.475	88.200
Czechia	94.029	0.781	92.400	93.450	94.100	94.575	95.100
Estonia	88.607	1.599	86.000	88.000	88.700	89.625	91.500
Croatia	96.300	1.132	94.800	95.125	96.800	97.200	97.800
Latvia	89.771	2.531	84.500	88.650	90.150	91.475	92.800
Lithuania	93.900	1.358	91.300	92.825	94.250	94.675	96.000
Hungary	88.186	0.475	87.500	87.925	88.200	88.475	89.200
Poland	94.657	0.298	94.100	94.450	94.650	94.800	95.200
Romania	82.757	1.385	80.700	81.900	82.450	83.975	84.700
Slovenia	95.536	0.580	94.700	95.025	95.600	95.800	96.900
Slovakia	93.214	1.460	90.700	92.250	93.200	94.525	95.300

Note: SD = Standard Deviation; Min = Minimum, 25% = First Quartile; 50% = Median; 75% = Third Quartile; Max = Maximum; *Source:* Authors' work

Table 3
Descriptive statistics for *neet*, or (one hundred minus) the NEET rate; by country

	Mean	SD	Min	25%	50%	75%	Max
Bulgaria	75.457	3.441	70.300	72.300	75.850	78.100	80.400
Czechia	80.393	1.406	78.700	79.525	80.000	81.375	83.200
Estonia	82.007	3.466	75.600	80.050	82.550	84.650	87.000
Croatia	78.586	3.373	73.700	76.050	78.750	81.225	83.600
Latvia	80.750	3.596	74.000	79.075	82.100	83.325	85.200
Lithuania	82.950	3.543	75.600	80.625	83.400	86.025	87.400
Hungary	78.071	4.175	72.500	74.150	78.100	81.200	86.700
Poland	81.014	1.739	78.400	79.600	80.800	82.425	84.300
Romania	78.486	2.245	75.600	76.725	78.050	79.400	83.600
Slovenia	87.643	2.402	83.300	85.975	88.400	89.325	91.700
Slovakia	75.779	2.905	71.800	73.200	76.150	77.600	82.100

Note: SD = Standard Deviation; Min = Minimum, 25% = First Quartile; 50% = Median; 75% = Third Quartile; Max = Maximum; *Source:* Authors' work

It may reveal even more when examining all three variables by year rather than by country. Tables 4, 5, and 6 examine, respectively, the Kaitz wage ratio (*wages*), (one hundred minus) the EDET rate of early departures from education or training (*edet*), and (one hundred minus) the NEET rate of young people not in employment, education, or training (*neet*), all arranged by the years 2008 through 2021 rather than by country. Table 4 reveals that the *wage variable generally increased* from 2008 to 2021. This suggests a region-wide trend toward greater income equality.

Table 4
Descriptive statistics for *wages*, the Kaitz wage ratio, by year

Year	Mean	SD	Min	25%	50%	75%	Max
2008	37.436	3.509	31.300	34.900	37.600	39.650	43.400
2009	38.791	3.345	34.300	36.050	38.300	41.750	44.200
2010	38.864	5.265	32.400	35.750	38.000	41.950	50.500
2011	39.073	5.701	32.400	35.450	37.500	42.150	51.700
2012	39.345	6.044	31.600	34.850	37.900	42.950	52.200
2013	40.745	6.284	32.600	36.150	39.200	43.950	53.200
2014	41.255	5.805	32.800	37.200	40.500	44.700	52.800
2015	42.373	5.243	34.400	38.850	41.600	45.950	52.400
2016	43.400	5.060	35.500	40.100	43.400	45.800	51.700
2017	43.973	4.224	37.100	41.900	43.200	47.050	51.300
2018	43.991	3.959	38.200	41.100	43.700	46.250	51.700
2019	43.736	4.116	38.500	41.150	43.200	46.300	52.500
2020	45.200	4.287	39.500	42.200	43.600	47.400	53.600
2021	44.745	5.013	37.900	41.400	44.400	47.450	55.200

Note: SD = Standard Deviation; Min = Minimum, 25% = First Quartile; 50% = Median; 75% = Third Quartile; Max = Maximum; *Source:* Authors' work

The *edet* variable changes very modestly over this period (Table 5). Year-by-year differences in the average EDET rate across central and eastern Europe were no greater than 1.636 percent. Indeed, the minimum annual value in region-wide average *edet* was observed in 2008, while the maximum value was observed in 2021. Because all variables have been oriented so that higher values are socially preferable, we see that early departures became less troublesome, even if only marginally, from 2008 to 2021.

Table 5
Descriptive statistics for *edet*, or (one hundred minus) the EDET rate; by year

Year	Mean	SD	Min	25%	50%	75%	Max
2008	90.409	4.766	84.100	85.600	92.500	94.650	95.600
2009	90.418	4.613	83.400	86.100	91.300	94.700	95.100
2010	90.936	4.705	80.700	88.200	92.100	94.900	95.300
2011	91.300	4.358	81.900	88.500	92.600	94.950	95.800
2012	91.318	4.244	82.200	88.800	93.500	94.600	95.600
2013	91.518	4.154	82.700	89.150	93.600	94.500	96.100
2014	91.491	4.584	81.900	88.300	93.300	94.550	97.200
2015	90.964	4.936	80.900	87.500	93.100	94.600	97.200
2016	91.109	4.766	81.500	88.100	92.600	94.950	97.200
2017	91.136	4.558	81.900	87.850	91.400	94.800	96.900
2018	91.491	4.337	83.600	87.750	91.700	95.300	96.700
2019	91.573	4.165	84.700	88.500	91.700	95.100	97.000
2020	91.936	4.002	84.400	89.700	92.400	94.500	97.800
2021	**92.045**	**4.004**	**84.700**	**89.100**	**92.700**	**94.400**	97.600

Note: SD = Standard Deviation; Min = Minimum, 25% = First Quartile; 50% = Median; 75% = Third Quartile; Max = Maximum; *Source:* Authors' work

Finally, the central and eastern European region experienced higher levels of young people with bad labor market outcomes after the global financial crisis from 2009 onward. This trend is reflected in Table 6's year-by-year summary of the *neet* variable. Full regional recovery in the NEET rate did not occur until 2019. Variability among countries tended to decline throughout the recovery – until the COVID-19 pandemic struck in 2020.

Table 6
Descriptive statistics for *neet*, or (one hundred minus) the NEET rate; by year

Year	Mean	SD	Min	25%	50%	75%	Max
2008	82.100	4.040	75.800	80.550	82.000	83.600	91.700
2009	79.027	4.558	73.800	75.900	78.400	81.200	89.400
2010	77.291	4.507	72.500	74.200	75.600	79.200	88.400
2011	77.545	4.576	71.200	74.750	77.700	79.200	88.400
2012	77.418	4.241	71.500	73.800	78.800	79.350	86.200
2013	77.709	4.467	70.300	74.450	78.400	81.150	83.900
2014	78.355	3.982	71.800	76.450	79.600	81.050	83.300
2015	79.736	3.866	74.400	76.450	79.600	83.150	85.400
2016	80.236	3.899	74.000	76.850	80.800	82.500	85.900
2017	81.827	3.735	76.600	78.200	82.200	84.300	87.500
2018	82.636	3.526	77.700	80.250	82.700	84.750	88.400
2019	83.018	3.267	77.900	81.000	82.500	85.100	89.500
2020	81.491	3.481	77.300	79.400	80.400	83.000	89.100
2021	83.055	4.229	75.600	80.500	82.400	86.450	89.900

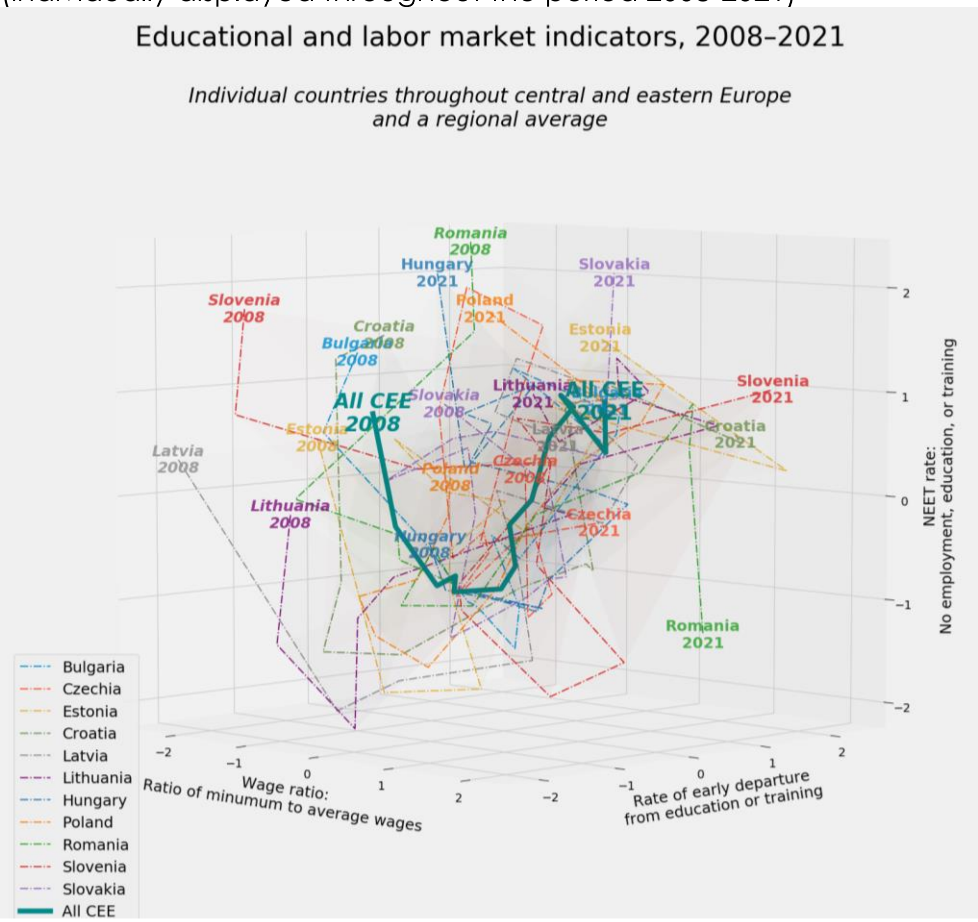
Note: SD = Standard Deviation; Min = Minimum, 25% = First Quartile; 50% = Median; 75% = Third Quartile; Max = Maximum; *Source:* Authors' work

The contrast between geographic and temporal overviews of the data also warrants commentary. Tables 1 through 3 organize the variables by country. Orderly rearrangement of the data is neither intuitive nor immediately evident. By contrast, temporal ordering by years in Tables 4 through 6 flows naturally. Even without elaborate computational intervention, some insights and patterns over time are visible to the naked eye. The clustering analysis that follows this exploratory description of the data reflects both geographic and temporal elements in the data. Above all, it must be remembered that all three variables are structured as *time series*. These series reflect autocorrelation within individual countries and geopolitical effects across borders. Therefore, clustering analysis should evaluate a continuous cross-section of data by multiple years, if at all possible, rather than individual years in isolation.

Three-dimensional projections

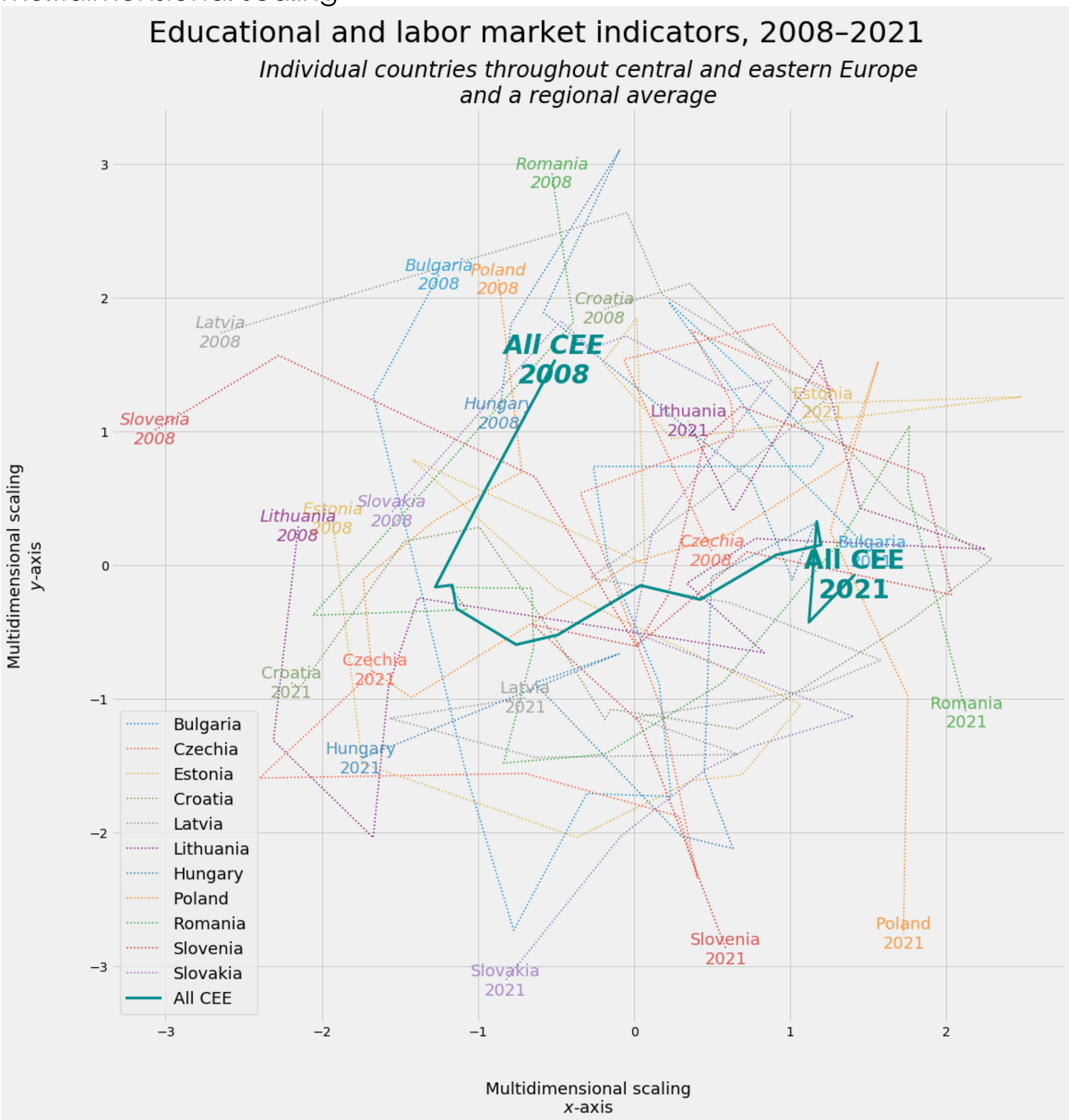
Analysis of the youth labor market begins with 3D projections. In the initial step, the standardization of raw values created a 3D projection of each country's performance in all three educational and labor market predictors (Figure 1). The average of all Cartesian coordinates representing these countries' Kaitz wage ratios, EDET rates, and NEET rates produces a global centroid, which is designated as the "All CEE" time series in Figure 1. The corresponding 2D manifold, enabled by multidimensional scaling, appears immediately afterward as Figure 1A.

Figure 1
3D projection of combined indicators for central and eastern European countries (individually displayed throughout the period 2008-2021)



Source: Authors' illustration

Figure 1A
2D projection of combined indicators for central and eastern European countries (individually displayed throughout the period 2008-2021), obtained with multidimensional scaling



Source: Authors' illustration

In Figure 1, standardized values are presented along three axes: (i) *wages:* The Kaitz wage ratio on the x-axis, (ii) *edet*: (One hundred minus) the EDET rate on the *y*-axis, and (iii) *neet*: (One hundred minus) the NEET rate on the z-axis.

These three variables do not appear in Figure 1A's MDS-enabled 2D manifold that encodes and visualizes the same information. Instead, the two-dimensional MDS manifold reports all three axes in Figure 1A – standardized values for *wages*, *edet*, and *neet* – as two arbitrarily designated x- and *y*-axes. As mathematical "translations" of the fuller 3D visualizations, this article's 2D manifolds are not intended to add analytical insights beyond the information contained in all three dimensions in the data. Rather, the 2D manifolds present clustering results in a way that is less visually demanding, and therefore perhaps more intuitively understandable.

Although this study's exclusive focus on clustering and unsupervised learning does not, strictly speaking, support distinctions between independent and dependent variables, let alone a search for causal inferences among variables, the NEET rate occupies the position on the z-axis that would ordinarily be reserved for a true target variable. To the extent that causal mechanisms could be hypothesized, sources such as Tudor et al. (2023) treat ultimate labor market outcomes such as employment, wages, and productivity as functions of macroeconomic and educational inputs. The Kaitz wage ratio measures the market for unskilled labor, while EDET is a surrogate for the overall educational attainment of the workforce. NEET, therefore, presents the strongest case for treatment as a target variable.

Values along each axis range roughly from –3 to +3, as should be expected of standardized data.

Countries moving toward higher values on all axes across this time period, such as Slovenia, indicate a favorable labor market with higher wage ratios and improved NEET and EDET conditions. Conversely, countries such as Bulgaria move toward lower values. These countries face suggesting challenges such as declining wage ratios or worsening NEET and EDET rates. With trajectories crossing the midsection of this 3D projection, Poland and Slovakia appear to have achieved moderate performance amid fluctuating labor markets.

In all three-dimensional plots, the upper right corner, representing higher values on all axes, is more desirable. Positive values on all axes indicate a more favorable youth labor market, with higher wage ratios and lower rates of NEET and EDET. The best performers exhibit upward trends across all axes, while the worst decline or stagnate. Moderate performers might show improvement in some areas, but not uniformly. Trends are influenced by economic policies, educational reforms, and broader socioeconomic factors. For instance, an improving trend could result from effective youth employment policies or educational programs reducing early departures from school or training. Stagnation might reflect persistent economic challenges or inadequate policy responses. More details follow the presentation of clustering results.

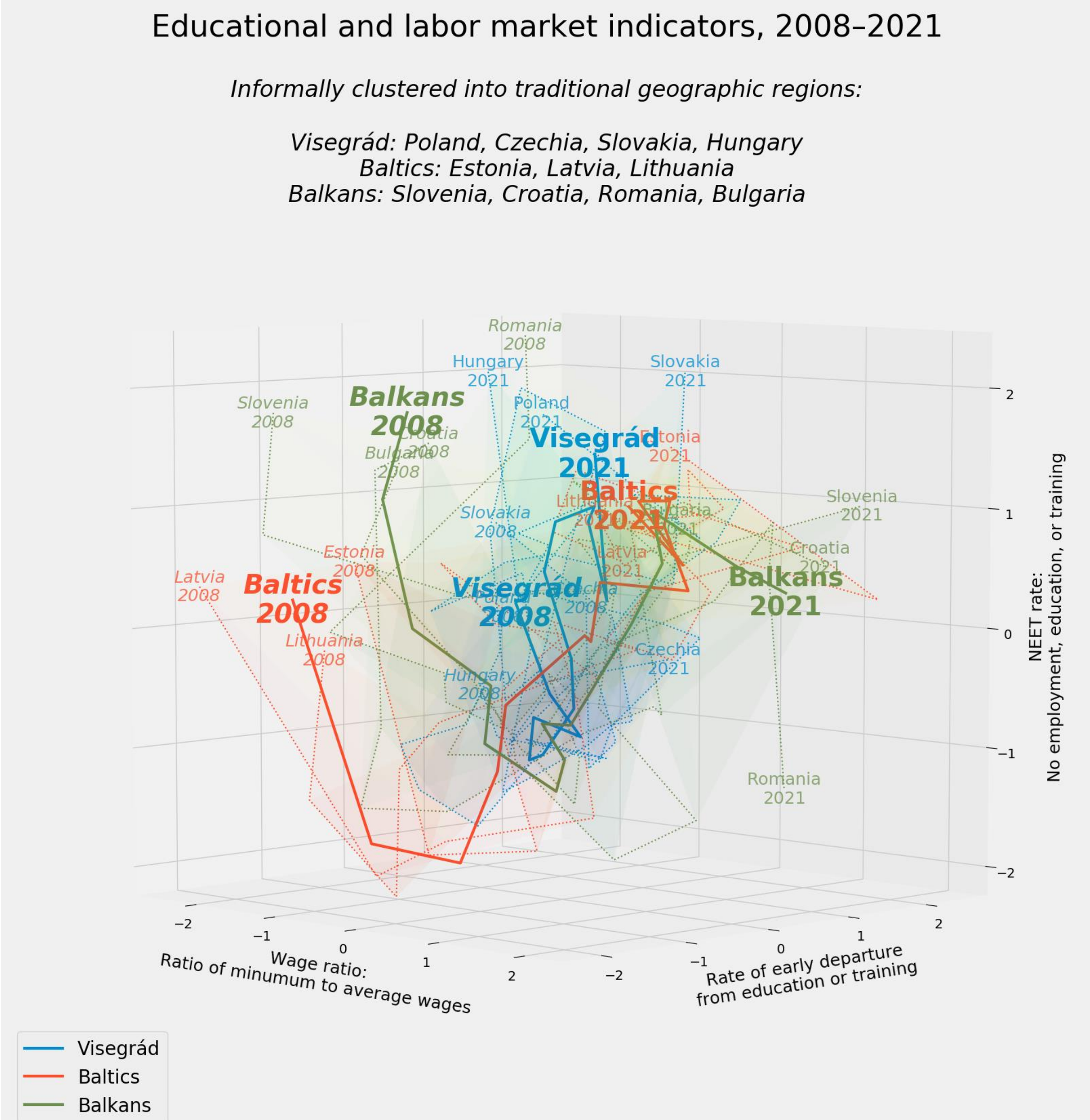
As in connection with Figure 2, an MDS manifold presents the same information in two dimensions (Figure 2A).

The second 3D projection (Figure 2) reports informal groupings of countries based on labor market variables – and, critically, on expert human judgment rather than strictly machine-derived definitions of clusters within central and eastern Europe. As in Figure 1, mean values for countries within a grouping produce a centroid in three-dimensional Cartesian space.

The Visegrád countries (Czechia, Hungary, Poland, Slovakia) are closely clustered, potentially indicating similar economic policies or labor market conditions influenced by their shared history and regional cooperation. The Baltics (Estonia, Latvia, Lithuania) also appear closely grouped, which could reflect their geographic proximity and comparable post-Soviet economic transitions. The Balkans (Bulgaria, Croatia, Romania, Slovenia) display a more dispersed pattern, perhaps due to varied economic conditions and differences in their levels of EU integration. As implied by the English word *balkanization*, this large and mountainous peninsula is one of Europe's most diverse regions.

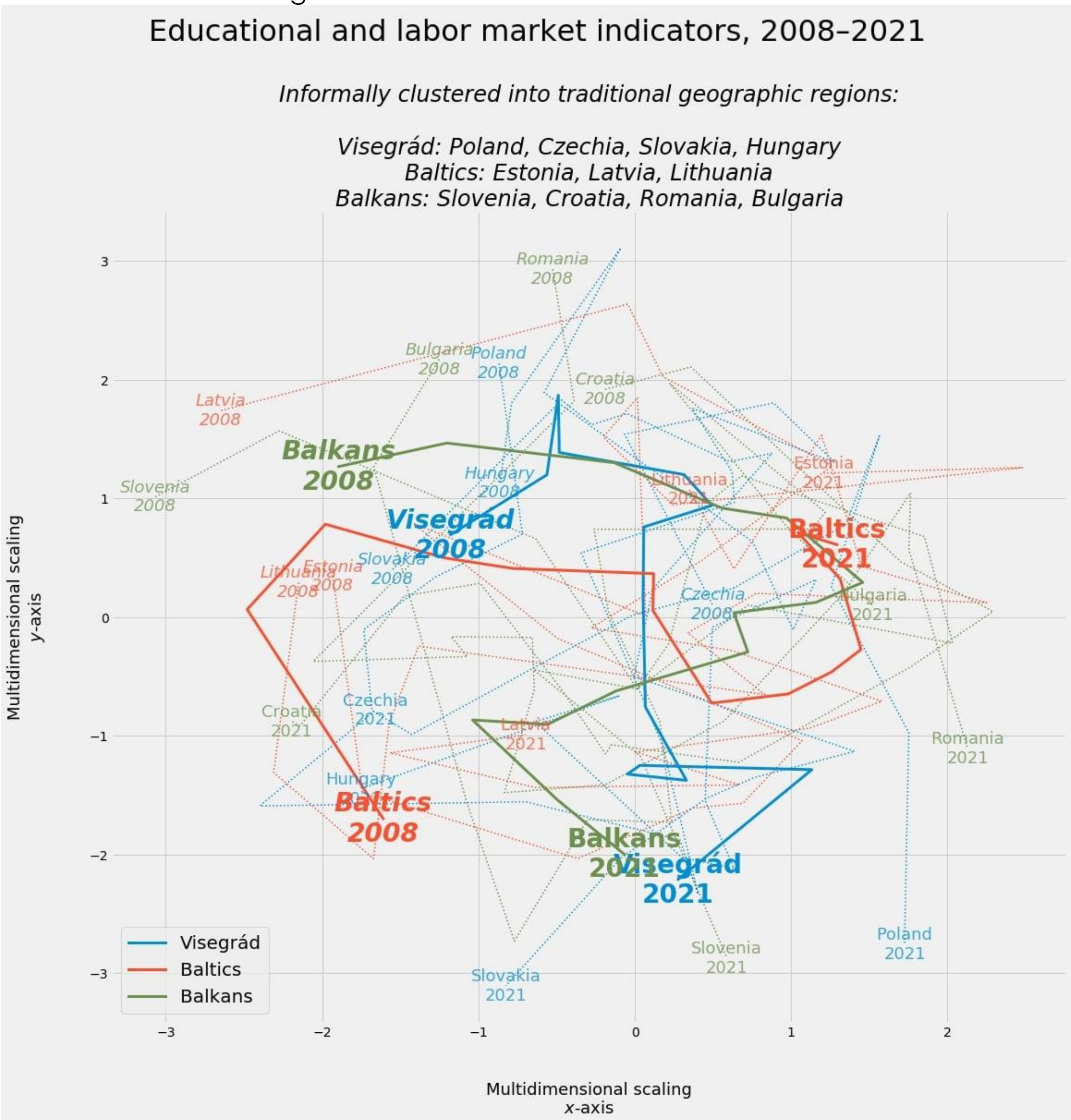
These informal clusters may reveal underlying historical, geographical, and political affinities, as well as shared economic experiences or reforms that influence labor market outcomes (Kunić, 2022). It will be instructive to compare qualitative or even intuitive human judgment against results drawn strictly from unsupervised machine learning.

Figure 2
3D projection of combined standardized indicators for central and eastern European countries (informally clustered throughout the period 2008-2021)



Source: Authors' illustration

Figure 2A
2D projection of combined standardized indicators for central and eastern European countries (informally clustered throughout the period 2008-2021) obtained with multidimensional scaling

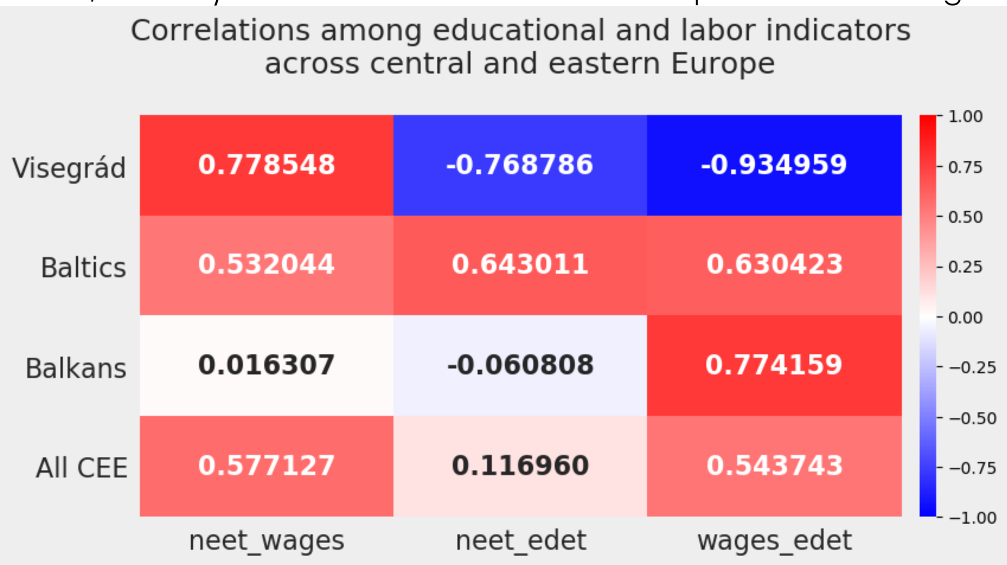


Source: Authors' illustration

A table of correlations for each informal geographic cluster shows how the three variables (the Kaitz wage ratio [or *wages*], EDET, and NEET) have starkly different relationships to one another by region (Figure 2B). The correlations in this table, and in other correlation tables throughout this article, cover all years from 2008 to 2021 inclusive. In principle, a favorable ratio of minimum to average wages, minimizing early departures from education and training, and avoidance of catastrophic NEET outcomes among young people should be mutually reinforcing. But a virtuous cycle of high wages for unskilled labor, persistence in education, and fuller employment may prove elusive. The effect of wages or education on ultimate labor outcomes may experience a lag. Those effects almost certainly differ by country and region.

Central and eastern Europe as a whole does exhibit a positive, mutually reinforcing relationship between the wage ratio, a positive interpretation of EDET, and a positive interpretation of NEET. All three correlations (NEET to wages, NEET to EDET, and wages to EDET) are mildly to modestly positive across these countries collectively. But only the Baltic region reflects the same, uniformly positive relationship. In the Balkans, neither wages nor EDET has any bearing on NEETs. At the same time, the wage ratio and this study's measure of EDET correlate very positively in the Balkans.

Figure 2B
Correlations among variables for each informal, conventional cluster (Visegrád, Baltics, Balkans) and for central and eastern Europe as an entire region



Source: Authors' illustration

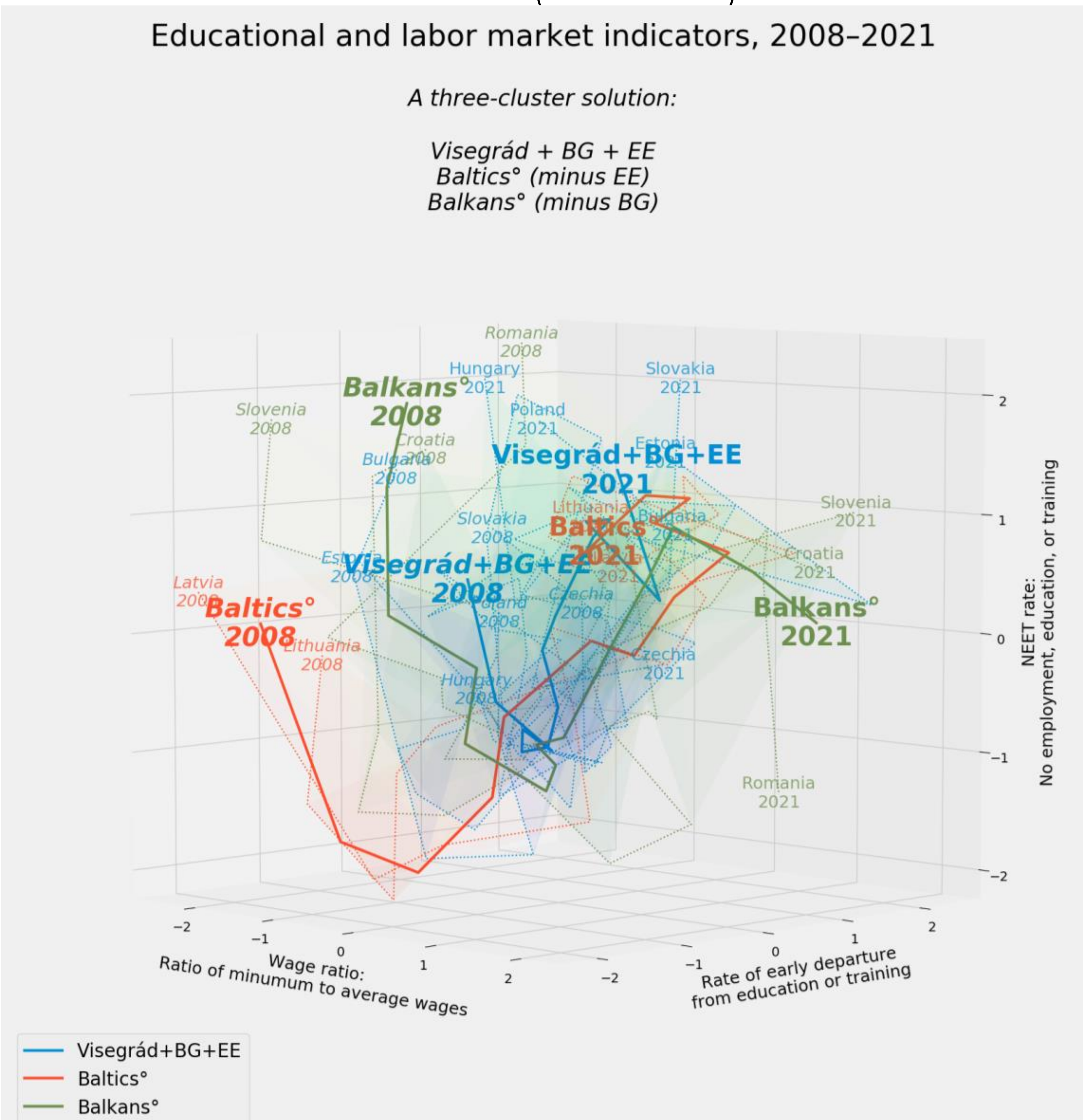
By contrast, the Visegrád countries exhibit a strongly negative relationship between EDET, on one hand, and either wages or NEET, on the other hand. Though at odds with the central and eastern European experience at large, a negative relationship could arise from the low opportunity cost of training or education during recessions and other disruptions of the lower-wage labor market. Cheap, abundant education or training may also lower the number of young people without employment, education, or training. The relative economic and political strength of the Visegrád countries, especially by contrast with poorer Balkan nations such as Bulgaria and Romania, may account for the different correlations.

Such stark differences may arise from diverse economic conditions and differences in policy. From 2008 to 2021, the Visegrád countries strengthened their cooperation in areas including labor policies, benefiting from EU funding and initiatives to improve youth employment (Bieszk-Stolorz & Dmytrów, 2020; Krzaklewska, 2013). The Baltic States, following EU accession, implemented reforms to address youth labor market integration, often focusing on vocational education and training enhancements (European Commission, 2017). The Balkan countries, with varied levels of EU integration, faced diverse challenges. Slovenia and Croatia, as EU members, took advantage of EU youth employment initiatives (Youth Wiki 2020, 2023a). Bulgaria and Romania worked to align with EU standards, addressing high NEET rates through national reforms and EU-supported programs (Institute for Market Economics, 2019; Neagu, Lendzhova, & Keranova, 2021; Toderiță, Damian, & Meiroşu, 2019).

Countries in all of these regions often increased minimum wages to improve living standards while combating youth unemployment and early departures from education through tailored policies, including labor market incentives and educational improvements. This 3D plot suggests that traditional country groupings may reflect more precisely mathematically prescribed agglomeration patterns revealed in subsequent 3D visualizations.

This study now shifts its focus to two formally selected cluster solutions. Cluster solution "A" is a three-cluster solution, and cluster solution "B" is a five-cluster solution. The former is displayed in Figure 3, and the latter, in Figure 4. Corresponding tables of correlations are also shown (Figure 3B and 4B).

Figure 3
3D visualization of Cluster solution "A" (three clusters)



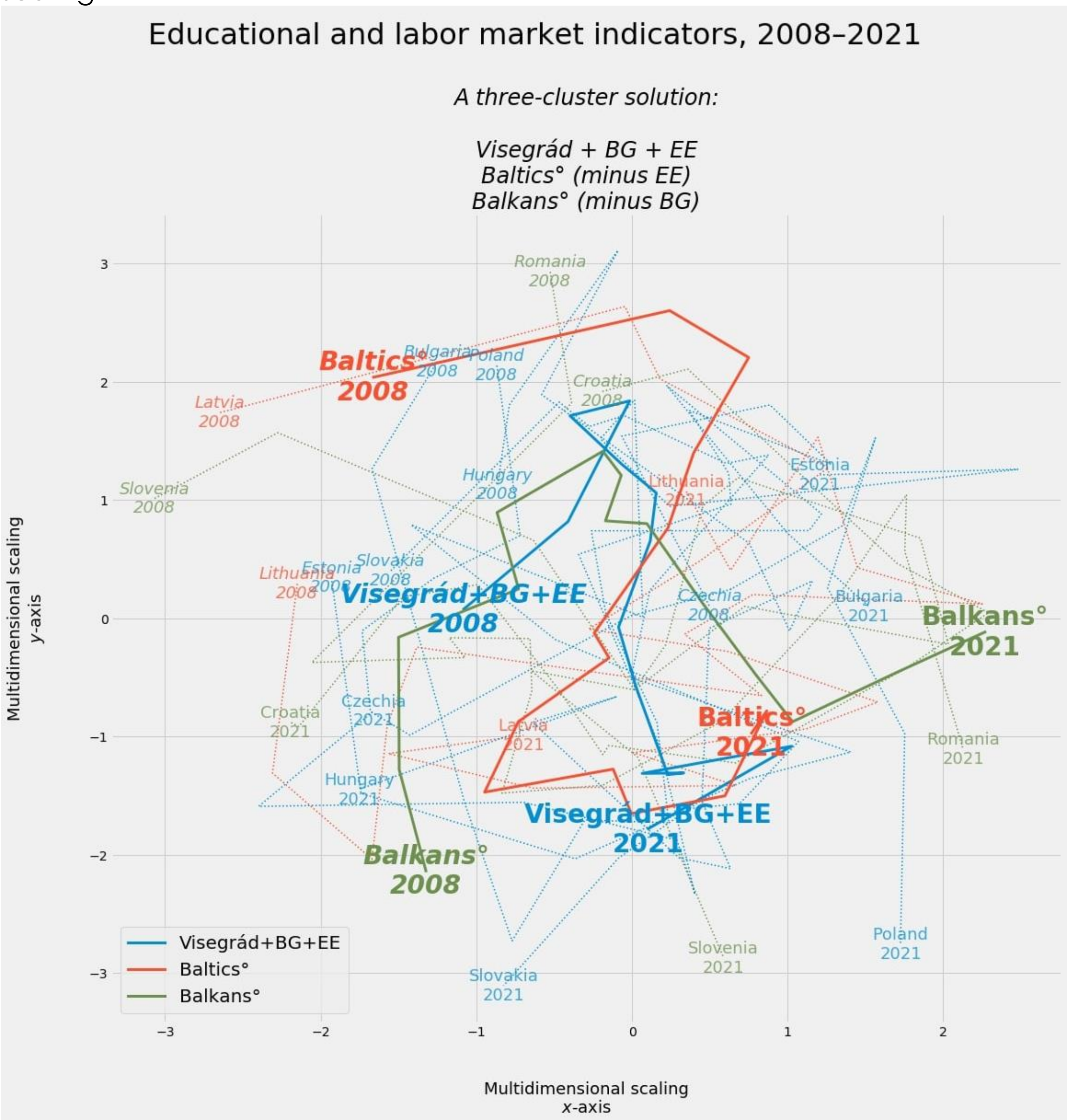
Source: Authors' illustration

The three-cluster solution also appears in two-dimensional form in Figure 3A. Over time, the Visegrád countries, joined by Bulgaria and Estonia, show a cluster that generally trends toward higher values on the wage ratio axis, suggesting relatively stronger wage growth compared to the other two clusters. Relative to the rest of central and eastern countries, minimum wages within this Visegrád-centered cluster have kept better pace with overall wage growth.

The Baltic cluster, excluding Estonia, displays a diverse trajectory, with some movement toward lower NEET and EDET rates. Latvia and Lithuania, the members of this shrunken Baltic cluster, could be considered moderate performers. Their most salient trait is modest improvement in educational performance, one not yet matched by wage growth for their lowest-earning workers.

The Balkan cluster, excluding Bulgaria, tends to have lower values across all variables. Uniformly negative movement across all three dimensions indicates persistent challenges in the labor market. The shrunken Balkan cluster of Croatia, Romania, and Slovenia appears to present the harshest educational and labor market conditions for young people in central and eastern Europe.

Figure 3A
2D visualization of Cluster solution "A" (three clusters) obtained with multidimensional scaling

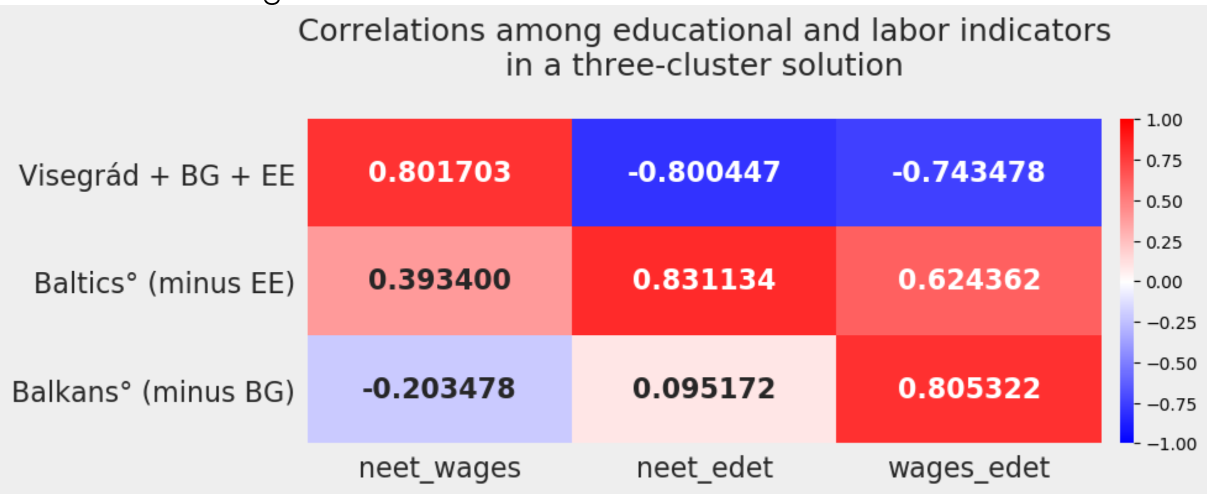


Source: Authors' illustration

Correlations among the three educational and labor market variables in the three-cluster solution "A" closely resemble their correlations in the informal division of central and eastern Europe into its traditional geographic components of the Visegrád Group, the Baltics, and the Balkans. The removal of Bulgaria from the Balkan cohort does have the effect of reversing the sign on the correlations between wages and NEET and between EDET and NEET. However, neither of those correlations had been large, in either direction, even with Bulgaria.

The reassignment of Estonia to an expanded group dominated by Visegrád does not alter the fundamental nature of either Visegrád or the Baltics. The reduced Baltic region of Latvia and Lithuania remains the one corner of eastern and central Europe where all three variables are positively correlated. The addition of Bulgaria and Estonia to an expanded Visegrád Group does not alter the unusual relationship among wage ratios, EDET, and NEET in the dominant countries of (north) central Europe.

Figure 3B
Correlations among variables within the three-cluster solution "A"



Source: Authors' illustration

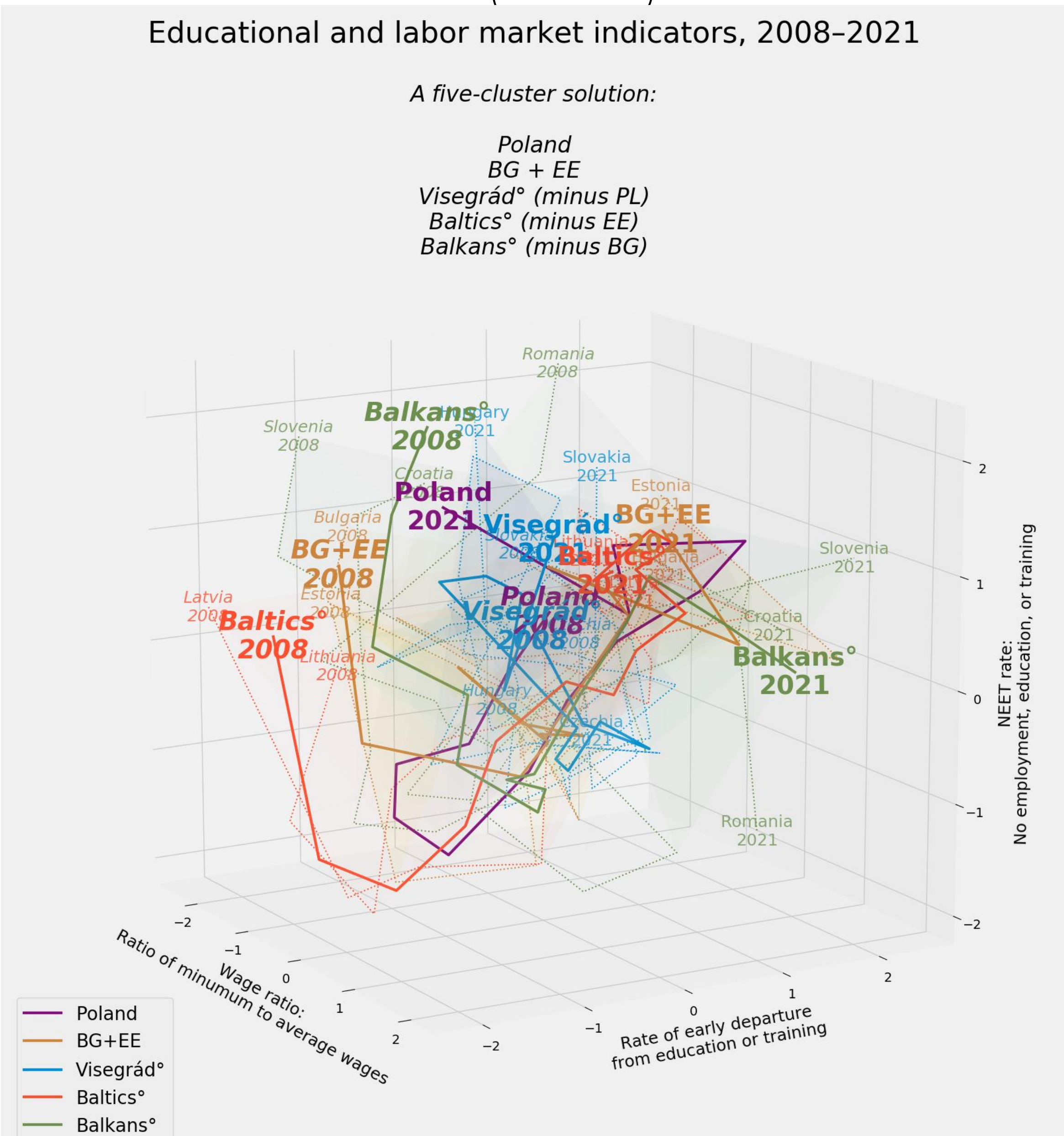
The results of the fourth 3D projection (Figure 4) provide a more granular view of central and Eastern Europe. By presenting five rather than three formal clusters, Poland stands out as a cluster of its own, suggesting unique labor market dynamics or policy outcomes relative to other countries. The presence of a singleton, in any clustering exercise, invites closer examination of that lone member's underlying traits.

The "BG+EE" cluster suggests similarities between Bulgaria and Estonia that isolate them from their respective regional groups. Despite their geographic separation, they share common pioneering trends in youth labor market reforms.

The remaining Visegrád countries, excluding Poland, still cluster together. The removal of Poland shows how Czechia, Hungary, and Slovakia have followed a distinct path relative to Poland. The mathematical definition of separate clusters supports the inference that these three Visegrád countries, whatever their internal variations, resemble each other more than Poland or, for that matter, other countries throughout central and eastern Europe.

The three-cluster solution ("A") and the five-cluster solution ("B") both deviate from conventional definitions of the three regions of central and eastern Europe. The reassignment of Bulgaria and Estonia, either to a larger cluster dominated by Visegrád, or to a new cluster consisting of those two countries, has the unavoidable effect of reducing the original, informal Baltic and Balkan clusters. In three-cluster solution A, the traditional regions of the Baltics and the Balkans each lose a member. Five-cluster solution B goes further. It isolates Poland entirely from the other members of the Visegrád Group.

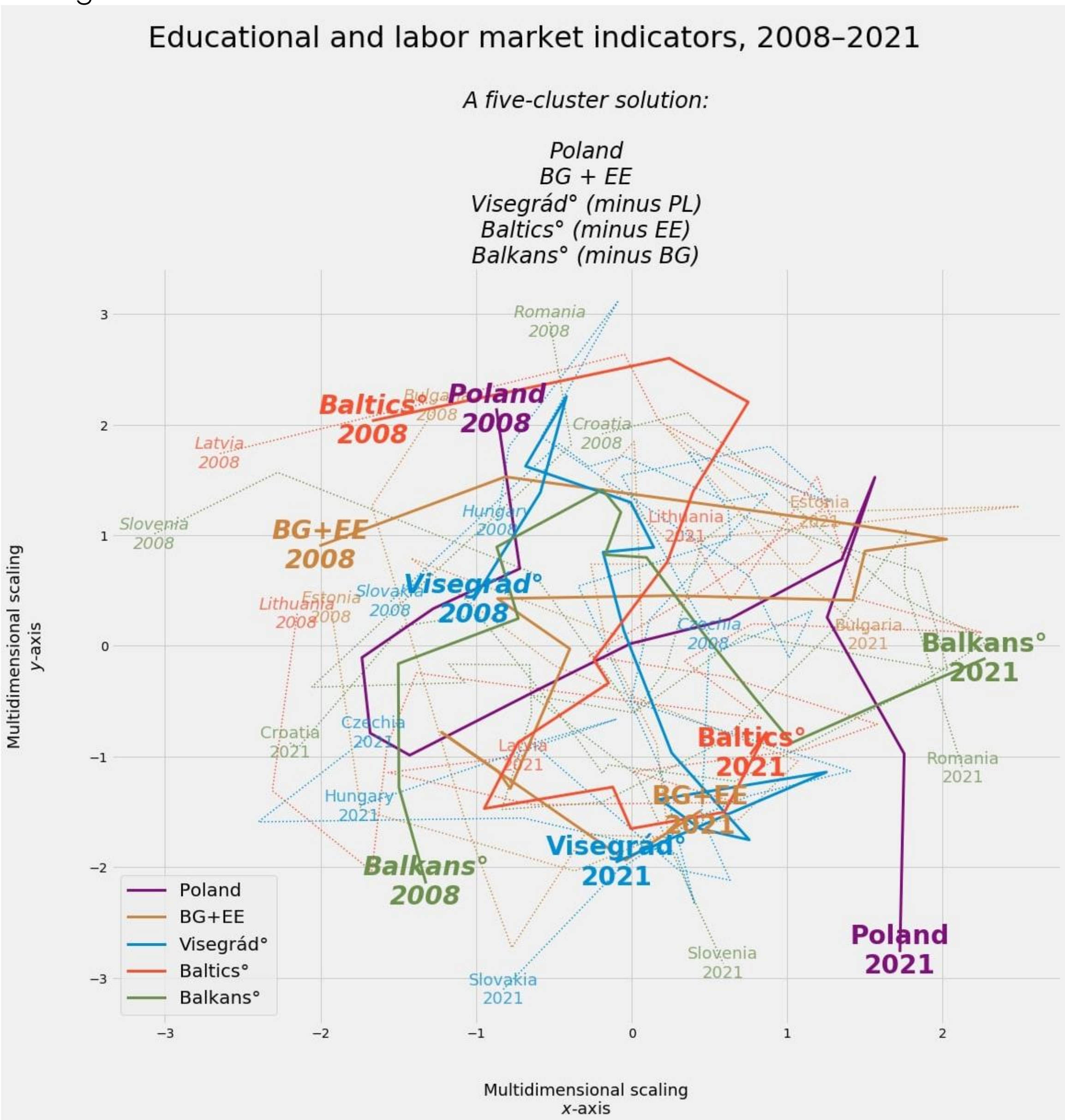
Figure 4
3D visualization of Cluster solution "B" (five clusters)



Source: Authors' illustration

A 2D variant of Figure 4 appears in Figure 4A. In cluster solution B, as depicted in Figures 4 and 4A, the Baltics (excluding Estonia) and the Balkans (excluding Bulgaria) maintain their closeness, perhaps due to shared regional characteristics or responses to economic conditions. The Baltic nations of Latvia and Lithuania show modest progress toward improving the school-to-work transition for young people, whereas a complex set of challenges peculiar to the Balkan region appears to have beset Romania and the ex-Yugoslav member-states of the EU.

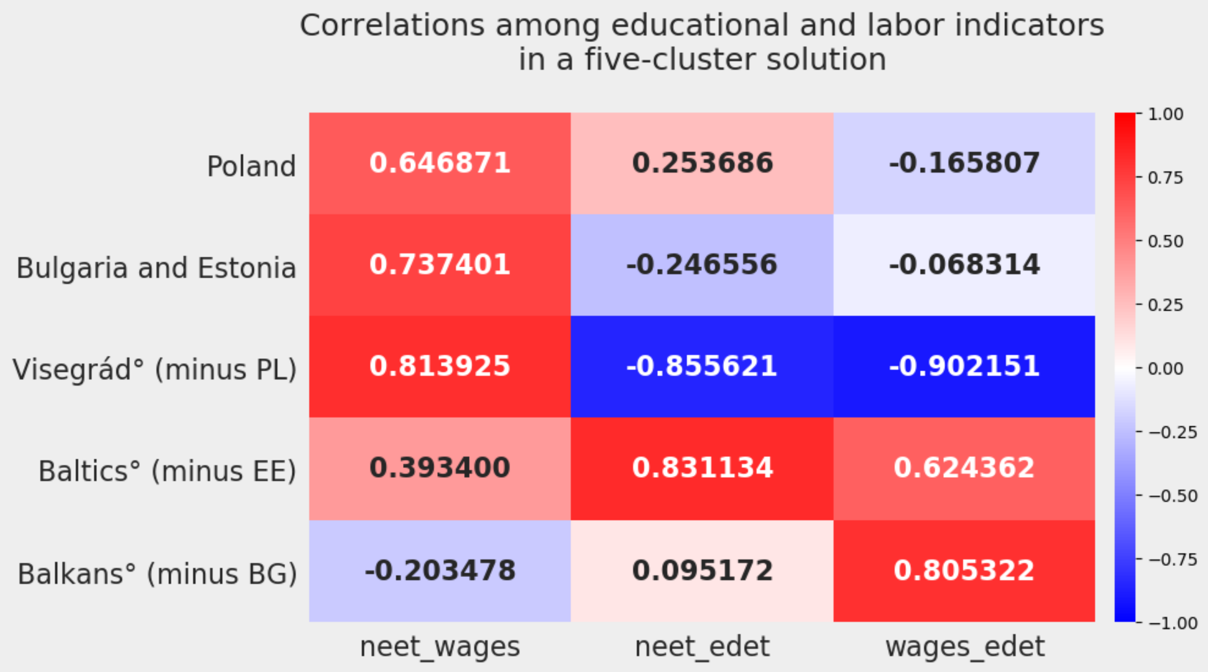
Figure 4A
2D visualization of Cluster solution "B" (five clusters) obtained with multidimensional scaling



Source: Authors' illustration

Correlations among variables within the five-cluster information realistically add only three new sets of information rather than five. The reduced Baltic° and Balkan° clusters from three-cluster solution A remain intact. Visegrád minus Poland retains its distinctive pattern of a negative correlation between EDET on one hand and the wage and NEET variables on the other hand. Poland shows a positive relationship between EDET and NEET. Bulgaria and Estonia, isolated on their own and removed from the expanded Visegrád cluster to which three-cluster solution A had assigned them, do reflect the Visegrád countries' collectively negative correlation between EDET and the wage and NEET variables. But the correlation between EDIT and NEET is very mildly negative, and between EDET and wages even more modestly negative.

Figure 4B
Correlations among variables within the five-cluster solution "B"



Source: Authors' illustration

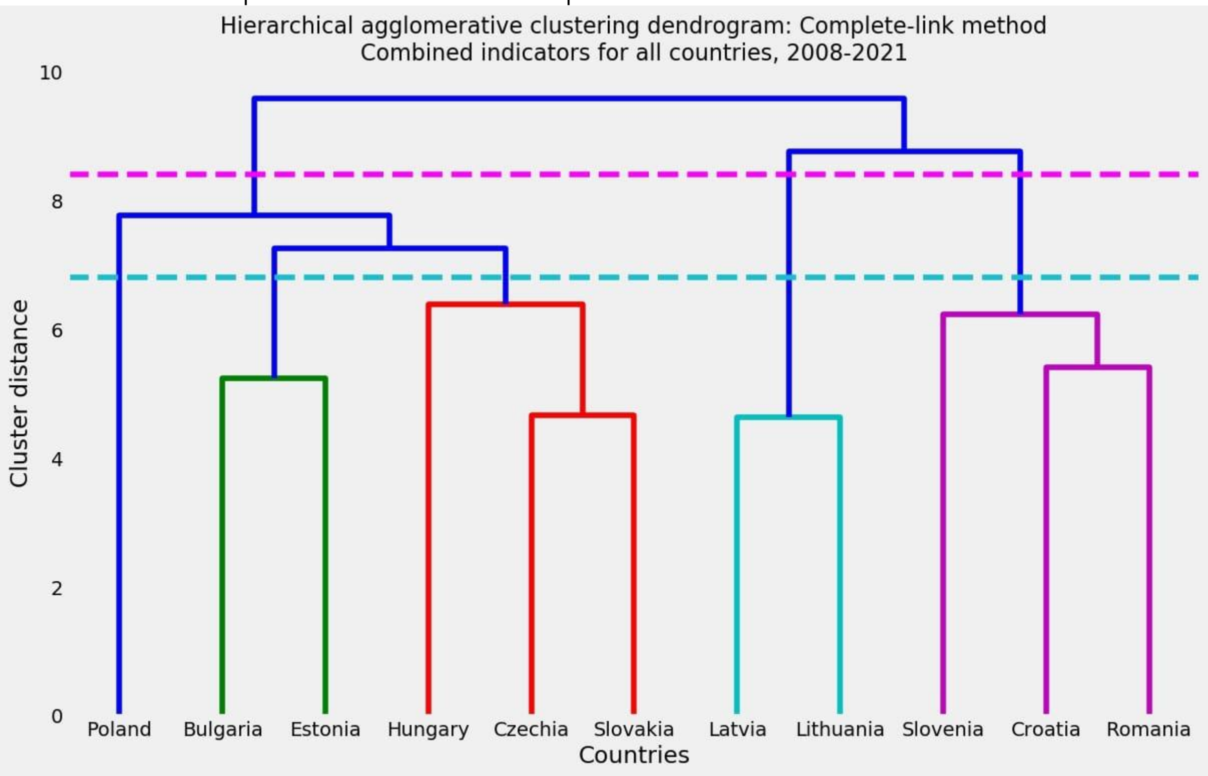
Dendrogram

The lone dendrogram in this study visualizes the results of hierarchical agglomerative clustering using complete linkage based on standardized Euclidean distances. Figure 5 shows how central and eastern European countries are arranged within three- and five-cluster solutions. Although these three- and five-cluster solutions have already been presented in Figures 3 and 4, the dendrograms in this section illustrate those results in a more mathematically formal way. Figure 5 also reveals a two-cluster solution not previously discussed. Thanks to their sheer simplicity, these visualizations also avoid the pitfalls that can beset complicated three-dimensional plots.

Figure 5 presents the three- and five-cluster solutions depicted in earlier 3D projections (Figures 3, 4). Within a five-cluster solution, Poland stands alone. Its status as a singleton indicates a distinct labor market profile from the rest of the group. Bulgaria and Estonia form a separate cluster, based on similarities in their labor market indicators. A larger cluster combines Hungary, Czechia, and Slovakia, implying these Visegrád countries share common labor market characteristics. Latvia, Lithuania, Slovenia, Croatia, and Romania cluster together at a higher distance, forming a plausible supercluster distinct from its counterpart, a Visegrád-dominated supercluster consisting of Poland, Czechia, Hungary, Slovakia, Bulgaria, and Estonia (indicated in Figure 4 as "Visegrád + BG + EE"). Dividing the non-Visegrád supercluster into two distinct components, one containing the Balkan nations of Croatia, Romania, and Slovenia, and other combining Latvia and Lithuania, yields a three-cluster solution.

Combining 3D projections with dendrograms enriches the analysis by providing a macroscopic and microscopic view of the data. While 3D projections give a spatial sense of the data, dendrograms offer detailed, formal insights into the hierarchical relationships and clustering process. The final clustering solutions are determined by considering the cluster distance at which clusters merge, with a larger vertical distance signifying less similarity. Significant increases in the dendrogram's vertical distances suggest natural divisions between clusters.

Figure 5
Dendrogram of combined indicators (wage ratios, EDET and NEET rates) for central and eastern European countries for the period 2008-2021



Source: Authors' illustration

Figure 5 presents at least two and possibly three potential clustering solutions. Notwithstanding previously presented three-cluster and five-cluster solutions, an even more simplified two-cluster solution distinguishes the Visegrád countries, Bulgaria, and Estonia from the rest of the region – namely most of the Baltic and Balkan subregions. Consequently, this analysis affirms previous selections and maintains both the three-cluster solution (displayed as "Cluster solution A" in Figure 3) and the five-cluster solution (displayed as "Cluster solution B" in Figure 4).

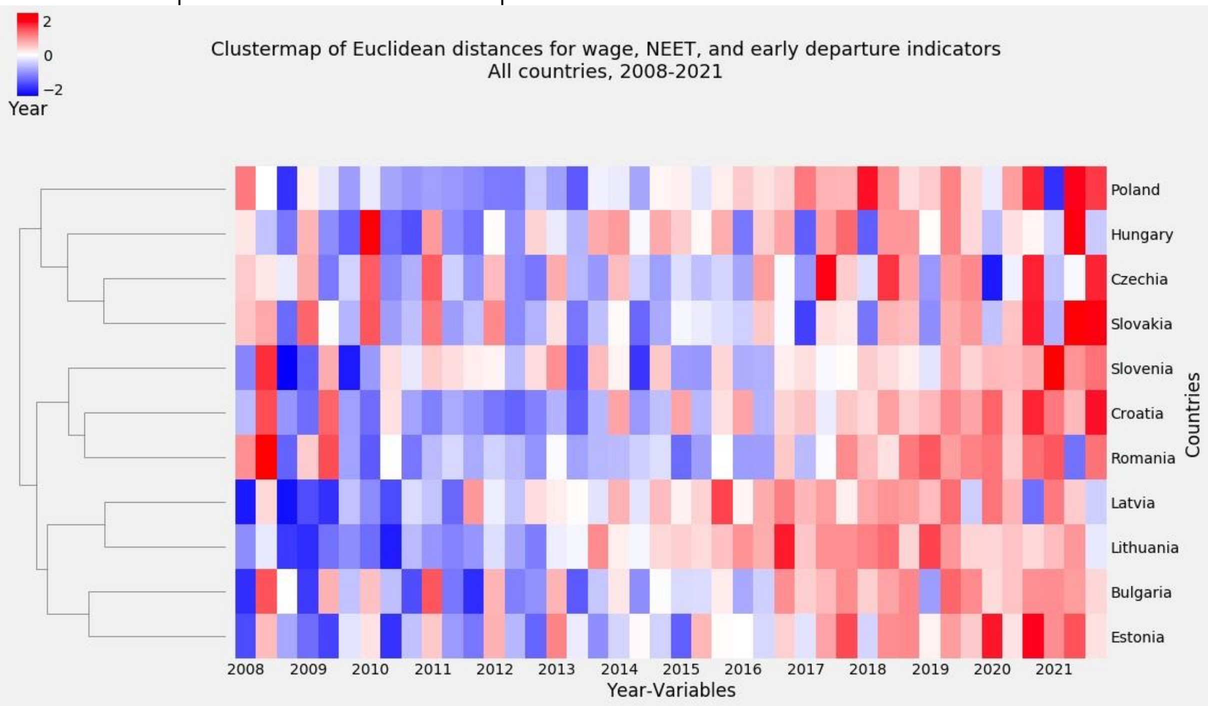
Heatmap

Figure 6 presents a heatmap, the last graphical tool in this study. Colors on a heatmap reflect the level of similarity or difference. Colors closer to blue indicate smaller standardized Euclidean distances and, therefore, denote higher similarity. Colors closer to red represent larger distances and, therefore, indicate lower similarity.

The color comparison does not directly separate countries from one another. Rather, they indicate distance from the Gaussian mean for the time series representing each feature. A blue reading for Poland, for instance, indicates a lower value for that country and for a particular year during the range from 2008 to 2021, relative to values for the Kaitz wage ratio, for instance. A red reading for a higher-than-average result for that variable in a country at a specific moment in time.

The resulting spatial, or distance-based, arrangement among central and eastern European countries appears along the vertical axis of each heatmap, while the horizontal axis shows the one-way progression of years from 2008 through 2021. Each row therefore represents a single country's performance throughout this period. Each column shows country-by-country differences during a single year.

Figure 6
Heatmap of combined indicators (wage ratios, EDET, and NEET rates) for central and eastern European countries for the period 2008-2021



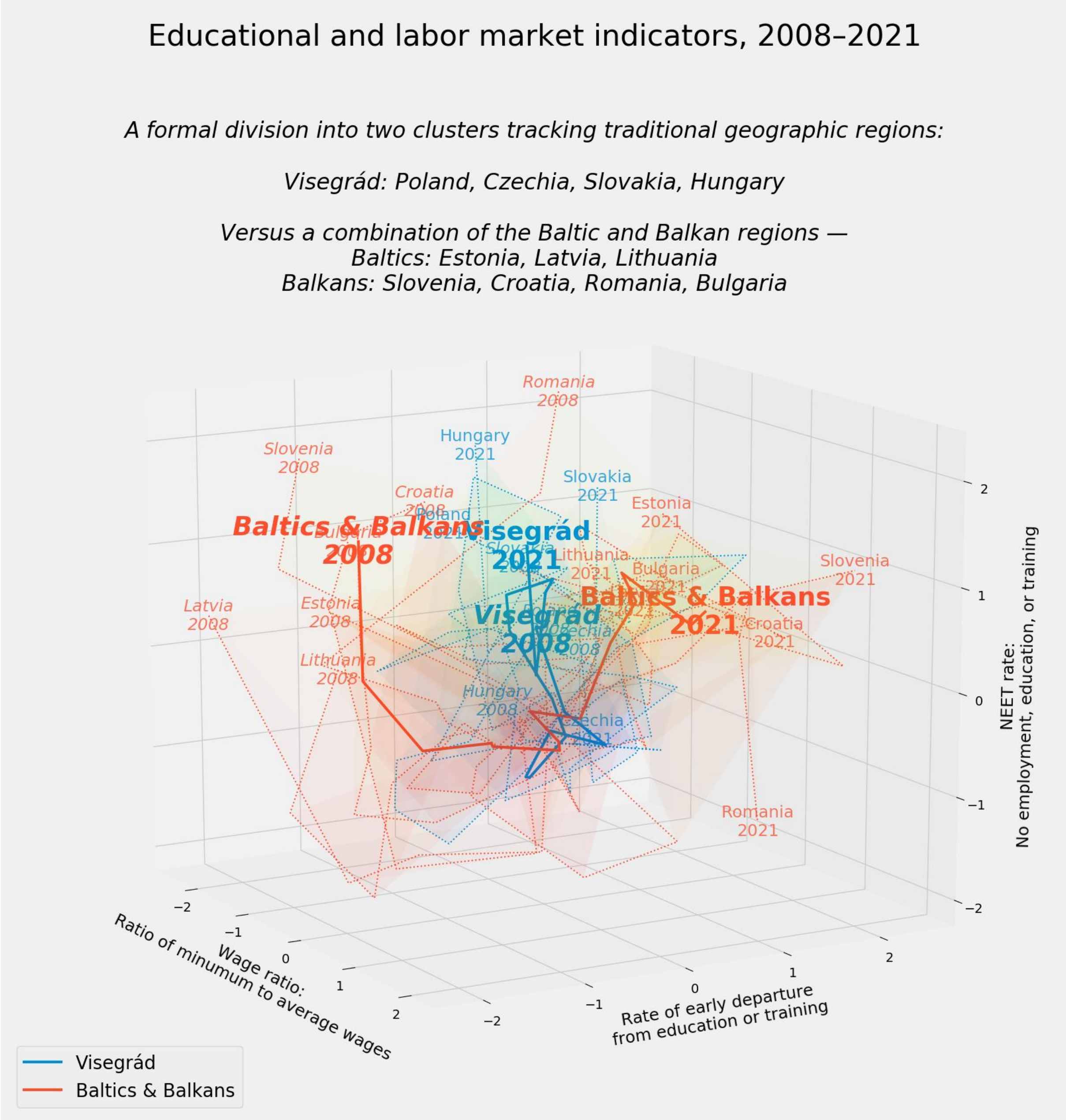
Source: Authors' illustration

The heatmap in Figure 6 also incorporates a dendrogram. This is an artifact of the Seaborn package for Python, which performs hierarchical clustering while generating heatmaps. Because the default linkage for hierarchical clustering within seaborn heatmaps is average rather than complete linkage, this heatmap has the incidental benefit of revealing clustering solutions that might differ from those prescribed by the dendrogram in Figure 5.

Figure 6 enables the distinction between two and five clusters. The former divides central and eastern Europe into two blocs: the Visegrád Group on one side and a grand union of the Baltics and Balkans on the other. The Baltic/Balkan union within Figure 6 does not distinguish perfectly between those regions. However, Bulgaria is closer to Estonia and other Baltic nations than it is to other Balkan countries.

Figure 6 also supports a five-cluster solution identical to that identified in Figure 4 (Cluster solution "B"). Under any clustering solution, Poland stands at a greater distance from other central and eastern European countries. The two-cluster configuration implied by the heatmap in Figure 6 is compelling. It consists simply of the Visegrád Group on one side and a grand north-south alliance of the Baltic and Balkan nations on the other side.

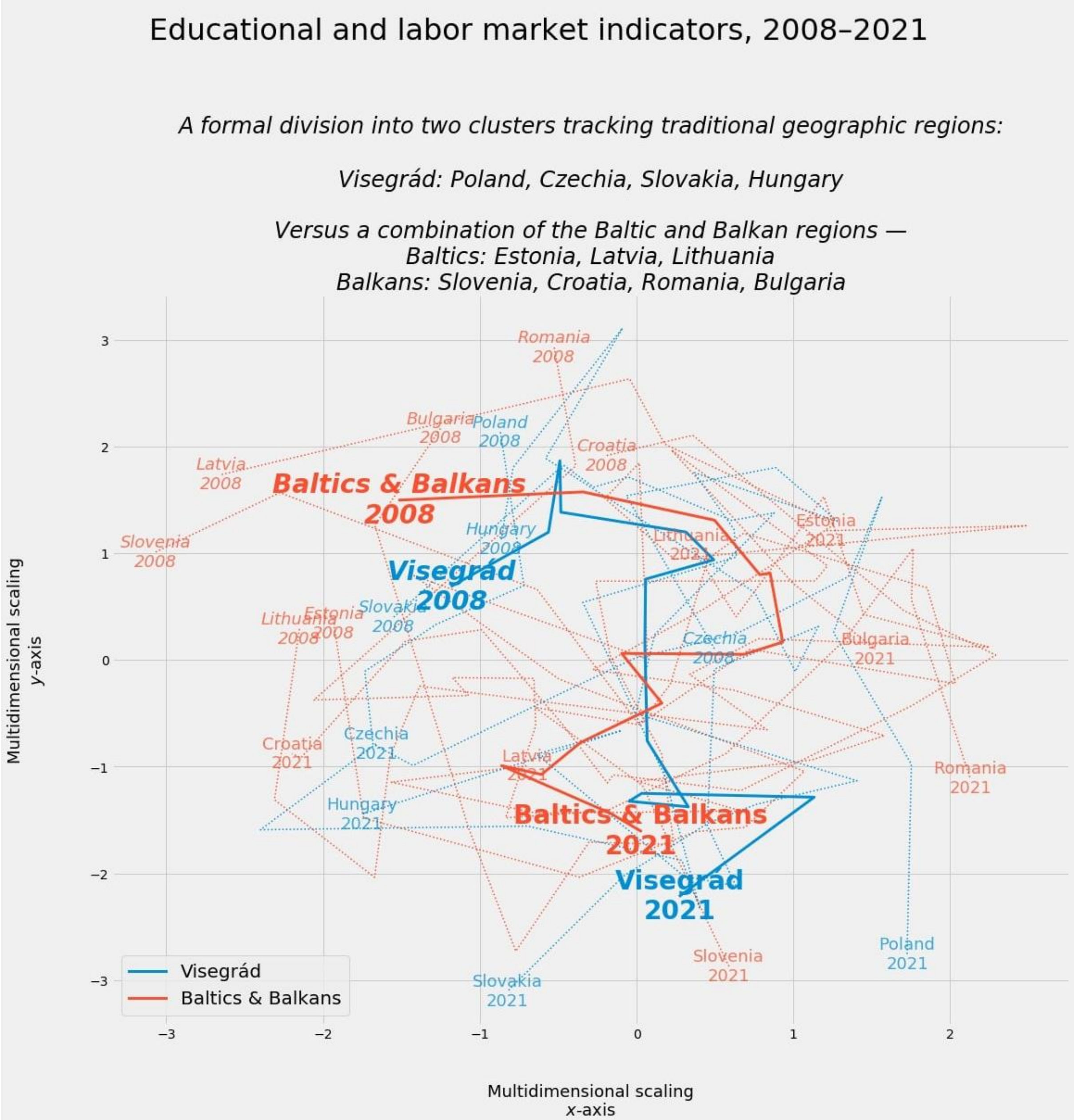
Figure 7
3D visualization of Cluster solution "C" (two clusters), as suggested by the average-linkage hierarchical clustering implied by the heatmap in Figure 6



Source: Authors' illustration

Figure 7 revisits the idea of a 3D projection, this time reduced to this two-cluster solution. The alignment of central and eastern Europe along one of the most geographically, culturally, and historically satisfying boundaries in the continent's history is an intuitive and satisfying result. Figure 7A presents the same information as a two-dimensional, MDS-enabled manifold.

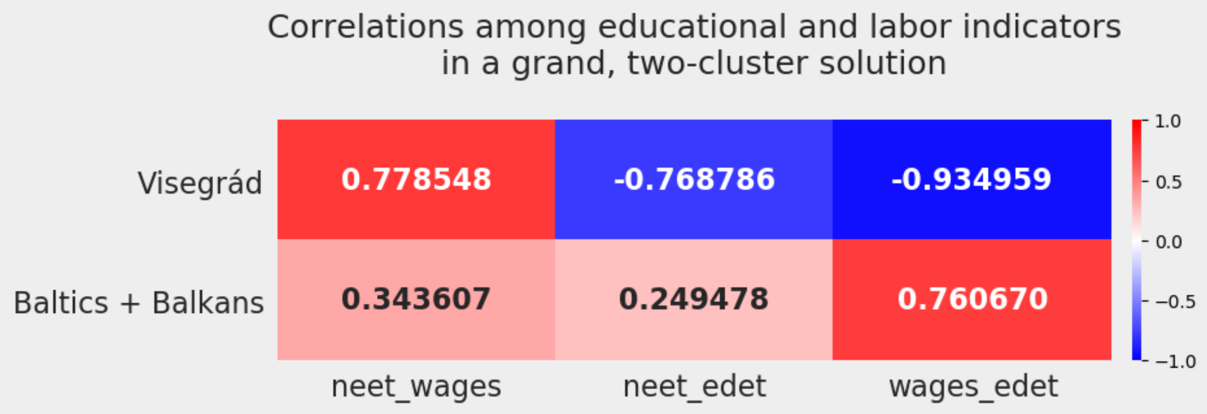
Figure 7A
2D visualization of Cluster solution "C" (two clusters), as suggested by the average-linkage hierarchical clustering implied by the heatmap in Figure 6



Source: Authors' illustration

Even with no more than two cohorts of countries, a correlation table reveals how the three underlying time series relate to one another (Figure 7B), which summarizes the correlations among the Kaitz, EDET, and NEET variables in the two clusters. The seven EU member-states along the Baltic Sea and on the Balkan Peninsula report positive correlations among all three variables. The four large states in the Visegrád Group report a negative relationship between EDET and the other variables. In microcosm, contemporary central and eastern Europe. The contrast between the Visegrád countries at the geographic and historic core of central and eastern Europe, on one hand, and the Baltics and Balkans, on the other hand, is reminiscent of the historical rivalry between the "inner six" nations that signed the original Treaty of Rome and the "outer seven" members of the European Free Trade Association (Kaiser, 1997).

Figure 7B
Correlations among variables within the grand two-cluster solution "C"



Source: Authors' illustration

Discussion

Having presented all the results of its clustering analysis, this paper now discusses the forces that may propel the youth labor market dynamics of central and eastern Europe. Again, this paper draws upon the reports of national and international organizations and national and EU policies aimed at education and school-to-work transition.

From the Baltic to the Black and Adriatic seas, central and eastern Europe stands out as a dynamic player within the European Union. Despite a common socialist history, the countries within this region exhibit diverse policies and responses to labor market challenges. Human judgment, uninformed by formal mathematical analysis, might be inclined to align these countries into three groups according to conventional geographical and political understandings: the Visegrád Group, the Baltics, and the Balkans.

However, hierarchical clustering enables deeper than superficial definitions by exploring economic reasons tied to labor market policies. This study has focused on three variables: wage policies and educational outcomes measurable by EDET and NEET rates. This research uses 3D projections to enhance our discussion of competing solutions, which offer two-, three-, and five-cluster solutions describing how countries in this region have strived to keep young people in school and to help them avoid unproductive labor market outcomes.

Wage ratios

Central and eastern European countries have actively increased legal minimum wages between 2008 and 2021 (ILOSTAT Explorer, 2024). Czechia experienced stagnation in minimum wage growth during the global financial crisis of 2008-2013 (Grossmann, 2021). Freezing wages at 8,000 CZK negatively affected the ratio of minimum to mean wages. Nonetheless, from 2008 to 2021, Czechia saw a 90 percent increase in the minimum wage.

Slovakia and Poland recorded the largest increases in wage ratio. Increases exceeding 10 percentage points can be attributed to significant hikes in minimum wages by 132 percent and 149 percent, respectively. Poland, in particular, exhibited the highest wage ratio throughout the period, exceeding 50 percent in 2021.

Among central and eastern European countries, Poland arguably distinguished itself with the best policy response during the global economic crisis by implementing countercyclical measures (Epstein et al., 2012; Piatkowski, 2015). Poland was the only

EU economy to avert a recession during that crisis. These reasons support clustering solutions that treat Poland as a singleton.

Latvia and Lithuania exhibited similar increases in their wage ratios, around 4 percentage points each during the observed period. In contrast, the Balkan nations of Slovenia, Croatia, and Romania, alongside Poland, ranked favorably in terms of wage ratio. Notably, Slovenia led all of central and eastern Europe with a 55.2 percent wage ratio, surpassing the EU's minimum wage directive threshold of 50 percent (European Parliament, 2022). Slovenia implemented two significant increases in the minimum wage in 2010 and 2018 (Slovenian Press Agency – STA, 2020).

The EDET rate

Early departures from compulsory schooling merit further discussion under a three-cluster solution (Figure 3). From 2008 to 2021, all four countries of the Visegrád Group witnessed an increase in EDET rates. On the other hand, Czechia, Slovakia, and Poland have already fallen below the EU's 2030 threshold of 9 percent. Hungary continues to struggle with a persistent rate of approximately 12 percent. Despite the implementation of programs and strategies aimed at reducing dropout rates among compulsory school attendees and providing support for disadvantaged students (such as the "Mid-Term Strategy on Early School Leaving 2014-2020", the "Springboard Programme", and the "Tanoda Programme") (Lénárt, 2021; Youth Wiki, 2023f), Hungary has not made significant progress.

Bulgaria and Estonia have managed to reduce their EDET rates, with more substantial reductions during times of crisis. This can be partially attributed to the lower opportunity costs of education during economic downturns, as finding employment becomes more challenging (Adamopoulou & Tanzi, 2017; Borjas, 2013; Long, 2014). Nonetheless, these countries still need to meet the EU's 2030 threshold.

Latvia and Lithuania also show promising results in NEET indicators. Notably, Latvia has reduced its dropout rate by more than half from 2008 to 2021, a success that can be linked to its long-term and sustainable growth strategy in anticipation of 2030. Both of these Baltic countries meet the EU threshold thanks to the development of early warning systems, such as electronic school diaries for parental monitoring of school absences, efforts to track and support low achievers, and youth homes for disadvantaged students.

Croatia and Slovenia boast the lowest EDET rates in all of central and eastern Europe, around 2 to 3 percent in 2021. Unsurprisingly, these countries lack a national strategy specifically targeting this issue.

Romania, conversely, has faced the highest EDET rates throughout the period from 2008 to 2021. Difficulties with dropouts arise largely from social exclusion affecting primarily four groups (Youth Wiki, 2023g): (a) young people aged 18-24, (b) young people from low-income families, (c) young people in rural areas, and (d) ethnic minorities such as the Roma. Despite Romania's efforts to reduce the EDET rate within the observed timeframe, it continues to lag the rest of Europe. Romania's EDET rate stands at 15.3 percent as of 2021.

The EU's Youth Guarantee program is aimed as much at solving the EDET issue as it is at addressing NEET challenges. The two concerns overlap considerably. To some extent, preventing early departures from education or training can also mitigate future failures to find work, training, or education. In other words, policies targeting the EDET rate may reduce the future incidence of NEET. Therefore, it is crucial to accurately identify the problem, develop tailored programs that respect the diversity of vulnerable groups, and actively implement and regularly update these programs to reintegrate as many young people as possible into the education system, labor market, and society at large.

The NEET rate

The countries of central and eastern Europe are working toward an EU-level goal of reducing the NEET rate to below 9 percent by 2030. Since its introduction in the 1990s in the United Kingdom, where it replaced the term "Status Zer0" (Istance, Rees, & Williamson, 1994, as cited in Mascherini, 2018), the NEET concept has evolved, particularly regarding the age groups it encompasses. While the EU target focuses on the 15-29 age group, our analysis concentrates on the 25-34 age group, where being NEET represents a more significant concern.

Cluster solution "A" identifies varied challenges related to NEET. Despite fluctuations, three of the four Visegrád countries (Hungary, Poland, and Slovakia) had reduced their NEET rates by 2021. Hungary showed the most significant decrease of 10.9 percentage points. By contrast, Czechia experienced stagnation, as its NEET rate stayed persistently close to 20 percent.

Bulgaria and Estonia are addressing similar challenges in retaining young people in secondary vocational education and training. School dropouts, especially from a track consciously designed to serve students who do not intend to continue their education after high school, are a crucial component of the NEET phenomenon. Secondary school dropouts often join the NEET population. Despite vastly different NEET rates (22 percent in Bulgaria and 13 percent in Estonia in 2021), both countries face common structural issues and have pursued similar national policies. Bulgaria has initiated programs (Youth Wiki, 2023b) to support the long-term unemployed and young people who are not taking part in the labor market ("Program for Training and Employment of Long-Term Unemployed Persons" and "Activation of Inactive Persons"). Estonia has launched similar initiatives (Youth Wiki, 2023c) such as "My First Job" and "Youth Prop Up."

Latvia and Lithuania experienced similar dynamics, with a NEET rate that peaked in 2009-2010. A general decrease in NEET rates ensued, thanks to successful national intervention projects (Youth Wiki, 2023d, 2023e) such as "Know and Do!" (Latvia) and "Discover Yourself" and "New Start" (Lithuania).

The Balkan trio of Croatia, Slovenia, and Romania saw an increase in NEET rates during 2008-2013. As of 2021, Slovenia is closest to the 2030 EU target with a 10.1 percent NEET rate, while Romania lags at 24.4 percent. Croatia falls in between. All three countries report higher NEET rates in 2021 compared to 2008.

This analysis of NEET rates justifies the preference for a three-cluster solution. However, it is important to consider other factors shaping the youth labor market, such as income distribution and the quality of the educational system. Incorporating more variables could yield further insights.

All of the countries in this study have embraced the EU's "Youth Guarantee" program (adopted in 2013; Council of the EU, 2013) aimed at addressing NEET challenges after the global financial crisis. The "Reinforced Youth Guarantee" (2020, Council of the EU, 2020) responded to the COVID-19 crisis, serving as EU guidelines for developing national programs tailored to local market needs. The long-term impact of this strategy on young people in the European labor market remains to be seen.

Conclusion

This paper has examined central and eastern Europe during a period marked by two global economic shocks: the global financial crisis (2008-2013) and the COVID-19 pandemic (2020-2021). It focuses on the challenges faced by young people. This paper has explored these challenges by analyzing wage ratios and EDET and NEET rates. This research makes a distinctive contribution through its novel graphical exploration of the youth labor market. Through 3D projections, dendrograms, and heatmaps, this study offers new insights into the dynamics underlying that market.

Hierarchical clustering with complete linkage, as applied to time series data spanning 2008 through 2021, reveals multiple cluster solutions. Two of those solutions appeared illustrative of the youth labor market's dynamics: a narrower three-cluster solution that divides this part of Europe into (1) the Visegrád Group plus Bulgaria and Estonia, (2) the remaining Baltic countries, and (3) the remaining Balkan countries.

A five-cluster solution further subdivides the first, Visegrád-dominated cluster into a Polish singleton, the geographically remote pair of Bulgaria and Estonia, and the remaining Visegrád countries of Czechia, Hungary, and Slovakia. The remaining two clusters consist, as they do in the three-cluster solution, of the Baltic states (minus Estonia) and the Balkan states (minus Bulgaria).

An alternative average-linkage approach to hierarchical clustering supports a historically and geographically intuitive bifurcation of central and eastern Europe into the Visegrád Group and a grand north/south coalition of the Baltic and Balkan countries.

The analysis highlighted standout performers: Slovenia excels in combating EDET challenges (alongside Croatia) and achieving high wage ratios (alongside Poland). Slovenia also boasts the region's lowest NEET rates. The Baltic countries have shown respectable progress.

These findings suggest that a strategic shift might be beneficial. Rather than solely promoting tertiary education, policymakers should place heightened focus on preventing dropout from compulsory education. This appears to be an especially significant issue in Bulgaria, Hungary, and Romania. Enhancing vocational education and training programs is crucial. As EU members, all countries in this study are aligned with common labor market strategies, especially the Youth Guarantee program. The enduring influence of political and economic legacies is apparent. Geopolitical proximity influences regional clustering and the unique challenges each bloc faces.

The primary limitations of this study arise from the same traits that supply its leading strengths. As a species of unsupervised machine learning, clustering analysis cannot directly support causal inferences or validate hypotheses regarding drivers of youth labor markets. Those research questions are better addressed by generalized linear methods and related forms of supervised machine learning. Studies such as Tudor et al. (2023) do combine clustering analysis with OLS regression.

A study combining unsupervised and supervised learning should be able to achieve at least two tasks beyond those attainable through clustering alone. First, studies incorporating regression should aspire to predict future labor market outcomes. Second, such studies should extract causal inferences related to hypotheses on the relationship of education, training, and other labor market inputs with ultimate labor market outcomes.

The identification of specific clusters within part or all of a multinational union such as the European Union is the primary contribution of a study such as this one. Predictions, causal inferences, and policy prescriptions may all be improved if researchers, guided by clustering analysis, can tailor their conclusions and recommendations according to geopolitical differences revealed by the data.

Other limitations of this research include the choice of average wage over median wage for calculating wage ratios due to incomplete data across several countries. In line with studies such as Dingeldey and Buttler (2023); Krpan, Gardijan Kedžo, and Žmuk (2023); and Tudor et al. (2023), future research should consider additional variables such as income distribution measures, socioeconomic indicators, and indicators of educational quality. More comprehensive collection and evaluation of distinct aspects of educational inputs and labor market outcomes could overcome some of the limitations of this study. In particular, a more detailed analysis of the NEET rate, distinguishing between active and inactive youth, could further refine policymakers' understanding.

Consistent with its stated research hypothesis, this article has classified similarities and differences in the school-to-work transition in central and eastern Europe. Conventional historical and geopolitical understandings sort the eleven countries in this article into the following three groups: the Visegrád countries, the Baltics, and the Balkans. Rigorous mathematical evaluation of three youth labor market variables, attained through hierarchical clustering and multidimensional scaling, supports two-, three-, and five-cluster solutions that partially reflect the conventional geopolitical taxonomy, but ultimately recommend more carefully tailored ontologies that should guide the formulation of labor policy in this vital region within the European Union.

Understanding the current labor market situation, as shaped by past governmental decisions, is vital to effective policymaking. The dynamic environment, evolving political landscape, and lingering historical influences all challenge the EU's convergence goals. Advanced economies are converging at a faster pace, while less developed ones are facing divergence pressures (Rambla & Scandurra, 2021). The European Union must ask whether convergence is a realistic goal, or whether the EU should strive instead for regional convergence. Central and eastern Europe stands at a crossroads, facing the challenge of meeting the EU's 2030 goals amid these complexities.

References

1. Adamecz, A. (2023). Longer schooling with grade retention: The effects of increasing the school leaving age on dropping out and labour market success. *Economics of Education Review*, *97*, 1-33. https://doi.org/10.1016/j.econedurev.2023.102487
2. Adamopoulou, E., & Tanzi, G., M. (2017). Academic Drop-Out and the Great Recession. *Journal of Human Capital, 11(1)*, 35-71. Retrieved from https://doi.org/10.1086/690650
3. Aggarwal, C., C., Hinneburg, A., & Keim, D., A. (2001). On the Surprising Behavior of Distance Metrics in High Dimensional Space. *Database Theory – ICDT 2001. Lecture Notes in Computer Science, Springer*, Berlin/Heidelberg, 1973, 420-434. https://doi.org/10.1007/3-540-44503-X_27
4. Arpaia, A., Cardoso, P., Kiss, A., Van Herck, K., & Vandeplas, A. (2017). Statutory Minimum Wages in the EU: Institutional Settings and Macroeconomic Implications. *IZA Policy Paper No. 124, Institute of Labor Economics*, 1-49.
5. Askenazy, P. (2003). Minimum wage, exports, and growth. *European Economic Review, 47(1)*, 147-164. https://doi.org/10.1016/S0014-2921(02)00187-3
6. Bernard, A., B., & Durlauf, S., N. (1995). Convergence in international output. *Journal of Applied Econometrics*, *10*(2), 97-108. https://doi.org/10.1002/jae.3950100202
7. Bernard, A., B., & Durlauf, S., N. (1996). Interpreting tests of the convergence hypothesis. Journal of Econometrics, *71*(1-2), 161-173. https://doi.org/10.1016/0304-4076(94)01699-2
8. Bezdek, J., C. (2021). *Elementary Cluster Analysis: Four Basic Methods that (Usually) Work*. Gistrup, Denmark: River Publishers.
9. Bieszk-Stolorz, B., & Dmytrów, K. (2020). Influence of Accession of the Visegrad Group Countries to the EU on the Situation in Their Labour Markets. *Sustainability 2020, 12(16)*, 6694. https://doi.org/10.3390/su12166694
10. Blatná, D. (2020). Analysis of the Relationship between the Share of Early Leavers from Education and Training and Selected Socio-Economic Indicators in the Czech Republic in the Period 2005 – 2018. The 14th International Days of Statistics and Economics, Prague, September 10-12, 2020, 136-145.

11. Bokros, L. (2013). *Accidental Occidental: Economics and Culture of Transition in Mitteleuropa, the Baltic and the Balkan Area*. Budapest: Central European University Press.
12. Borjas, G., J. (2013). *Labor Economics* (6th ed.). New York, United States: McGraw-Hill Companies, Inc.
13. Brown, C., Gilroy, C., & Kohen, A. (1982). The effect of the minimum wage on employment and unemployment. *Journal of Economic Literature*, *20*(2), 487-528. Retrieved from https://www.jstor.org/stable/2724487
14. Burzykowski, T., Rousseau, A.-J., Geubbelmans, M., & Valkenborg, D. (2023). Introduction to Machine Learning. *American Journal of Orthodontics & Dentofacial Orthopedics, 163(5)*, 732-734, https://doi.org/10.1016/j.ajodo.2023.02.005
15. Council of the European Union (2013). Council Recommendation of 22 April 2013 on establishing a Youth Guarantee. Retrieved from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32013H0426%2801%29
16. Council of The European Union (2020). Council Recommendation of 30 October 2020 on A Bridge to Jobs – Reinforcing the Youth Guarantee and replacing the Council Recommendation of 22 April 2013 on establishing a Youth Guarantee 2020/C 372/01. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?toc=OJ%3AC%3A2020%3A372%3ATOC&uri=uriserv%3AOJ.C_.2020.372.01.0001.01.ENG
17. Council of the European Union (2021). Council Resolution on a strategic framework for European cooperation in education and training toward the European Education Area and beyond (2021-2030), 2021/C, 66/01. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32021G0226%2801%29&qid=1617706323424
18. Cox, M., A., A., & Cox, T., F. (2008). Multidimensional scaling. In Chen, C.-H., Härdle, W., & Unwin, A. (Ed.). *Handbook of Data Visualization* (pp. 315-347). Berlin: Springer. https://doi.org/10.1007/978-3-540-33037-0_14
19. Cuestas, J., C., Monfort, M., Ordóñez, J. (2021). The education pillar of the Europe 2020 strategy: a convergence analysis. *Empirica*, *48*(4), 1113-1129. https://doi.org/10.1007/s10663-021-09504-x
20. De Luca, G., Mazzocchi, P., Quintano, C., & Rocca, A. (2020). Going Behind the High Rates of NEETs in Italy and Spain: The Role of Early School Leavers. *Social Indicators Research, 151*(1), 345-363. https://doi.org/10.1007/s11205-020-02370-3
21. Dingeldey, I., & Buttler, D. (2023). School-to-work transition regimes in post-socialist countries: An introduction. In Buttler, D., Ławrynowicz, M., & Michoń, P. (Ed.). *School-to-Work Transition in Comparative Perspective* (pp. 156-173.). Cheltenham, UK, and Northampton, Massachusetts, USA: Edward Elgar Publishing. https://doi.org/10.4337/9781800370111.00015
22. Domingos, P. (2012). A Few Useful Things to Know about Machine Learning. *Communications of the ACM, 55*(10), 78-87. https://doi.org/10.1145/2347736.2347755
23. El Bouchefry, K., & de Souza, R., S. (2020). Learning in Big Data: Introduction to Machine Learning. In Škoda, P., & Adam, R. (Ed.), *Knowledge Discovery in Big Data from Astronomy and Earth Observation: AstroGeoInformatics*, pp. 225-249. St. Louis, Mo.: Elsevier. https://doi.org/10.1016/B978-0-12-819154-S.00023-3
24. Epstein, N., Goretti, M., Llaudes, R., & Velculescu, D., (2012). Poland: A Beacon of Resilience in Europe. In Bakker, B., B., & Klingen, C. (Ed.). *How Emerging Europe Came Through the 2008/09 Crisis* (pp. 155-162). International Monetary Fund. Retrieved from https://www.elibrary.imf.org/display/book/9781616353810/ch012.xml
25. European Commission (2010). EUROPE 2020 – A strategy for smart, sustainable and inclusive growth. Retrieved from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A52010DC2020
26. European Commission (2017). Baltic Alliance for Apprenticeships. Enhancing the attractiveness of VET systems in the Baltic states through work-based learning and apprenticeships. Retrieved from https://ec.europa.eu/social/BlobServlet?docId=18440&langId=en

27. European Parliament, Council of the European Union (2022). Directive (EU) 2022/2041 of the European Parliament and of the Council of 19 October 2022 on adequate minimum wages in the European Union. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022L2041
28. Evangelista, P., F., Embrechts, M., J., & Szymanski, B., K. (2006). Taming the Curse of Dimensionality in Kernels and Novelty Detection. In: Abraham, A., de Baets, B., Köppen, M., Nickolay, B. (Ed.) Applied Soft Computing Technologies: The Challenge of Complexity. Advances in Soft Computing, Vol 34. Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-31662-0_33
29. Fialová, K., & Mysíková, M. (2020). Minimum Wage and Youth Employment in Regions of the Visegrád Countries. *Eastern European Economics,* *59*(1), 82-102. https://doi.org/10.1080/00128775.2020.1816474
30. Furlong, A. (2006). Not a very NEET solution: representing problematic labour market transitions among early school-leavers. *Work, Employment and Society, 20*(3), 553-569. https://doi.org/10.1177/0950017006067001
31. Geiger, B., C., & Kubin, G. (2012). Relative Information Loss in the PCA. *Proceedings of the 2012 IEEE Information Theory Workshop*, Lausanne, Switzerland, pp. 562-566. https://doi.org/10.1109/ITW.2012.6404738
32. Ghica, L.-A. (2008). Rhetorical Strategies, Institutional Dilemmas: The Visegrád Group and the Baltic Cooperation Facing the EU and NATO Accession Process. *Analele Universității din Bucureşti. Seria Ştiințe Politice*, *10*(10), 75-86. Retrieved from https://www.ceeol.com/search/article-detail?id=89903
33. Grossmann, J. (2021). The Effects of Minimum Wage Increases in the Czech Republic. *CERGE-EI Working Paper Series No. 679*. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3784096
34. Hair, J., F., Black, W., C., Babin, B., J., & Anderson, R., E. (2018). *Multivariate Data Analysis* (8th ed.). Andover, Hampshire, United Kingdom: Cengage Learning EMEA.
35. Institute for Market Economics (2019). *Assessment of the people not in employment, education and training (NEETs) in Bulgaria and policy measures to effectively address their integration*. European Commission, Luxembourg: Publications Office of the European Union, 1-70. https://doi.org/10.2767/768177
36. Hout, M. C., Papesh, M. H., & Goldinger, S. D. (2013). Multidimensional scaling. *WIREs Cognitive Science*, *4*(1), 93-103. https://doi.org/10.1002/wcs.1203
37. International Labour Organization. Statutory nominal gross monthly minimum wage – Annual. ILOSTAT Explorer. Retrieved from https://www.ilo.org/shinyapps/bulkexplorer31/?lang=en&id=EAR_4MMN_CUR_NB_A [January 31, 2024]
38. Istance, D., Rees, G., & Williamson, H. (1994). *Young People Not in Education, Training or Employment in South Glamorgan*. Cardiff, UK: South Glamorgan Training and Enterprise Council.
39. Jolliffe, I., T. (2002). *Principal Component Analysis* (2nd ed.). New York, United States: Springer-Verlag New York, Inc.
40. Kaiser, W. (1997). Challenge to the community: the creation, crisis and consolidation of the European Free Trade Association, 1958-72. *Journal of European Integration History*, *3*(1), 7-33.
41. Kaitz, H. (1970). Experience of the Past: The National Minimum. *Youth unemployment and minimum wages*, Bulletin 1657, United States Department of Labor, Bureau of Labor Statistics, 30-54.
42. Korotaj, T., Chen, J., M., & Kurnoga, N. (2023). Hierarchical Clustering of CEE Countries According to Educational and Labour Market Indicators. *Proceedings of the 17th International Symposium on Operational Research* (SOR'23, Bled, Slovenia), 270-276.
43. Kriegel, H.-P., Kröger, P., & Zimek, A. (2009). Clustering high-dimensional data. *ACM Transactions on Knowledge Discovery from Data,* *3*(1), 1-58. https://doi.org/10.1145/1497577.1497578
44. Krpan, M., Gardijan Kedžo, M., & Žmuk, B. (2023). Exploring the link between education length and employment outcomes among youth in Europe: A hierarchical clustering

approach. *Business Systems Research*. *14*(2), 190-213. https://doi.org/10.2478/bsrj-2023-0019
45. Krzaklewska, E. (2013). Visegrad Youth – Comparative review of the situation of young people in the V4 countries. *Council of Europe*. Warszava, Poland.
46. Kunić, J. (2022). The Future of the European Union and the European Integration. In Dimitrijević, D., & Mileski, T. (Ed.). *International Organizations: Serbia and Contemporary World* (1 (2022), 448-457). https://doi.org/10.18485/iipe_ioscw.2022.1.ch26
47. Lénárt, A. (2021). Hungary. Profile commissioned by EASNIE for the Global Education Monitoring Report 2021 – Central and Eastern Europe, the Caucasus and Central Asia – Inclusion and education: All means all. Retrieved from https://gem-report-2020.unesco.org/wp-content/uploads/2021/02/Hungary.pdf
48. Lenhart, O. (2017). The impact of minimum wages on population health: evidence from 24 OECD countries. *European Journal of Health Economics, 18*(8), 1031-1039. https://doi.org/10.1007/s10198-016-0847-5
49. Long, B., T. (2014). The Financial Crisis and College Enrollment: How Have Students and Their Families Responded? In Brown, J., R., & Hoxby, C., M. (Ed.). *How the Financial Crisis and Great Recession Affected Higher Education* (pp. 209-233). Chicago, United States: University of Chicago Press. Retrieved from https://www.nber.org/system/files/chapters/c12862/c12862.pdf
50. Macedo, E., Nairz-Wirth, E., Araújo, H. C., & Szalai, J. (2020). Drawing lessons from early school leavers in a social justice context: Introductory comments. *European Educational Research Journal, 19(5)*, 387-397. https://doi.org/10.1177/1474904120951658
51. Marimont, R., B.; & Shapiro, M., B. (1979). Nearest neighbour searches and the curse of dimensionality. *IMA Journal of Applied Mathematics*. *24(1)*, 59-70. https://doi.org/10.1093/imamat/24.1.59
52. Mascherini, M. (2018). Origins and future of the concept of NEETs in the European policy agenda. In O'Reilly, J., Leschke, J., Ortlieb, R., Seeleib-Kaiser, M., & Villa, P. (Ed.). *Youth Labor in Transition: Inequalities, Mobility, and Policies in Europe* (pp. 503-529). Oxford Academic. Retrieved from https://academic.oup.com/book/26877/chapter/195940054
53. Mehmetaj, N., & Zulfiu Alili, M. (2020). Youth Graduates Employment Country Studies of Albania and North Macedonia in Economics Programs. *ENTRENOVA — ENTerprise Research InNOVAtion,* 6(1), 275-290. Retrieved from: https://econstore.eu/bitstream/10419/224695/1/26-ENT-2020-Mehmetaj-275-290.pdf
54. Müller, A., C., & Guido, S. (2017). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. Sebastopol, Cal.: O'Reilly Media.
55. Neagu, G., Lendzhova, V., & Keranova, D. (2021). The Specifics of the NEETs Population in Bulgaria and Romania. *Postmodernism Problems, 11*(2), 112-123. https://doi.org/10.46324/PMP2102112
56. Nič, M. (2016). The Visegrád Group in the EU: 2016 as a Turning-Point? *European View*, *15*(2), 281-290. https://doi.org/10.1007/s12290-016-0422-6
57. Nielsen, F. (2016). Hierarchical clustering. In *Introduction to HPC with MPI for Data Science* (195-211). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-21903-5_8
58. O'Higgins, N., & Moscariello, V. (2017). *Labour market institutions and youth labour markets: Minimum wages and youth employment revisited*. Employment Working Paper No. 223. Geneva, Switzerland: International Labour Office, Employment Policy Department.
59. Olaode, A., Naghdy, G., & Todd, C. (2014). Unsupervised Classification of Images: A Review. *International Journal of Image Processing*, *8*(5), 325-342. https://www.cscjournals.org/library/manuscriptinfo.php?mc=IJIP-918
60. Patzina, A., & Wydra-Somaggio, G. (2020). Early Careers of Dropouts from Vocational Training: Signals, Human Capital Formation, and Training Firms. *European Sociological Review, 36*(5), 741-759. https://doi.org/10.1093/esr/jcaa011

61. Piatkowski, M. (2015, June 12). *Four ways Poland's state bank helped it avoid recession*. Retrieved from https://www.brookings.edu/articles/four-ways-polands-state-bank-helped-it-avoid-recession/
62. Portela Pruaño, A., Rodríguez Entrena, M., J., Torres Soto, A., & Nieto Cano, J., M. (2022). Why vulnerable early school leavers return to and re-engage with education: Push and pull reasons underlying their decision. *Intercultural Education, 33*(2), 156-172. https://doi.org/10.1080/14675986.2021.2018209
63. Rambla, X., & Scandurra, R. (2021). Is the distribution of NEETs and early leavers from education and training converging across the regions of the European Union? *European Societies, 23*(5), 563-589. https://doi.org/10.1080/14616696.2020.1869282
64. Sakellariou, A., Sanoudou, D. & Spyrou, G. (2012). Combining multiple hypothesis testing and affinity propagation clustering leads to accurate, robust and sample size independent classification on gene expression data. *BMC Bioinformatics, 13*, 270. https://doi.org/10.1186/1471-2105-13-270
65. Schmarje, L., Santarossa, M., Schröder, S.-M., & Koch, R. (2021). A Survey on Semi-, Self- and Unsupervised Learning for Image Classification. *IEEE Access*, *9*, 82146-82168. https://doi.org/10.1109/ACCESS.2021.3084358
66. The Slovenian Press Agency – STA (2020, August 24). *Slovenia Has Smallest Difference Between Minimum & Average Wages in EU*. Retrieved from https://www.total-slovenia-news.com/business/6818-slovenia-has-smallest-difference-between-minimum-average-wages-in-eu
67. Supreme Court of the United States (1921). New York Trust Company v. Eisner. *United States Reports*, *256*, 345-350.
68. Toderiţă, A., Damian, A., & Meiroşu, C. (2019). *Access of young vulnerable NEETs to Youth Guarantee measures in Romania. Current status and way forward.* CRPE Policy Memo no. 76, 2019. Retrieved from https://www.crpe.ro/wp-content/uploads/2019/09/CRPE-Policy-Memo76_Access-of-young-NEETs-coming-from-vulnerable-groups-to-YG-1.pdf
69. Tudor, S., Cilan, T., F., Năstase, L., L., Ecobici, M., L., Opran, E., R., Cojocaru, A., V. (2023). Evolution of interdependencies between education and the labor market in the view of sustainable development and investment in the educational system. *Sustainability*. *15*(5), 3908. https://doi.org/10.3390/su15053908
70. Valkenborg, D., Rousseau, A.-J., Geubbelmans, M., & Burzykowski, T. (2023). Unsupervised learning. *American Journal of Orthodontics & Dentofacial Orthopedics, 163*(6), 877-882. https://doi.org/10.1016/j.ajodo.2023.04.001
71. Vasilescu, M., D., Stănilă, L., Popescu, M., E., Militaru, E.; & Marin, E. (2024). Using panel data clustering regression analysis to revisit income inequalities in the European Union. *Applied Economics Letters*, 1-6. https://doi.org/10.1080/13504851.2024.2358186
72. Vodă, A., I., Bercu, A-M., & Sebestova, J. (2021). Is a higher minimum wage associated with a higher youth employment rate? A panel data analysis. *European Journal of International Management, 16*(4), 631-646. https://doi.org/10.1504/EJIM.2021.118569
73. Watson, D., S. (2023). On the Philosophy of Unsupervised Learning. *Philosophy & Technology*, *38*, 28. https://doi.org/10.1007/s13347-023-00635-6
74. Williams, N., & Mills, J., A. (1998). Minimum wage effects by gender. *Journal of Labor Research, 19*(2), 397-414. https://doi.org/10.1007/s12122-998-1022-2
75. Xu, N., Finkelman, R., B., Dai, S., Xu, C., & Peng, M. (2021). Average Linkage Hierarchical Clustering Algorithm for Determining the Relationships between Elements in Coal. *ACS Omega*, *6*(9), 6206-6217. https://doi.org/10.1021/acsomega.0c05758
76. Youth Wiki (2020). Youth policies in Croatia – 2020. European Commission. Retrieved from https://national-policies.eacea.ec.europa.eu/sites/default/files/2021-09/Croatia_2020_0.pdf
77. Youth Wiki (2023a). Youth Policy Governance, National youth strategy: Slovenia. Retrieved from https://national-policies.eacea.ec.europa.eu/youthwiki/chapters/slovenia/13-national-youth-strategy
78. Youth Wiki (2023b). Employment and Entrepreneurship, Integration of Young People in the Labour Market: Bulgaria. Retrieved from https://national-

policies.eacea.ec.europa.eu/youthwiki/chapters/bulgaria/36-integration-of-young-people-in-the-labour-market

79. Youth Wiki (2023c). Employment and Entrepreneurship, Integration of Young People in the Labour Market: Estonia. Retrieved from https://national-policies.eacea.ec.europa.eu/youthwiki/chapters/estonia/36-integration-of-young-people-in-the-labour-market
80. Youth Wiki (2023d). Employment and Entrepreneurship, Integration of Young People in the Labour Market: Latvia. Retrieved from https://national-policies.eacea.ec.europa.eu/youthwiki/chapters/latvia/36-integration-of-young-people-in-the-labour-market
81. Youth Wiki (2023e). Employment and Entrepreneurship, Integration of Young People in the Labour Market: Lithuania. Retrieved from https://national-policies.eacea.ec.europa.eu/youthwiki/chapters/lithuania/36-integration-of-young-people-in-the-labour-market
82. Youth Wiki (2023f). Education and Training, Preventing early leaving from education and training (ELET): Hungary. Retrieved from https://national-policies.eacea.ec.europa.eu/youthwiki/chapters/hungary/63-preventing-early-leaving-from-education-and-training-elet
83. Youth Wiki (2023g). Education and Training, Preventing early leaving from education and training (ELET): Romania. Retrieved from https://national-policies.eacea.ec.europa.eu/youthwiki/chapters/romania/63-preventing-early-leaving-from-education-and-training-elet

About the authors

Tomislav Korotaj is a Teaching and Research Assistant at the Department of Statistics, Faculty of Economics and Business, University of Zagreb, Croatia. He holds a bachelor's degree in business economics and a master's degree in accounting and auditing from the University of Zagreb. He teaches the following courses: Statistics, Business Statistics, and Methods of Multivariate Analysis. He is pursuing a PhD in Economics, focusing on the youth labour market, wage policies, and income inequality. In addition to the above, his research interests include cluster analysis, panel data models, and school-to-work transition. The author can be contacted at **tkorotaj@efzg.hr**.

James Ming Chen is Professor of Law and Justin Smith Morrill Chair in Law at Michigan State University, U.S.A. He holds a degree in law, magna cum laude, from Harvard University and a master's degree in data science from Northwestern University. He is a member of the American Law Institute and a senior fellow of the Administrative Conference of the United States. He is licensed to practice law in Virginia and the District of Columbia. His research interests include law (particularly administrative law, competition law, and economic regulation), machine learning, mathematical and behavioural finance, and computational social science. The author can be contacted at **chenjame@law.msu.edu**.

Nataša Kurnoga is a Full Professor (Tenured) at the Department of Statistics, Faculty of Economics and Business, University of Zagreb, Croatia. She teaches the following courses at different study levels: Statistics, Business Statistics, Quantitative Methods, and Methods of Multivariate Analysis. She has participated in various scientific research projects and published numerous papers in scientific journals, conference proceedings, and books. She is a reviewer for numerous international scientific journals. Her research interests include business and economics statistics, especially multivariate data analysis methods. The author can be contacted at **nkurnoga@efzg.hr.**

sciendo

Strategic Categorization of Dairy Cow Farms in Croatia using Cluster Analysis

Maja Petrač, Krunoslav Zmaić
Faculty of Agrobiotechnical Sciences Osijek, University J.J. Strossmayer, Croatia
Jaka Žgajnar
Biotechnical Faculty, University of Ljubljana, Slovenia

Abstract

Background: The milk processing sector in the Republic of Croatia faces numerous challenges. It is a distinctly bipolar structure, with some entities resembling the largest milk producers in the EU, while many small and medium-sized dairy farms struggle to remain competitive and achieve further progress. To formulate effective policy, it is important to differentiate between these types and address their key challenges. **Objectives:** The aim is to find the most representative solution that will help us define typical dairy farms and upgrade a SiTFarm tool (Slovenian Typical Farm Model), enabling us to assess the situation in Croatia comprehensively. **Methods/Approach:** Cluster analysis was conducted using empirical data obtained from the Croatian Agency for Agriculture and Food. The analysis involved applying both hierarchical and non-hierarchical clustering techniques. **Results:** Two cluster analysis scenarios are presented, differing in the variables used. In each scenario, 16 relatively homogeneous clusters of farms were obtained. Diversity was minimized within these clusters, and they effectively explain the dairy business in Croatia. **Conclusions:** The results of this analysis thus represent an important starting point for further analysis of the dairy sector in Croatia. These findings could help policymakers identify the types of farms that would benefit most from targeted investments to enhance efficiency, economic viability, and environmental sustainability.

Keywords: cluster analysis; typical farms; dairy sector; mathematical programming; farm model

JEL classification: C38, Q13
Paper type: Research article

Received: May 12, 2024
Accepted: Jul 2, 2024

Citation: Petrač, M., Zmajić, K., & Žgajnar, J. (2024). Strategic Categorization of Dairy Cow Farms in Croatia using Cluster Analysis. *Business Systems Research, 15*(2), 140-153.
DOI: doi.org/10.2478/bsrj-2024-0021

Introduction

Over the years, the number of farms, livestock, and milk production in the Republic of Croatia has steadily declined, reflecting a typical consolidation process, as described by Gonzalez-Mejia et al. (2018). Small farms with a few animals are either closed or transferred to arable production (Mijić and Bobić, 2021), there was, however, some increase in larger farms (Očić et al., 2023). As the situation in the milk production sector in the Republic of Croatia is far from good, and the measures implemented obviously do not lead to improvement, a more detailed analysis of the sector is needed. Očić et al. (2023) see the limited number of competitive farms as the main challenge facing the Croatian dairy sector. Thus, they advocate for various agricultural and rural development measures aimed at empowering especially smaller farms and enhancing the overall competitiveness of the dairy sector. To achieve this goal, the development of appropriate models for a more detailed analysis of the sector and support, especially policy decision-makers, is a logical follow-up. So, different models are needed for decision-makers to have a better insight into what is happening on certain types of agricultural holdings and for their decisions to be based on facts (Ciaian et al., 2013).

The common agricultural policy (CAP) strategic plans (CSP) of the EU member states place increasing emphasis on the use of models that enable simulation at the farm level or the level of the selected aggregate (Lovec et al., 2020). They suggest that each member state should choose its agricultural policy priorities and, in accordance with common EU principles, determine the type, allocated funds, and scope of individual measures.

In such a manner, models become the main tool for generating scenarios or 'what-if' analyses because the effects of policies differ by farm type. It can simulate how a particular scenario, for example, a change in agricultural resource prices or agricultural or environmental policy, may affect a set of performance indicators (Ciaian et al., 2013). In the farm models, mathematical programming is most often applied, as well as models based on econometric and simulation approaches Pečnik and Žgajnar (2022). As it is impossible to carry out the analysis at the level of every agricultural farm, it is carried out at the level of typical representatives. Farms should be classified into smaller categories with common characteristics, which are called typical agricultural farms (Alvarez et al., 2018). These are generally real or hypothetical farms that best represent the situation in a certain segment of a particular sector (representative households) and allow generalization at the aggregate level. Poczta et al. (2020) identified five different types of dairy farms based on a cluster analysis of EU-wide Farm Accountancy Data Network (FADN) data. Based on average indicators, Croatian farms were classified into the first type, along with Slovenia, Austria, Poland, and Romania.

Additionally, Očić et al. (2023), using FADN data, analysed the dairy sector in Croatia, employing a different approach from the one utilized in our study. In their study, which relied on FADN data, they designed three types of farms focused only on the number of cows in the herd. Our goal, however, is to delve even deeper into researching the sector and categorize it into several types, enabling us to provide additional support to different stakeholders. The potential and structure of dairy herds within the sector vary significantly, and tailored solutions are needed to enhance efficiency.

Before the actual creation and application of the farm model in the dairy sector in the Republic of Croatia, it is necessary to define typical farms focusing on dairy production. These farms should be categorized according to main common characteristics and production endowments. Building on these insights, we will establish the initial parameters for production plans per each farm type. With the assistance of experts and stakeholders, we will refine and calibrate these plans, and all needed technological parameters to reflect the realities of the sector accurately. Therefore, the main purpose of this work is to obtain typical farms by applying cluster analysis based on available data following the example of similar studies. It does not require any assumptions about the number and structure of the categories into which the data will be distributed, but the categorization is done based on similarity between the data of farms. The aim is to find the most representative solution that will help us define typical dairy farms and, in a further step, upgrade a SiTFarm (Slovenian Typical Farm Model) tool, enabling us to assess the milk production situation in Croatia comprehensively. Namely, SiTFarm is a microsimulation tool based on mathematical programming and serves as an example of a bioeconomic farm model (BEFM). It enables various analyses at the level of the agricultural production plan, with results that can also be aggregated at the sector level (Žgajnar and Kavčič, 2024). The primary purpose of SiTFarm is to facilitate analyses from the perspective of income sustainability for typical farms that are representative of a certain number of real farms. This model does not require FADN data, allowing for more detailed analyses of smaller farms that would otherwise not be included in the sample. The model calculates various economic indicators and accommodates different CAP interventions at multiple levels and under various conditions, considering the socio-economic context of the analysis.

Dairy farms are integral to the milk sector, and their development directly influences milk production. As emphasized by Parzonko et al. (2024), there is a wide diversity of dairy farms across EU countries concerning the scale and technology of milk production, and this applies to Croatia as well. Cluster analysis will be conducted for both family farms and farms owned by legal entities, covering the entire milk sector in the Republic of Croatia. In the future, this will also enable a comprehensive analysis of all three aspects of sustainability, where significant differences exist between individual types of farms, as well as regions and production conditions.

In the continuation of the paper, we first present the database used for the analysis, including brief descriptive statistics. A description of the methodology employed follows this. In the results chapter, we analyse the possibilities and the influence of various variables on the design of typical farms. We conclude with key findings and guidelines for future work.

Methodology

Data

Empirical data was obtained from the Croatian Agency for Agriculture and Food (Agency). This data is collected from all agricultural holdings in Croatia that deliver milk. Farmers report in a standardized format, including dates and other relevant details. This data is at the farm level. The database contains a list of all registered farms in the Republic of Croatia engaged in milk production.

The database consisted of 4198 dairy cow farms that supply milk in the Republic of Croatia. There was no comprehensive farm database in the Republic of Croatia that contained the information we needed (e.g., number of cows, amount of milk delivered, breed structure, area under cultivation, main crops, land structure, number of employees, location). Unfortunately, not all the needed data were available, so we worked with the data that were accessible to us. Thus, the resulting database was combined from the available date of the Agency dated 2022, with issues such as multiple identifiers and the removal of duplicates and some mistakes, such as farms with zero cows. After arranging the resulting database (connecting data from different farm databases, removing duplicates and inactive farms, etc.), 3393 farms remained for analysis. Of the total number of farms, 3331 farms are family farms, while 62 farms have the status of a legal entity and constitute a special category. These are large, highly specialized dairy farms playing a key role in milk production. Among the family farms, there are also a few large-scale milk producers.

As can be seen from Table 1b, it is a very diverse group of dairy farms. We have chosen four variables that clearly show the differences between the two categories of farms (Table 1a). We tracked the number of cows (NOC) in the herd, representing the average annual stock in the year (2022). Milk production (ADOM) was monitored on the farm level, including the annual production of milk supplied to dairies. Additionally, we monitored agricultural production on the arable land, which is essential for fodder and cash crop cultivation. These crops often supplement or, in some cases, even replace milk production. This trend is expected to continue, especially as further structural changes occur. As highlighted in the introduction, many small and medium dairy farms often shift to arable production and abandon milk production (Mijić and Bobić, 2021). Further, we included two additional variables: the number of plant cultures (NOPC) and the area under cultivation (AUC). These variables shed light on the farm's land use, feeding strategies, and development potential, providing a more comprehensive characterization of the farms.

Table 1a
List of variables at the farm level

Variable Name	Variable
NOC	Number of cows
ADOM	Annual delivery of milk (kg)
NOPC	Number of plant cultures
AUC	The area under culture (ha)

Source: Author's work

Table 1b
Descriptive statistics for dairy farms

Variable Name	Mean	SD	Min	Max	Q1	Q2	Q3
			Family farms (n = 3331)				
NOC	14.78	20.47	1	456	5	10	18
ADOM	73257	146015	21	2799071	14438	34840	76882
NOPC	6.29	2.27	1	18	5	6	8
AUC	23.68	31.68	0.15	469.05	7.64	14.1	27.4
			Legal entity dairy farms (n = 62)				
NOC	337.74	682172	1	4051	28	88.5	369.75
ADOM	2853309	5996202	9938	34601195	124963.25	473117	3111024
NOPC	6.1	3.28	1	21	4	6	7
AUC	2076.19	12946.85	2.38	101427.88	31.06	104.84	306.78

Source: Author's work

The variables and their descriptive statistics are presented in Table 1b. The first category of farms includes 3331 family farms that submitted milk and were included in the resulting register as such (Table 1b). As can be seen from Table 1b, the first category includes farms that raise an average of 14.78 dairy cows, with the median being 10 cows. On 75% of these farms, the number of dairy cows is less than 18, while only 25% of the farms have more than 18 animals. High variability is indicated by a high standard deviation (SD) value. The same applies to milk production. Here, the differences are even greater because, in addition to herd size, lactation milk yield, which reflects the intensity of production, also plays an important role. This indicator further underscores the high variability of family dairy farms. The median is 34840 kg of milk per average farm, with 75% of farms producing less than 76882 kg of milk per year. We can observe that on family farms engaged in milk production, the number of agricultural activities averages 6.3. The variability here is significantly lower, indicating that, in most cases, it involves home-produced fodder. However, in 25% of cases, we can also find examples of more diversified production plans where the number of plant cultures exceeds 5. The last indicator, the area under cultivation, again shows greater variability. On average, dairy farms cultivate 23.68 ha, with a median of 14.1 ha, and the largest 25% of farms more than 27.4 ha.

The second category is 62 farms that have the status of a legal entity (Table 1b). In this example, we observe that the average herd size is notably higher, with 337.74 dairy cows per herd. Additionally, among the largest farms, 25% have a herd size exceeding 369.75, and 75% exceed 28 dairy cows. These are particularly large farms that play a pivotal role in the dairy sector. Annual milk production correlates closely with herd size, indicating significant variability among farms in this aspect. This variability suggests the presence of both intensive forage farms and those employing economies of scale with extensive forage from grasslands. Furthermore, these farms cultivate an average of 6 crops, although there is slightly higher variability compared to family farms. As indicated in the Table 1b, these farms are notably large, averaging 2076 ha of cultivated land.

Cluster analysis and scenarios

The cluster analysis was first performed in relation to all the mentioned variables. First, hierarchical (agglomerative) grouping was performed using Ward's method. Agglomerative grouping starts from a single object, in our case, from a dairy farm. In the first step, each dairy farm forms one cluster. The two most similar dairy farms are grouped into one cluster. Then, a new dairy farm is added to that cluster, or the other two dairy farms are grouped into a new cluster. The merging continues according to mutual similarities until all subgroups are merged into one cluster. The matrix of similarity (distance) between all objects (dairy farms) is the basis for this method. It is a symmetrical matrix with dimensions $n \times n$.

One of the methods used to determine the similarity between clusters is Ward's method or the minimum variance method. This method minimizes the sum of squares between any two clusters that could be formed. Agglomerative grouping is more often used in practice because it is more implemented in computer programs than, for example, divisive grouping. The graphical presentation of the results of agglomerative and divisive grouping can be graphically presented in the form of a two-dimensional hierarchical diagram, known as a dendrogram.

After hierarchical clustering, non-hierarchical clustering was performed, i.e., the k-means algorithm (Scitovski and Sabo, 2020). The k-means algorithm belongs to the optimization algorithms of non-hierarchical clustering that enables the redistribution of objects. It organizes objects (dairy farms) into a predetermined number of clusters, k, and then iteratively reassigns objects to clusters until a specified numerical criterion is met. Achieving the criteria is closely related to achieving the goal of cluster analysis, which consists of finding as compact and better-separated clusters as possible. Therefore, the goal is to minimize the distance between objects within a cluster and maximize the distance between the clusters themselves.

Given that all variables are numeric, the squared Euclidean distance $d(x, y)$ is defined as

$$d(x,y) = \sum_{i=1}^{n} (x_i - y_i)^2 \qquad (1)$$

where $x, y \in S$, S is the data set, $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$, $n \in \mathbb{N}$, and x_i, y_i, $\forall i \in \{1,2,\dots,n\}$, are numeric variables.

All algorithms were applied to standardized z-score so that all variables had equal weight in the cluster analysis. Let $x^{(1)}, x^{(2)}, \dots, x^{(n)} \in S$ be any objects $1 \le k \le n$. The formula calculates the Z-score.

$$z_k = \frac{x_k{}^{(j)} - \mu_k}{\sigma_k} \quad , \qquad 1 \le j \le n, \qquad (2)$$

where μ_k is arithmetic mean and σ_k is the standard deviation. Among the many solutions, two final solutions were chosen for each category. Cluster analysis was conducted for each category using the IBM SPSS Statistics V22.0 software package.

Cluster analysis solutions are not unique and depend on various elements of the analytical procedure, such as the choice between hierarchical or non-hierarchical methods and different algorithms within each method. The selection of variables used to measure similarity also influences the solution. Therefore, it is important to consider the impact of each decision carefully when choosing variables and conducting cluster analysis.

In this paper, two cluster scenarios are analysed for each category. The first scenario describes the clusters in which all the mentioned variables (Table 1a) are included separately for family farms (Table 2) and legal entities (Table 3). In the second scenario, only two variables, NOC and ADOM, were included for both categories (family farms Table 4 and legal entities Table 5).

The utilization of two scenarios serves the purpose of analysing the disparities and the typology of farms that emerge as prototypical representatives within their respective groups. Simultaneously, we will monitor which types of agricultural holdings are more rational from a professional standpoint and will serve as the foundation for designing typical farms for further modelling with the SiTFarm tool.

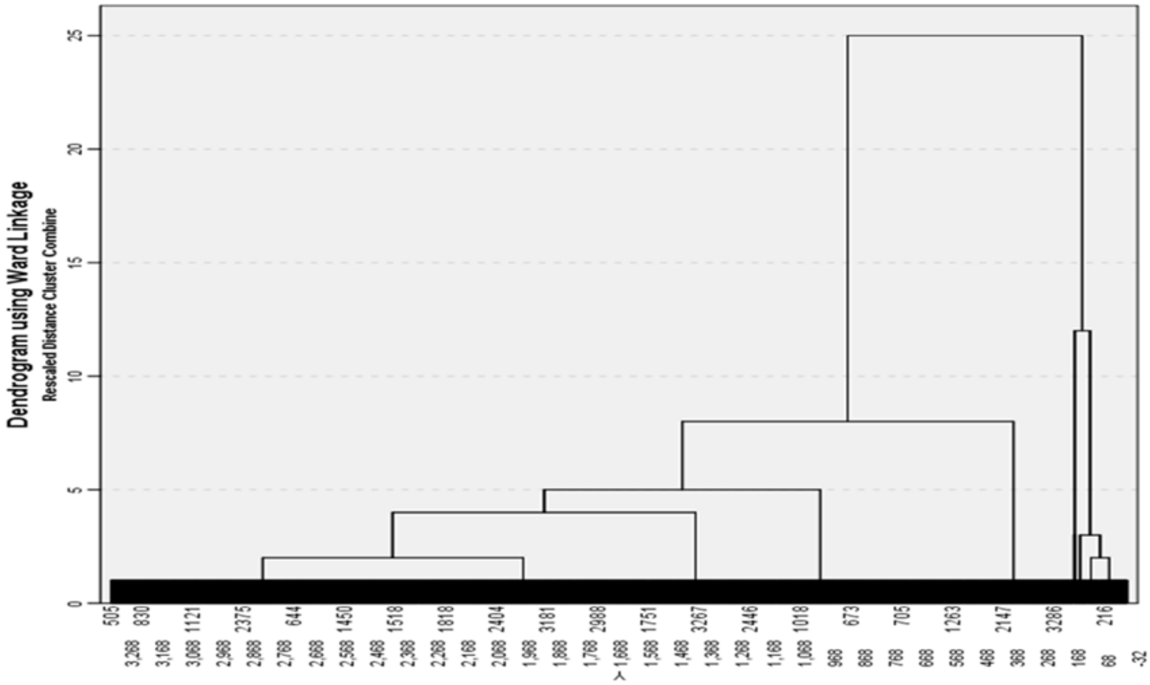
Results

Scenario 1: Four variables

Family dairy farms

The subsequent step in the analysis is to determine the optimal number of clusters using the dendrogram. While the dendrogram assists in this determination, researchers and experts generally make the final estimation. As depicted in the dendrogram, the analysis suggests the presence of 10 clusters (Figure 1). The data structure was analysed, revealing distinct categories of farms with shared characteristics. As the number of clusters increases, the heterogeneity within clusters decreases. Following the hierarchical analysis (dendrogram), a non-hierarchical algorithm (k-means) was applied to the data, with a maximum of 10 clusters imposed. Table 2 presents the composition of all clusters post k-means implementation, illustrating the distribution of farms across the 10 clusters. It delineates which farms belong to each cluster, their respective characteristics, the number of farms within each cluster, and the average number of cows per cluster.

Figure 1
Dendrogram – Scenario 1 – Family dairy farms



Source: Author's work

In Scenario 1, with four variables, 45% of farms in the category of family farms belong to Cluster 1. The average number of cows in this cluster (7.70) is smaller than the average number of cows in the Republic of Croatia (14.78), the average annual milk delivery (30107 kg) is lower than the national average of 73275 kg, the average area of land per farm (10.67 ha) is also smaller than the national average. This implies that cluster 1 consists of very small dairy farms. In some cases, these are also semi-subsistence farms. Cluster 2 is also relatively similar to Cluster 1, but farms in Cluster 2 have much more land and cash crops than farms in Cluster 1. Namely, there are many farms with a few cows and much arable land, and these are usually not farming whose primary activity is milk production. We observe significant variability in the number of cows (NOC) across individual clusters, with relatively common minimum values, particularly within the first five clusters and partially in the sixth cluster.

However, the average values vary considerably, indicating distinct farming technologies and production intensity. Notably, clusters four through six comprise farms involved in milk production on a semi-professional level, despite the relatively low average production per dairy cow, which limits the scope for major investments needed for further growth and development of these farms. From cluster 6 onward, the milk yields of the dairy cows are sufficient to indicate production development potential. This is also noted by Žgajnar and Kavčič (2024) in their study on Slovenian dairy farms. Additionally, livestock density tends to be higher in these clusters. The cultivated areas stand out in certain clusters (2, 3, and 5), reflecting a significantly larger amount of cultivation and, consequently, a significantly lower livestock density.

This suggests that within these clusters, there are also farms where milk production is not the primary economic activity. However, on a certain share of farms, the rationale may also stem from extensive fodder production and low livestock density. To substantiate the latter, additional analyses will be necessary, wherein we will endeavour to assess the impact of the region. Furthermore, based on available data from other sources and expert assessment, we aim to determine the production potential of the land. For the stated reason, the variables AUC and NOPC are excluded from the analysis in Scenario 2 with two variables. The two largest farms are in special clusters (clusters 9 and 10). In both cases, the number of cows is very high. The essential difference between them is that one represents intensive management, while the other, on the contrary, exemplifies extensive cultivation with a significantly lower milk yield. There are two types of dairy farms, as mentioned by Gonzalez-Mejia et al. (2018) in their study, which differ in intensity, and each strives for profitability in milk production in its way.

Table 2
Cluster structure – Scenario 1 - Family dairy farms

Clusters	Variables								
	Number of farms	Average NOC	Min NOC	Max NOC	Average ADOM	Average NOPC	Average AUC	Yield per cow	Land area per cow
1	1496	7.70	1	40	30107.85	4.56	10.67	3973.20	2.09
2	873	8.41	1	25	33634.14	7.84	15.13	4020.36	2.70
3	310	17.85	1	53	77299.30	10.24	44.31	4409.71	3.57
4	464	24.99	6	56	122814.55	5.96	34.39	5096.74	1.69
5	25	46.52	8	111	216429.72	8.80	220.43	4831.20	8.04
6	124	53.60	23	124	342978.66	7.10	70.73	6569.65	1.45
7	27	98.26	64	157	738 311.22	6.78	110.40	7793.64	1.19
8	10	201.50	153	238	1543495.06	6.80	192.35	7671.41	0.95
9	1	317.00	317	317	2799071.00	8.00	469.05	8829.88	1.48
10	1	456.00	456	456	2560156.00	6.00	370.27	5614.38	0.81
Σ	**3 331**	**14.78**	**1**	**456**	**73257.69**	**6.29**	**23.68**	**4329.80**	**2.34**

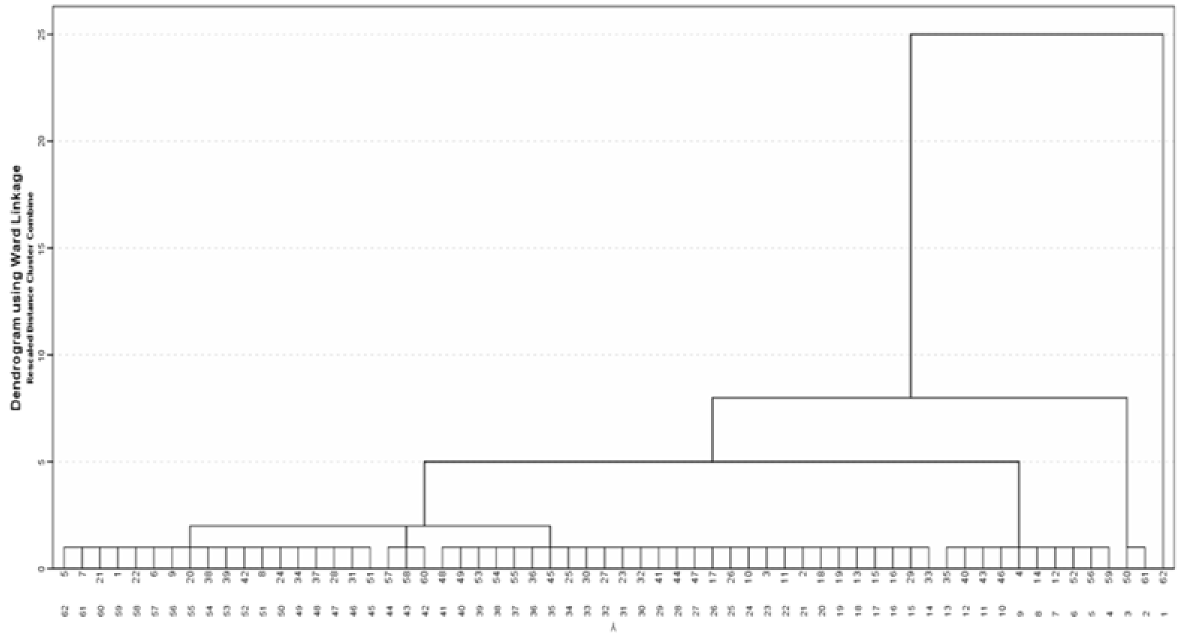
Source: Authors work

Legal entity in dairy production

In the milk processing sector, larger farms are of key importance in Croatia. However, there are also large differences between them in terms of the number of dairy cows in cultivated areas and production potential. The dendrogram for legal entity suggests 6 clusters (Figure 2). Following the hierarchical analysis (dendrogram), a non-hierarchical algorithm (k-means) was applied to the data, limited to a maximum of 6 clusters. Table 3 shows the structure of all clusters in the category of legal entity after the implementation of k-means.

In scenario 1, considering four variables, 61% of farms in the category of legal entities belong to cluster 1. The average number of cows in this cluster (66) is lower than the average number of cows of legal entities in the Republic of Croatia (338). In this cluster, the average annual milk delivery is 436356 kg. This implies that cluster 1 consists of relatively small agricultural holdings. So, it is quite like cluster 4 in family farms (Table 2), except that these farms achieve a significantly lower lactation milk yield per cow (5627 kg). On these farms, the intensity of production is also modest, which is also reflected in the lower livestock density. However, it is notable that, aside from the first cluster, there are no significant differences in the average milk yield achieved in other clusters. Cluster 2 contains farms with a higher average number of cows (246 cows) in a herd. Cluster 3 is a large dairy farm that, due to its technology, has a significantly higher milk production per cow (9073 kg). The three largest farms in the country are each in special clusters (clusters 4 and 6). These are farms with over 2000 cows, and they belong to the category of largest farms on the EU scale (Parzonko et al., 2024). The fourth cluster is interesting, as it has a distinct livestock density. Given the higher milk production, this suggests that breeding relies on purchased fodder. The largest farm has over 4000 cows and, unlike the farms in cluster 4, has much arable land (101428 ha), and a lot of different crops are cultivated on those lands (21). Hence, the last cluster represents a distinct type, engaging not only in milk production but also in crop production. This is evident from the large surface area per dairy cow, indicating a dual focus on both dairy farming and crop cultivation (Table 3).

Figure 2
Dendrogram – Scenario 1 – Legal entity



Source: Author's work

Table 3
Cluster structure – Scenario 1 - Legal entity

Clusters	Variables								
	Number of farms	Average NOC	Min NOC	Max NOC	Average ADOM	Average NOPC	Average AUC	Yield per cow	Land area per cow
1	38	66.42	1	227	436355.81	4.74	97.33	5626.9	1.68
2	11	245.91	11	754	1996714.73	10	634.38	6562.2	3.36
3	9	631.44	387	1332	5713704.89	5.11	238.35	9073.1	0.48
4	2	2 406.5	2057	2756	21518697	5.5	251.14	8896.6	0.11
5	1	1164	1164	1164	9297902	10	13971.76	7987.9	12
6	1	4 051	4051	4.05	34601195	21	101427.9	8541.4	25.04
Σ	**62**	**337.74**	**1**	**4051**	**2853309.96**	**6.1**	**2076.19**	**6483.7**	**2.3**

Source: Author's work

Scenario 2: Two variables

Family dairy farms

In this scenario, we present the results obtained when only two variables were considered in the analysis: the number of dairy cows (NOC) and the annual delivery of milk (ADOM). A dendrogram was constructed using the same procedure as in Scenario 1, followed by the implementation of the k-means algorithm.

Table 4 shows the structure of all clusters after the implementation of k-means for family dairy farms. Although the dendrogram suggested 7 clusters, the analysis was done with 10 clusters for the sake of comparison with Scenario 1. The obtained analysis (see Table 4) shows that the variable NOC for obtained clusters has a smaller range of disjoint intervals. The same holds for the ADOM variable since it is a correlation variable.

This analysis shows that we get a different set of types of typical farms (Table 4) because this focus is on milk production, it is expected that individual clusters follow the logic of increasing herds, where they are more homogeneous than in the scenario 1 of the first analysis (Table 2). Even in scenario 2, cluster 1 is numerically the most represented, with an average number of 5.7 cows in a herd. It is a category of dairy farms that largely belongs to self-sufficient farms with milk production of 3802 kg of milk. It is also a relatively homogeneous category, with a maximum of 15 animals in the herd. This is a special category that is expected to be abandoned in future structural changes. Appropriate solutions will need to be found to slow down this trend, as these farms are important both from the point of view of social and environmental sustainability. Interestingly, in scenario 2 with two variables, cluster 2 is quite different.

Table 4
Cluster structure – Scenario 2 - Family dairy farms

Clusters	Variables								
	Number of farms	Average NOC	Min NOC	Max NOC	Average ADOM	Average NOPC	Average AUC	Yield per cow	Land area per cow
1	1.889	5.70	1	15	20374.81	5.95	12.98	3801.96	2.96
2	994	16.77	7	32	73871.13	6.68	27.01	4640.77	1.64
3	304	33.27	16	59	168547.71	6.82	43.52	5328.34	1.31
4	101	58.07	30	111	373897.32	7.10	74.78	6719.71	1.30
5	30	96.50	56	157	699186.73	6.87	130.98	7645.72	1.42
6	6	182.17	153	204	1275357.46	7.17	136.78	7197.46	0.73
7	3	204.67	200	214	1636 69.67	7.00	247.45	7999.35	1.22
8	2	230.50	223	238	2103315.93	4.50	204.51	9123.60	0.88
9	1	317.00	317	317	2799071.00	8.00	469.05	8829.88	1.48
10	1	456.00	456	456	2560156.00	6.00	370.27	5614.38	0.81
Σ	**3331**	**14.78**	**1**	**456**	**73257.69**	**6.29**	**23.68**	**4329.80**	**2.34**

Source: Author's work

It is a category of medium-sized farms, which, like cluster 3, includes dairy farms that will face the challenge of restructuring soon. In both categories (2 and 3), we can expect that there will continue to be a cessation of production on one side and, for those with sufficient production resources, an increase in and maintenance of milk production. For the latter, greater investments and appropriate incentives will be needed so that agricultural holdings can follow the trend and achieve higher productivity. This is especially true for cluster 3, where the average production is over 5000 kg of milk. Clusters 4 and 5 combine medium-sized dairy farms with average 58 and 96 dairy cows. It is a type of family farm that is professionally engaged in farming and has development potential. This is also confirmed by the findings of Žgajnar and Kavčič (2024), who analysed a similar type of farm under Slovenian conditions. The next clusters (5 to 10) are numerically rather modestly represented. However, except for the last type of cluster, which includes only one farm, they are significant from the perspective of both the number of cows and the annual delivery of milk.

Legal entity in dairy production

Table 5 displays the composition of all clusters after the implementation of the k-means algorithm for legal entities. Although the dendrogram suggested 5 clusters, an analysis was conducted with 6 clusters for comparison with Scenario 1.

Table 5
Cluster structure – Scenario 2 - Legal entity

Clusters	Variables								
	Number of farms	Average NOC	Min NOC	Max NOC	Average ADOM	Average NOPC	Average AUC	Yield per cow	Land area per cow
1	45	71.73	1	275	467688.46	5.51	135.47	5574.9	2.01
2	12	529.33	333	824	4806974.92	7.08	556.52	9029.6	1.22
3	2	1248	1164	1332	10268474.5	6.5	7009.66	8212.8	6.02
4	1	2057	2057	2057	17657987	5	300.83	8584.3	0.15
5	1	2756	2756	2756	25379407	6	201.45	9208.8	0.07
6	1	4051	4051	4051	34601195	21	101427.9	8541.4	25.04
Σ	**62**	**337.74**	**1**	**4051**	**2853309.96**	**6.1**	**2076.19**	**6483.7**	**2.3**

Source: Author's work

In scenario 2, interesting types of legal entities emerge. Similar to the first scenario (Table 3), the most strongly represented farms are in cluster 1 (Table 5), where the average number of dairy cows in the herd does not deviate significantly from the first analysis, nor does the average milk yield per dairy cow (Table 5). However, it is a fairly variable category depending on the number of dairy cows in the herd. Similar to the first scenario, it will be necessary to establish at least two types of typical farms for this category in the future analysis when developing the starting points for typical farms in the model, as the average value does not adequately represent the variability among farms. The upper quartile, which holds greater significance for milk processing, is certainly of interest. In scenario 2, cluster 2 is similar to cluster 3 of the scenario 1(Table 3). In clusters 3 to 6, there are distinctly large agricultural holdings. There are no significant differences compared to the first analysis.

Conclusion

Summary of research

By utilizing cluster analysis on the data of dairy farms in Croatia, we aimed to identify the characteristics of basic types of dairy farms. This information will facilitate a deeper understanding of the milk production sector in Croatia. Therefore, the results of this research, as suggested by Gonzalez-Mejia et al. (2018), can be used for further research to model scenarios including economic components (e.g. gross margin per cow, gross margin per litre of milk, gross margin per ha), social aspect (e.g. labour) and environmental impacts (e.g. land use per cow, resource depletion, GHG).

The results obtained confirm the distinct polarity of farms in milk production observed in Croatia. In the initial phase, we examined the available data, which encompasses four variables. However, due to the presence of highly atypical dairy farms in this sector (such as those with a small herd and share of milk production but an extremely large area of land), we analyzed two scenarios, with a substantial variation in the variables included. It emerged that the analysis considering solely the herd size and total milk production at the annual level is more rational and yields more homogeneous groups of dairy businesses. This corresponds to scenario 2, focusing on two variables. In certain cases, however, Scenario 1 (four variables) indicates specific types of farms where there is a significant amount of land devoted to cash crops rather than milk production. The latter is less important from the perspective of analyzing the dairy sector, particularly concerning small herds. This is a special category that is expected to be abandoned in future structural changes or transformed into cash crop production. This coincides with the findings of Mijić and Bobić (2021). However, there are also cases where the volume of milking production remains significant, and these farms often have diversified production plans as one of the strategies of risk management.

The disadvantage of the database used in this study is, however, the lack of all the necessary data for a precise analysis of a single typical farm using the farm model. In future analysis, it is recommended to extend the time series and, following the example of Klímová et al. (2022), analyse from a temporal perspective.

Theoretical contributions

Based on the analysed data, to encompass the entire sector, it will be necessary to define at least 15 baseline types of dairy farms in Croatia. It is a significantly larger number than defined in previous research for Croatia, but the approach we intend to apply in the continuation of the research requires a more detailed analysis with a larger number of typical farms.

Of these 15 baseline types, two-thirds will be family farms, ranging from small self-sufficient farms to medium-sized and large family farms. These types are crucial from both social and environmental sustainability perspectives. From the perspective of structural changes, we anticipate that these types of farms will undergo the most significant transformations in the future. This will likely involve consolidation, reflected in a declining number of farms. This is also noted to a certain extent by Mijić and Bobić (2021).

A distinct category will be agricultural companies dealing with milk production, which serve as the pillar of the milk processing sector in Croatia. The analysis suggests that it would be sufficient to have six such production types in this category.

In the next step of defining farm types, given the diverse growing conditions, it would be prudent to include the influence of the region, as conditions vary significantly across regions. Additional indicators, such as the scope of grass, fodder, and cash crops, should be used to provide information on land use and feeding strategies, helping to characterize farms more accurately.

Practical implications

The results of this analysis thus represent an important starting point for further analysis of the dairy sector in Croatia. These will be typical farms that will be defined in more detail at workshops with consultants and experts in the field and will be further adjusted and upgraded with the Slovenian farm model SiTFarm (Pečnik and Žgajnar, 2022). This model will help to evaluate various economic, technological, and environmental indicators with predetermined production constraints.

Indeed, the diverse and varied category of agricultural holdings, as illustrated in this study, achieve very different levels of production efficiency, and face distinct challenges. Given the number of farms in each of the groups, our results corroborate the findings of Očić et al. (2023), who identify the limited number of competitive farms as the primary challenge facing the Croatian dairy sector.

Consequently, addressing these challenges will require tailored agricultural policy measures suited to the specific needs of each category as well as the type of farm. In such a manner, certain policy measures that would increase the profitability of one typical farm do not mean that they would increase the profitability of another farm. So, the same measure is not equally effective for diverse categories of dairy farms. The key question here revolves around identifying which agricultural policy measures effectively address the primary challenges faced by farms, with the aim of achieving greater economic, environmental, and social sustainability. This also suggests a potential improvement in monitoring the effectiveness of individual measures at the farm level, similar to the approach taken by Klímová et al. (2022) in evaluating public support for innovation and measuring changes in the turnover of funded companies. By analysing the time series, we could determine which measures achieve the greatest effectiveness for each type of intervention. Such results could be useful to policymakers in getting information on dairy farm needs, such as which type of measure they need and how much they can adapt to the given situation. Certainly, it will be necessary to support certain types of family farms with appropriate investments to improve efficiency and enhance both economic and environmental sustainability. In the long term, such an approach will help reduce dependence on budgetary payments and increase overall sustainability.

References

1. Alvarez, S., Timler, C. J., Michalscheck, M., Paas, W., Descheemaeker, K., Tittonell, P., Andersson, J. A., & Groot, J. C. J. (2018). Capturing farm diversity with hypothesis-based typologies: An innovative methodological framework for farming system typology development. *PLOS ONE, 13*(5), e0194757. https://doi.org/10.1371/journal.pone.0194757
2. Ciaian, P., Espinosa, M., Gomez Y. Paloma, S., Heckelei, T., Langrell, S., Louhichi, K., Sckokai, P., Thomas, A., & Vard, T. (2013). *Farm level modelling of the CAP: a methodological overview*. Publications Office of the European Union.
3. Gonzalez-Mejia, A., Styles, D., Wilson, P., & Gibbons, J. (2018). Metrics and methods for characterizing dairy farm intensification using farm survey data. *PLOS ONE, 13*(5), e0195286. https://doi.org/10.1371/journal.pone.0195286
4. Klímová, V., Žítek, V., & Lelková, T. (2022). Public Support for Innovation: Changes in Turnover of Granted Companies. *Business Systems Research Journal, 13*(1), 120-137. https://doi.org/10.2478/bsrj-2022-0008
5. Lovec, M., Šumrada, T., & Erjavec, E. (2020). New CAP Delivery Model, Old Issues. *Intereconomics, 55*(2), 112-119. https://doi.org/10.1007/s10272-020-0880-6
6. Mijić, P., Bobić, T. (2021). Milk production and challenges in transition from conventional to robotic milking in Croatia. *Acta Scientiarum Polonorum Zootechnica, 20*(3), 59-64. https://doi.org/10.21005/asp.2021.20.3.08
7. Očić, V., Šakić Bobić, B., Grgić, Z. (2023). Economic analysis of specialized dairy farms in Croatia according to FADN. *Mljekarstvo, 73*(1), 50-58. https://doi.org/10.15567/mljekarstvo.2023.0106
8. Parzonko, A., Parzonko, A. J., & Bórawski, P. (2024). Economic Competitiveness of Dairy Farms from the Top Milk-Producing Countries in the EU: Assessment in 2014-2021. *Agriculture, 14*(1), 123. https://doi.org/10.3390/agriculture14010123
9. Pečnik, Ž., & Žgajnar, J. (2022). Resilience of dairy farms measured through production plan adjustments. *Journal of Central European Agriculture, 23*(1), 207-219. https://doi.org/10.5513/jcea01/23.1.3372
10. Poczta, W., Średzińska, J., & Chenczke, M. (2020). Economic Situation of Dairy Farms in Identified Clusters of European Union Countries. *Agriculture, 10*(4), 92. https://doi.org/10.3390/agriculture10040092
11. Scitovski, R., & Sabo, K. (2020). Cluster analysis and recognition of geometric objects (in Croatian). Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel za matematiku
12. Žgajnar, J., & Kavčič, S. (2024). Understanding the Dairy Sector in Slovenia: A Modeling Approach for Policy Evaluation and Decision Support. *Sustainability, 16*(14), 6009. https://doi.org/10.3390/su16146009

About the authors

Maja Petrač is a university specialist in actuarial mathematics, employed at the Faculty of Agrobiotechnical Sciences Osijek University of Josip Juraj Strossmayer Osijek (Croatia) as a lecturer in the scientific field of mathematics. She is currently a student in the postgraduate doctoral study of Agroeconomics at the Faculty of Agrobiotechnical Sciences in Osijek, and her research interests are mathematical programming and farm models. The author can be contacted at **mpetrac@fazos.hr**.

Krunoslav Zmaić is a Full professor at the Faculty of Agrobiotechnical Sciences Osijek, University of Josip Juraj Strossmayer Osijek (Croatia). He teaches a course on Agroeconomics and Agricultural and Rural Policy. His areas of scientific interest are the application of economic theory in optimizing the production and distribution of food and fibre products, macro-agroeconomic issues of agricultural development, problems of sustainable development of rural areas, and the evaluation of the impact of agricultural and rural policy in the function of integrated rural development. He is actively involved in several other research projects. The author can be contacted at **kzmaic@fazos.hr.**

Jaka Žgajnar is an Associate Professor at the Biotechnical Faculty, University of Ljubljana (Slovenia). He teaches courses on farm management, animal health economics, and the utilization of mathematical programming techniques in agriculture. His research primarily focuses on various mathematical programming techniques for analyzing decision-making in agriculture. He specializes in farm-level modelling and conducting impact assessment analyses of Common Agricultural Policy measures. He is actively involved in several research projects. The author can be contacted at **jaka.zgajnar@bf.uni-lj.si**.

Estimating Asymmetric Fuel Price Responses in Croatia

Karol Szomolányi, Martin Lukáčik, Adriana Lukáčiková
University of Economics in Bratislava, Slovakia

Abstract

Background: According to many studies, the transmission of oil prices to retail fuel prices is asymmetric. Fuel prices react faster if oil prices rise and more slowly if oil prices fall. Different standard econometric procedures lead to different results. The Linex approach, which is based on formulating the non-linear adjustment cost function, reflects the theory. It uses the generalised method of moments to estimate the reaction functions, which demands many observations. **Objectives:** The paper investigates the price asymmetry in the Croatian retail fuel market using standard approaches and the Linex approach. **Methods/Approach:** The simple and dynamic asymmetry models, error correction models, threshold autoregressive co-integration, and the Linex approach are used to verify the hypothesis of asymmetric reactions of gasoline and diesel prices in Croatia. **Results** The results using the standard methods are mixed, while the Linex approach indicates price asymmetry, the size of which is measured with the average price bias. The results correspond to other studies worldwide. **Conclusions** The authors' preferred Linex approach detects price asymmetries, even with large data samples with frequent changes in trends and volatilities. According to the approach, the question is not whether prices are formed asymmetrically but the size of the asymmetry.

Keywords: retail prices; error correction model; threshold autoregressive cointegration; linear-exponential adjustment cost function; generalised method of moments

JEL classification: C26, C51, Q41
Paper type: Research article

Received: Feb 8, 2024
Accepted: Jun 27, 2024

Acknowledgements: The Grant Agency of Slovak Republic - VEGA, supports this paper by grant no. 1/0052/24, "Estimation of Key Structural Parameters of Current Macroeconomic Models".

Citation: Szomolányi, K., Lukáčik, M. & Lukáčiková, A. (2024). Estimating Asymmetric Fuel Price Responses in Croatia. *Business Systems Research, 15*(2), 154-176.
DOI: doi.org/10.2478/bsrj-2024-0022

Introduction

Fuel price shifts, especially their growth, are perceived very sensitively by the population in every country because they represent a critical factor in determining the transport price. On the other hand, transport services enter into the pricing of most other goods and services as a nonnegligible component. Therefore, every gasoline and diesel price increase is perceived very negatively, and a decrease in fuel prices, on the other hand, is usually overlooked. This psychological aspect is often reflected in the opinion that the reaction to the adjustment of gasoline prices is not the same in the case of a drop in oil prices on the stock exchange (or world markets) as in the case of an increase. This effect was named by Bacon (1991) as the "rockets and feathers phenomenon". What causes this perception of reality?

In addition to the population's feelings, four theoretical explanations for asymmetric fuel price-making are known. Borenstein et al. (1997) suggested the first three (a short review is also provided by Brown et al., 2000 and Radchenko, 2005), and Douglas et al. (2010) assume the last.

The first theoretical rationale is called oligopolistic coordination theory. Firms try to assure competitors they keep a tacit agreement by reacting asymmetrically to oil price changes.

The second explanation results from the cost of production and inventory adjustment. Firms spread the adjustment over time because adjusting production and inventory levels is costly (Borenstein et al., 2002).

The search theory is the third theoretical reason why prices are changing asymmetrically. The high oil price volatility allows the firm to take advantage of consumers' high search costs and temporarily increase its margin after the increase in the oil price.

The fourth explanation is the theory of strategic interactions (Okun, 1981). Price makers (including "selfish") try not to pit against themselves a minority of consumers who tend to "punish" firms for unjustified or insufficiently explained price increases (Rotemberg, 2011, p. 953). There is a tendency to use the crude oil price increase to adjust retail fuel prices by additional unexpected costs or the effects of an increase in demand. All the theories should be considered by the chosen methodology investigating price asymmetries.

The primary objective of this paper is to delve into the phenomenon of price asymmetry in the retail fuel market in Croatia. The study employs standard approaches but also introduces a unique method using the Linex adjustment cost function, which the authors have modified explicitly for this study (Szomolanyi et al. 2020, 2022a, 2022b). This innovative approach allows for estimating the average fuel price bias that directly results from asymmetric price making, a topic of significant interest to academic researchers and economists.

Frequent changes in trends and volatilities characterise energy prices. Therefore, many authors tend to split the data sample into more subsamples (Bagnai et al., 2018; Bumpass et al., 2019; Cipcic, 2021). This decision may come at the cost of too much loss of degrees of freedom. The methodological part will demonstrate that the generalised method of moments (GMM) is appropriate for estimating the reaction function of fuel prices derived by the Linex function. The asymptotic properties of GMMs demand larger datasets (a few hundred or more). The methodological part also discusses that the aforementioned theoretical justifications for asymmetric price transition can be mathematically formulated as a non-linear adjustment cost function. Therefore, the Linex approach corresponds to the theoretical starting points.

The research hypothesis is whether fuel prices react asymmetrically, i.e., faster on average to oil price increases than decreases. An alternative option is symmetrical price adjustment; the absolute values of the reactions are, on average, the same. The average price of fuels is higher for asymmetric reactions than for symmetric ones. The average bias is the difference between the average price for asymmetric responses and the average price for symmetrical reactions.

The central research question of this paper is to explore whether a systematic price bias exists in the Croatian retail fuel market. The entire sample of Croatian retail gasoline price data and an asymptotically consistent estimate of the econometric specification of the price reaction function derived from the theoretical basis of price asymmetries will be used. Notably, this approach, novel in the context of retail fuel pricing in Croatia, will be compared with estimates based on econometric methods commonly used in global studies.

The result is a comprehensive, complex analysis of the relationship between Croatia's gasoline and diesel retail prices and the crude oil prices in world markets. The procedures mentioned can also be used to analyse price asymmetry, for example, in agriculture, energy, or other markets. As Deltas et al. (2020) point out, the study's standard research approaches also lead to mixed results. According to the Linex adjustment cost function approach, the retail fuel prices adjust asymmetrically. The corresponding price bias is slightly higher for gasoline.

Retail fuel price decreases are more sluggish than price increases in Croatia. Such price-making systematically produces a bias, which is, on average, positive in Croatia. The average gasoline price bias is higher than diesel in Croatia. This result is in contrast with the European average. The average European retail fuel price biases are higher than those of Croatia.

Literature Review

According to many studies, starting with Bacon (1991), retail fuel prices respond to changes in crude oil prices asymmetrically. However, Perdiguero-Garcia (2013) provided a meta-analysis of price asymmetries in the gasoline market, finding that the significant variation in the outcomes reported makes drawing definitive conclusions difficult. Deltas et al. (2020) observed that using different methods, data samples, and frequencies can yield varying results when examining the pass-through of crude oil prices to retail fuel prices. Some studies split the sample into two or more sets to find evidence for the rockets' and feathers' effects (Bagnai et al., 2018; Bumpass et al., 2019; Cipcic, 2021)

Historically, studies have employed various econometric methods to detect asymmetric responses of output prices to input prices. For instance, Bacon (1991) tested the price asymmetry hypothesis using a quadratic adjustment function. Other approaches used in empirical studies of retail fuel price asymmetries include the distributed lag model (Karrenbrock, 1991), the first differences model (Duffy-Deno, 1996), cointegration techniques (Borenstein et al., 1997), and the vector autoregressive model (Balke et al., 1998; Kang et al., 2019). More recently, authors of empirical studies have utilised threshold models with multiple regimes (Douglas et al., 2010; Bagnai et al., 2018; Torrado et al., 2020; Gosinska et al., 2020), broadening this field's methodological landscape.

The most recent papers continue to show mixed results. Torrado et al. (2020) confirmed retail gasoline price asymmetries in Spain and Germany but not in France from 2011 to 2017. Gosinska et al. (2020) found asymmetric retail fuel pricing in Poland from 2000 to 2016. Bragoudakis et al. (2021) did not confirm asymmetric adjustments of gasoline prices to changes in oil prices in Greece after 2010. Using Greek data immediately after an unannounced and non-negligible increase in consumption taxes in 2010, Genakos et al. (2022) confirmed that the asymmetric gasoline price reaction is higher with lower competitiveness. Asane-Otoo et al. (2022) used daily German data from 2014 to 2018 and pooled-panel asymmetric error correction models to confirm the asymmetries in most cases. The authors note that temporal aggregation of station-level price data leads to inaccurate inferences and could account for the inconclusive findings in the literature.

In the context of the study, it is crucial to note the findings of Cipcic (2021), who investigated the asymmetric reactions of gasoline and diesel prices in several post-communist countries, including Croatia. Her results, spanning from January 2005 to June 2013, did not confirm asymmetry in the countries during the entire period under review. However, her findings did reveal asymmetric fuel price reactions in some countries from January 2009 to June 2013, underscoring the importance of temporal considerations in such analyses.

Methodology

During its evolution, the methodology of analysing asymmetric reactions of business prices has undergone extensive development, accompanied by many exciting ideas or modifications of existing methods, many of which are still used today. Due to its importance in estimating the development of prices, it is a constantly dynamically developing area within economic analyses. It is not technically possible to present all the methods used in one article, so this section categorises them into several primary groups, in which it presents the workhorses of each approach.

The primary group corresponding to the initial period of asymmetry research are Simple Models of Asymmetry, which use dummy binary variables to capture price increases and decreases. These are applied to products with critical factors determining the examined prices. There may be more factors, and the asymmetry effects may persist for extended periods, which is why these models have been dynamised.

The second group consists of Error Correction Models. They were used to reflect the impact of the non-stationarity of data-generating processes on the analysis, particularly the procedures for dealing with them. The possibility to model, in addition to short-term first differences, also the original levels of non-stationary variables meant a significant shift in methodology. This type of model allowed distinguishing between short-term and long-term asymmetric responses. Through gradual development, non-linear versions were also proposed, considering differences other than ordinary ones and solving the non-stationarity of processes.

The third type of model is the Threshold Autoregressive Cointegration Model. Using them, analysts responded to the criticism that, in the case of asymmetry, the classic cointegration test used in the previous type of models is inappropriate because it can lead to incorrect conclusions. Models from this group differ in how they search for the threshold value. For all three mentioned groups, analyses in the form of vector models were proposed, which examine the connection of the investigated commodities in several related markets.

The last of the methods (Linex) presented is an approach modified by the authors using the adjustment costs function. The principle of this approach is based on the idea that changes in economic processes reflecting shifts in input prices are not costless. A non-linear functional form can express asymmetric adjustment costs. In the beginning, four theoretical explanations of asymmetric price adjustment were presented. The adjustment costs function can formulate all four. The authors consider the approach advantageous because it reflects all the known theories of asymmetric price adjustment.

Using the Linex approach and the Netherlands and U.K. data, Pfann et al. (1993) reported that the costs of firing production workers are lower than the hiring costs. However, the opposite is true for the non-production workers (Adda et al., 2003, pp. 243-4). Surico (2007a, 2007b, 2008) used the approach to analyse the U.S. and EMU monetary policy asymmetries.

Simple Models of Asymmetry

First, a simple asymmetry model intuitively uses dummy variables of price increases and decreases. The product of price and each of these dummy variables leads to an estimate of a pair of asymmetry parameters, whose equality is tested by the F test. This procedure has been used since Tweeten et al. (1969).

The basic simple model of price asymmetry for fuel prices has the form:

$$y_t = \beta_0 + \gamma_0^+ x_t^+ + \gamma_0^- x_t^- + u_t \qquad (1)$$

where y_t is regressand and the average weekly retail price of gasoline or diesel in time *t*; x_t^+ denotes the key regressor – the average weekly crude oil price in time *t* equals x_t if its value has increased over the last period and zero otherwise and x_t^- is the average weekly price of oil in time *t* equals x_t if its value has decreased over the last period and zero otherwise. The coefficients γ_0^+ and γ_0^- are the very parameters that, if the hypothesis of their equality is rejected, it means pricing asymmetry. Conversely, if the linear hypothesis of the equality of these parameters cannot be rejected, prices change symmetrically.

In addition to the primary determinant of the investigated fuel price - the oil price in this case, other essential factors influencing the price of fuels can be assumed. Several analyses confirmed the interconnectedness of gasoline and diesel prices in the fuel market, gradually leading to the development of multi-equation models.

The second model, which extends the basic model by an additional explanatory variable, is a model in the form:

$$y_t = \beta_0 + \gamma_0^+ x_t^+ + \gamma_0^- x_t^- + \delta_0 z_t + u_t \qquad (2)$$

where z_t is another regressor (the price of another fuel – the gasoline price in the equation of diesel price and vice versa in time *t*, and the statistical significance of the coefficient δ_0 confirms the validity of the influence of another factor. Asymmetry in models (1) and (2) is present if the null hypothesis of $\gamma_0^+ = \gamma_0^-$ is rejected. The F test can test this linear hypothesis in the linear model.

The last presented models in this group are dynamic models, similar to Karrenbrock (1991), in which instead of the price level, its change is examined:

$$\Delta y_t = \beta_0 + \sum_{i=0}^{s} \gamma_i^+ \Delta x_{t-i}^+ + \sum_{i=0}^{q} \gamma_i^- \Delta x_{t-i}^- + u_t \qquad (3)$$

The duration of the rise response may not be the same as the fall, which will be reflected in the difference between the *s* and *q* values. The cumulative effect of price variation can be tested with the hypothesis: $\sum_{i=0}^{s} \gamma_i^+ = \sum_{i=0}^{q} \gamma_i^-$. The F test can again be used to check for asymmetry.

Error Correction Models

The development of most commodities' prices tends to be non-stationary because their essential characteristics (distribution moments) change over time. For example, the price of the studied commodity grows over time because inflationary pressures act on it so that the average will grow over time. However, the non-stationarity of the processes generating the investigated variables can lead to spurious regressions and conclusions that identify factors that do not influence them as determinants of prices. The solution to this problem was the second model group, the Error Correction Model.

The basis of this methodology is an auto-regressive distributed lag model of order one with two variables:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \gamma_0 x_t + \gamma_1 x_{t-1} + u_t \tag{4}$$

and if other essential factors are assumed, then with three (or more) variables:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \gamma_0 x_t + \gamma_1 x_{t-1} + \delta_0 z_t + \delta_1 z_{t-1} + u_t \tag{5}$$

where y_t is the average weekly retail price of gasoline or diesel in time t; x_t is the average weekly crude oil price in time t; z_t is another relevant regressor in time t (for example, the price of another fuel); u_t is a stochastic term in time t and β_0, β_1, γ_0, γ_1, δ_0, and δ_1 are unknown parameters of this regression model.

The model (4) can be rewritten as the error correction model (Engle et al., 1987):

$$\Delta y_t = \beta_0 + \gamma_0 \Delta x_t + (\beta_1 - 1)\left[y_{t-1} - \frac{(\gamma_0 + \gamma_1)}{1-\beta_1} x_{t-1} \right] + u_t \tag{6}$$

and model (5) as the error correction (ECM) model:

$$\Delta y_t = \beta_0 + \gamma_0 \Delta x_t + \delta_0 \Delta z_t + (\beta_1 - 1)\left[y_{t-1} - \frac{(\gamma_0 + \gamma_1)}{1-\beta_1} x_{t-1} - \frac{(\delta_0 + \delta_1)}{1-\beta_1} z_{t-1} \right] + u_t \tag{7}$$

which contains the original (one period-lagged) variables in the levels (deviations from the long-run equilibrium) and their first differences (the short-run relationship). Suppose a positive unit change of the regressor has an identical influence on the regressand as a negative unit change. In that case, distinguishing between them is not needed. The overall response with one parameter for one regressor, as in the reversible models (6) and (7) can be estimated. If this restriction is not valid, the estimation results can be improved by specifying increases ($\Delta^+ x_t$ and $\Delta^+ z_t$) and decreases ($\Delta^- x_t$ and ($\Delta^- z_t$) of the explanatory variables as separate variables and also by separating the positive and negative deviations from the long-run equilibrium relationship.

The asymmetric irreversible error correction model (Granger et al., 1989):

$$\Delta y_t = \beta_0 + \gamma_0^+ \Delta^+ x_t + \gamma_0^- \Delta^- x_t + \lambda^+ e_{t-1} \times D(e_{t-1} > 0) + \lambda^- e_{t-1} \times D(e_{t-1} \le 0) + u_t \tag{8}$$

where $e_{t-1} = y_{t-1} - \frac{(\gamma_0 + \gamma_1)}{1-\beta_1} x_{t-1}$ is one period-lagged deviation from the long-run equilibrium relationship; $D(e_{t-1} > 0)$ is a dummy variable that equals one if $e_{t-1} > 0$ and equals zero otherwise; $D(e_{t-1} \le 0)$ is a dummy variable that equals one if $e_{t-1} \le 0$ and equals zero otherwise; λ^+ and λ^- are the corresponding adjustment parameters.

The asymmetric irreversible error correction (A-ECM) model is in the form:

$$\Delta y_t = \beta_0 + \gamma_0^+ \Delta^+ x_t + \gamma_0^- \Delta^- x_t + \delta_0 \Delta z_t + \lambda^+ e_{t-1} \times D(e_{t-1} > 0) + \lambda^- e_{t-1} \times D(e_{t-1} \le 0) + u_t \tag{9}$$

where $e_{t-1} = y_{t-1} - \frac{(\gamma_0 + \gamma_1)}{1-\beta_1} x_{t-1} - \frac{(\delta_0 + \delta_1)}{1-\beta_1} z_{t-1}$ is one period-lagged deviation from the long-run equilibrium relationship; $D(e_{t-1} > 0)$ is a dummy variable that equals one if $e_{t-1} > 0$ and equals zero otherwise; $D(e_{t-1} \le 0)$ is a dummy variable that equals one if $e_{t-1} \le 0$ and equals zero otherwise; λ^+ and λ^- are the corresponding adjustment parameters, β_0, γ_0^+, γ_0^-, and δ_0 are also parameters of this regression model.

The models (6) and (7) are obtained from models (8) and (9) using restrictions $\lambda^+ = \lambda^-$ and $\gamma_0^+ = \gamma_0^-$. The rejection of the hypothesis $\lambda^+ = \lambda^-$ (LR Symmetry) indicates asymmetry in adjusting the long-run equilibrium. The rejection of the hypothesis $\gamma_0^+ = \gamma_0^-$ (SR Symmetry) indicates short-term adjustment asymmetry. Both hypotheses can also be tested jointly. In cases where models (4) and (5) have a more extensive dynamic structure (γ_i), models (8) and (9) will also be more extensive.

Assume that there are cointegrating relationships between the crude oil price and the retail fuel prices for gasoline and diesel, individually or jointly. The vector error correction model (VECM) should be used to look for a long-term equilibrium relationship.

Similarly, the single-equation error correction model is auto-regressive, so the vector error correction model is a vector auto-regressive model. It can be shown by the vector auto-regressive model of order two denoted VAR(2):

$$\mathbf{y_t} = \mathbf{\Phi D_t} + \mathbf{\Pi}_1\mathbf{y_{t-1}} + \mathbf{\Pi}_2\mathbf{y_{t-2}} + \mathbf{u_t} \quad (10)$$

where $\mathbf{y_t}$ is the vector of variables (gasoline, diesel and oil prices) in time *t*; $\mathbf{D_t}$ is the matrix of deterministic terms (constant, trend, …) in time *t*; $\mathbf{u_t}$ is the vector of stochastic terms in time *t* and $\mathbf{\Phi}$, $\mathbf{\Pi_1}$ and $\mathbf{\Pi_2}$ are the matrices of unknown parameters of this model.

The model (10) can be rewritten as the vector error correction model (*VECM*) of order one:

$$\Delta\mathbf{y_t} = \mathbf{\Phi D_t} + \boldsymbol{\alpha\beta}^T\mathbf{y_{t-1}} + \mathbf{\Phi}_1\Delta\mathbf{y_{t-1}} + \mathbf{u_t} \quad (11)$$

where $\boldsymbol{\alpha\beta}^T = \left(\mathbf{\Pi}_1 + \mathbf{\Pi}_2 - \mathbf{I}\right)$ and $\mathbf{\Phi}_1 = -\mathbf{\Pi}_2$. Matrix $\boldsymbol{\beta}$ is called a co-integration matrix with co-integration vectors as columns and matrix $\boldsymbol{\alpha}$ is called a loading matrix. The test of co-integration in VECM is realized by Johansen's procedure (Johansen, 1988, 1991) by the lambda trace statistics depending on the specification of the deterministic components $\mathbf{\Phi D_t}$ of model (11).

The asymmetric form of this irreversible vector error correction (*A-VECM*) model is:

$$\boldsymbol{\Delta y_t} = \boldsymbol{\Phi D_t} + \boldsymbol{\alpha}^+[\boldsymbol{\beta}^T\boldsymbol{y}_{t-1} \odot D(\boldsymbol{\beta}^T\boldsymbol{y}_{t-1} > \mathbf{0})] + \boldsymbol{\alpha}^-[\boldsymbol{\beta}^T\boldsymbol{y}_{t-1} \odot D(\boldsymbol{\beta}^T\boldsymbol{y}_{t-1} \le \mathbf{0})] + \boldsymbol{\Phi}_1^+\boldsymbol{\Delta}^+\boldsymbol{y}_{t-1} + \boldsymbol{\Phi}_1^-\boldsymbol{\Delta}^-\boldsymbol{y}_{t-1} + \boldsymbol{u_t} \quad (12)$$

where the multiplication operation $\odot$ in square brackets of model (12) does not represent the matrix product, but the element-wise product (product of elements in the same positions); $\boldsymbol{\beta}^T\mathbf{y_{t-1}}$ is the vector of one period lagged deviations from the long-run equilibrium relationships; $D(\boldsymbol{\beta}^T\mathbf{y_{t-1}} > \mathbf{0})$ is the vector of a dummy variable; its element equals 1 if corresponding element of $\boldsymbol{\beta}^T\mathbf{y_{t-1}}$ is positive and equals 0 otherwise; similarly $D(\boldsymbol{\beta}^T\mathbf{y_{t-1}} \le \mathbf{0})$ is the vector of a dummy variable; its element equals 1 if corresponding element of $\boldsymbol{\beta}^T\mathbf{y_{t-1}}$ is not positive and equals 0 otherwise; $\boldsymbol{\alpha}^+$ and $\boldsymbol{\alpha}^-$ are the loading matrices of corresponding adjustment parameters and $\mathbf{\Phi_1}^+$ and $\mathbf{\Phi_1}^-$ are also matrices with some pairs of the asymmetric parameters of this model. Model (11) is obtained from model (12) using restrictions $\mathbf{\Phi_1}^+ = \mathbf{\Phi_1}^-$ and $\boldsymbol{\alpha}^+ = \boldsymbol{\alpha}^-$.

Threshold Autoregressive Cointegration Models

Engle et al. (1987) approach is based on a symmetric long-run relationship. A different solution to the problem than Granger et al. (1989) was proposed by Enders et al. (1998), who introduced Threshold Autoregressive Cointegration (TAR). If the adjustment to the long-run equilibrium is asymmetric, the cointegration test is mis-specified. To overcome the problem, Enders et al. (2001) replace the standard augmented Dickey-Fuller test equation with the following threshold autoregressive process:

$$\Delta e_t = I_t\rho_1 e_{t-1} + \left(1 - I_t\right)\rho_2 e_{t-1} + \varepsilon_t \quad (13)$$

where e_t is the deviation from the long-run equilibrium relationship (residual). If the errors are serially correlated, equation (13) can be augmented with the lagged differences of e_t as in the standard augmented Dickey-Fuller test.

Indicator function I_t is defined to depend on the lagged values of the residuals, according to the following scheme:

$$I_t = 1 \text{ if } e_{t-1} > 0 \text{ and } I_t = 0 \text{ otherwise} \quad (14)$$

alternatively, it is defined to depend on the lagged values of the first differences of residuals:

$$I_t = 1 \text{ if } \Delta e_{t-1} > 0 \text{ and } I_t = 0 \text{ otherwise} \quad (15)$$

The relationships (13) and (14) are called TAR cointegration. In contrast, the relationships (13) and (15) are known as momentum TAR (or M-TAR) cointegration. In M-TAR models, the threshold is placed on the variation of e_{t-1} rather than on e_{t-1}.

The null hypothesis $\rho_1 = \rho_2 = 0$ of no cointegration can be tested through an F test. The adjustment is symmetric for nonzero $\rho_1 = \rho_2$; thus, the Engle-Granger approach is a special case of (13) and (14).

In the case of the rejection of the null hypothesis in (13), the analysed variables are cointegrated, and the asymmetric ECM representation can be written as:

$$\Delta y_t = \lambda_{up} e_{t-1}^{up} + \lambda_{down} e_{t-1}^{down} + \sum_{i=0}^{p} \gamma_i \Delta x_{t-i} + \sum_{i=0}^{r} \delta_i \Delta z_{t-i} + \sum_{i=1}^{s} \alpha_i \Delta y_{t-i} + u_t \quad (16)$$

where $e_{t-1}^{up} = I_t e_{t-1}$, $e_{t-1}^{down} = (1 - I_t) e_{t-1}$. When ρ_1 is less than ρ_2, the increases tend to persist, whereas the decreases tend to revert quickly toward equilibrium.

Adjustment Cost Function in Linear-Exponential Form

In this part, the linear exponential adjustment cost function form is presented, and the reaction function specification is derived, as proposed by the authors in their earlier papers (Szomolanyi et al., 2020, 2022b). Consider the function in the form:

$$F\left[p_t, E_{t-1}(c_t)\right] = \frac{-\gamma\left[p_t - kE_{t-1}(c_t)\right] + \exp\left\{\gamma\left[p_t - kE_{t-1}(c_t)\right]\right\} - 1}{\gamma^2} \quad (17)$$

where p_t is the retail gasoline or diesel price, c_t is the crude oil price, E_{t-1} denotes the expectation conditional upon the information available at the time t–1, k is the technology coefficient. If the asymmetry coefficient γ is negative, the negative difference $p_t - kE_{t-1}(c_t)$ is costlier for the price-making firm than the positive.

The first-order condition of the firm choosing the output price to minimise the cost function (17) is the general form of the firm's reaction function.

$$\frac{-1 + \exp\left\{\gamma\left[p_t - kE_{t-1}(c_t)\right]\right\}}{\gamma} = 0 \quad (18)$$

By applying the l'Hôpital's rule, if γ tends to zero, the reaction function is linear:

$$\lim_{\gamma \to 0}(p_t) = kE_{t-1}(c_t) \quad (19)$$

Performing a second-order Taylor expansion of the exponential terms in (18), solving the equation for p_t and, prior to the generalised method of moments estimation (GMM) of the short-run relation, replacing the expected values with actual, and taking the first differences of the relation, the econometric specification of the firm's price reaction function is obtained:

$$\Delta p_t = k\Delta c_t - \frac{1}{2}\gamma\Delta\left[\left(p_t - kc_t\right)^2\right] + u_t \quad (20)$$

where Δ denotes the first difference operator, u_t is a stochastic term containing the first differences of terms of the third or higher orders of the expansion.

Assuming that the changes in the input prices Δc_t are a normally distributed process with zero mean and variance σ^2, the estimates of the average price biases are:

$$E(\Delta p_t) = -\frac{k^2\gamma}{2}\sigma^2 \quad (21)$$

The orthogonality conditions implied by the rational expectation hypothesis make the GMM a natural candidate to estimate the (20). Standard errors have been computed using the Newey-West procedure. The most important feature of the procedure explained by Newey et al. (1987) is its consistency in the presence of both heteroskedasticity and the autocorrelation of unknown forms.

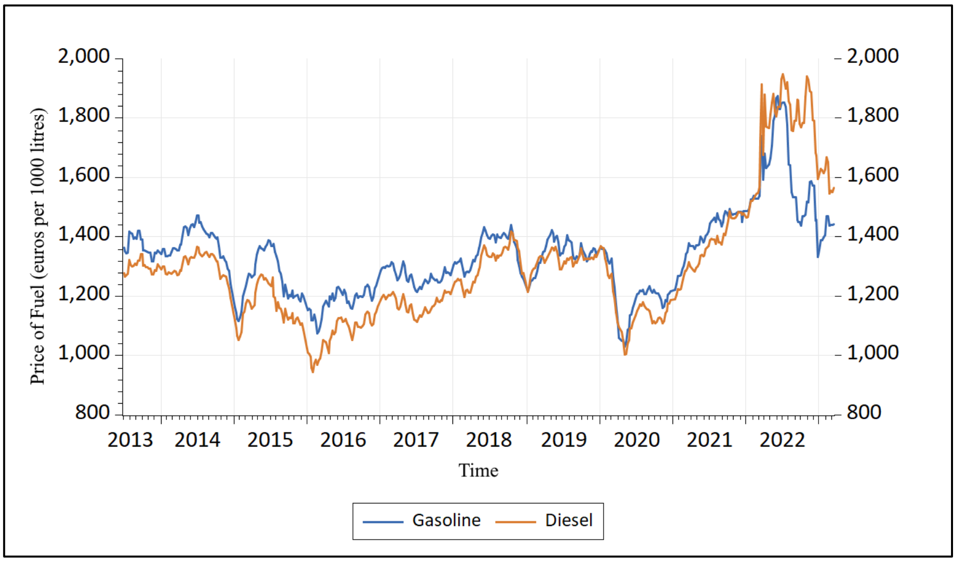
If retail gasoline and diesel prices are correlated, the relationship (20) can be estimated as the system of two equations, each for the given fuel price. The correlation is tested using the Breusch-Pagan test (Breusch et al., 1980).

Data

The analysis uses the weekly Croatian retail gasoline and diesel price data from the European Commission Weekly Oil Bulletin (energy.ec.europa.eu, 2023) and the daily Spot Brent crude oil prices obtained from the US Energy Information Administration website (eia.gov, 2023). The crude oil prices were converted to euros per 1000 litres using the daily exchange rate data series obtained from the European Central Bank (ECB.europa.eu, 2023). The daily data series were aggregated to the weekly by averaging.

The analysis uses the entire sample of published data, including the COVID-19 pandemic period. However, the common practice is sometimes different. Some studies split the sample into two or more sets to find evidence for the rockets' and feathers' effects (Bagnai et al., 2018; Bumpass et al., 2019; Cipcic, 2021).

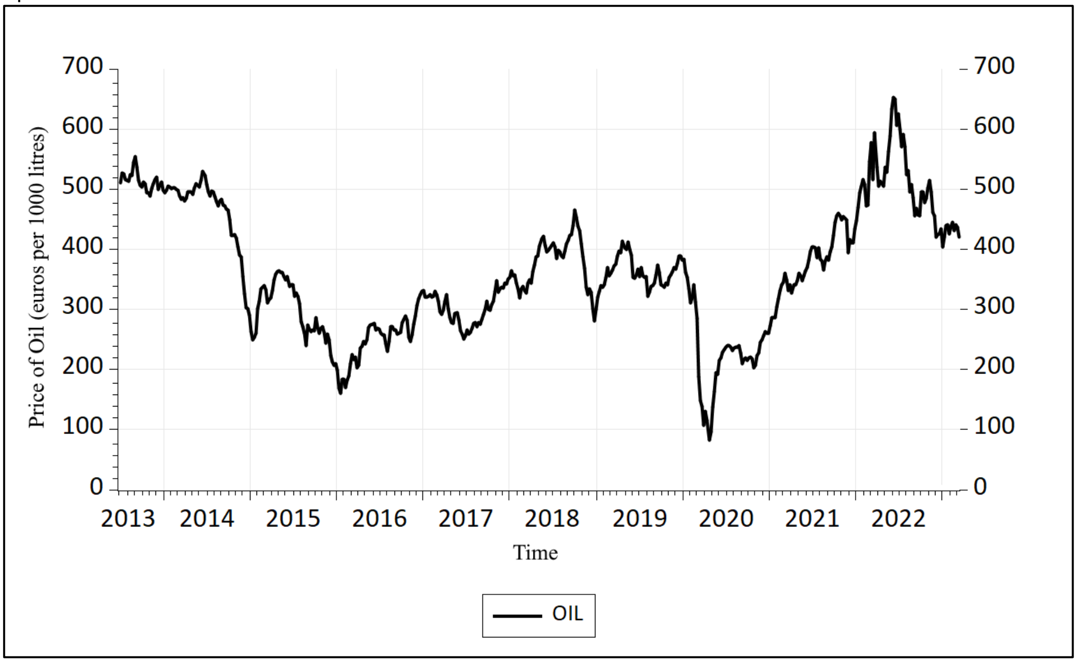
Figure 1
Retail Gasoline and Diesel Prices in Croatia – Time Series



Source: Author's illustration

Figure 1 graphically presents the weekly Croatian retail gasoline and diesel price data from January 2013 to March 2023, representing 483 data observations. The time series volatility has enlarged in the post-pandemic period. Since this period, the retail diesel price has been higher than the gasoline price almost continuously, which has happened only sporadically before. At the same time, both time series have practically identical courses.

Figure 2
Spot Brent Crude Oil Prices – Time Series



Source: Author's illustration

Figure 2 shows the weekly Spot Brent crude oil prices during the same period as fuel prices. All three data series have a very similar course, with a significant drop at the beginning of the pandemic period. Nevertheless, the decline in the third and fourth quarters of 2014, which continued slowly until 2016, was also interesting. The slowdown in global economic activity and shocks in the oil markets caused it. Fuel prices also copied this development.

When analysing data with similar characteristics, a changing trend, or volatility, it is customary to split the data set into several samples. Given the data used in the analysis, partitioning into changing trend periods would lead to selections with an insufficient data sample range, which is insufficient to use the Newey-West method and resistant to autocorrelation. Therefore, the entire range of available data is employed due to the adequacy of comparing individual procedures.

The given time series course indicates the non-stationarity of the processes generating these data. The non-stationarity was tested using the augmented Dickey-Fuller (ADF) unit root test (Dickey et al., 1981). The results of the ADF test for individual deterministic specifications (constant and trend, only constant and none) are shown for each variable and its first difference in Table 1. Test results in bold mean the rejection of non-stationarity (the first difference of all the data used). The null hypothesis of non-stationarity of the original variables was not rejected. This means that all prices are generated by non-stationary processes integrated by order one and are suitable for the search for co-integration, which is also indicated by an interconnected course.

Table 1
Augmented Dickey-Fuller Unit Root Test (Test of Non-stationarity)

Gasoline	Const	Trend	y_{t-1}	ΔGasoline	Const	Trend	y_{t-1}
const + trend	26.943	0.011	-0.022	const + trend	-0.834	0.002	**-0.911**
[τ stat]	[2.388]	[1.352]	[-2.511]	[τ stat]	[-0.354]	[0.292]	**[-18.90]**
const	23.436		-0.018	const	-0.237		**-0.911**
[τ stat]	[2.133]		[-2.149]	[τ stat]	[-0.202]		**[-18.92]**
none			0.0002	none			**-0.911**
[τ stat]			[-0.263]	[τ stat]			**[-18.94]**
Diesel	**Const**	**Trend**	y_{t-1}	**ΔDiesel**	**Const**	**Trend**	y_{t-1}
const + trend	18.594	0.019	-0.018	const + trend	-0.704	0.006	**-1.119**
[τ stat]	[2.019]	[1.748]	[-2.191]	[τ stat]	[-0.254]	[0.596]	**[-23.36]**
const	13.135		-0.010	const	0.731		**-1.118**
[τ stat]	[1.513]		[-1.450]	[τ stat]	[0.529]		**[-23.37]**
none			0.0003	none			**-1.117**
[τ stat]			[0.289]	[τ stat]			**[-23.39]**
Oil	**Const**	**Trend**	y_{t-1}	**ΔOil**	**Const**	**Trend**	y_{t-1}
const + trend	4.163	0.004	-0.015	const + trend	-1.020	0.004	**-0.854**
[τ stat]	[1.508]	[0.849]	[-2.179]	[τ stat]	[-0.726]	[0.741]	**[-19.37]**
const	5.094		-0.014	const	-0.118		**-0.853**
[τ stat]	[2.011]		[-2.141]	[τ stat]	[-0.167]		**[-19.36]**
none			-0.001	none			**-0.853**
[τ stat]			[-0.752]	[τ stat]			**[-19.38]**

Note: Test results in bold mean the rejection of non-stationarity. Values of tau statistics are in square brackets.
Source: Authors' calculations

Results

The results of the models from each group of methodological procedures are presented in individual tables. The section structure comes from the methodology section structure.

Simple Models of Asymmetry

Table 2 shows the results of estimations of models with the primary determinant oil price (1) and with an extended model supplemented by a change in the price of the second fuel (2). In the models, the corresponding autoregressive errors are estimated by appropriate methods to eliminate the autocorrelation problem. Table 3 shows the estimation results of dynamic models with a gradually increasing number of oil price lags included in the model. Standard errors have been computed using the Newey-West procedure.

The significance of the statistical tests is indicated by the number of asterisks, with three indicating statistical significance at the 1% level, two at the 5%, and one at the 10%. The probability values are enclosed in square brackets. The estimated parameters of all models in Table 1 are statistically significant, at least at the 5% significance level. Price symmetry was rejected in the models for gasoline prices but not diesel prices.

Table 2
Simple Models of Croatia's Fuel Prices Asymmetry

Simple Models	γ_0^+	γ_0^-	δ_0	Symmetry
gasoline model (1)	0.384***	0.395***	---	***F* = 6.182**
(std. err.)	(0.050)	(0.053)	---	**[0.013]**
diesel model (1)	0.351***	0.362***	---	*F* = 3.095
(std. err.)	(0.064)	(0.069)	---	[0.079]
gasoline model (2)	0.103**	0.109**	0.597***	***F* = 4.413**
(std. err.)	(0.045)	(0.046)	(0.016)	**[0.036]**
diesel model (2)	0.067**	0.069**	0.808***	*F* = 0.119
(std. err.)	(0.030)	(0.032)	(0.185)	[0.731]

Note: Three asterisks indicate statistical significance at the 1% significance level, two at the 5%. Test results in bold mean the rejection of symmetry. The probability values are in square brackets.
Source: Authors' calculations

Table 3
Dynamic Models of Croatia's Fuel Prices Asymmetry

Models with 1 Lag	Last Coefficients		Last Signif. Coefficients		Symmetry
	γ_1^+	γ_1^-		γ_1^-	
gasoline model (3)	-0.103	0.286*	-	0.286*	*F* = 0.373
(std. err.)	(0.227)	(0.167)	-	(0.167)	[0.542]
diesel model (3)	-0.195	-0.003	-	-	*F* = 0.024
(std. err.)	(0.350)	(0.263)	-	-	[0.878]
Models with 2 Lags	γ_2^+	γ_2^-	γ_2^+	γ_2^-	**Symmetry**
gasoline model (3)	1.049***	0.815***	1.049***	0.815***	*F* = 0.001
(std. err.)	(0.175)	(0.125)	(0.175)	(0.125)	[0.515]
diesel model (3)	1.357***	0.849***	1.357***	0.849***	*F* = 0.915
(std. err.)	(0.371)	(0.230)	(0.371)	(0.230)	[0.393]
Models with 3 Lags	γ_3^+	γ_3^-	γ_2^+	γ_3^-/γ_2^-	**Symmetry**
gasoline model (3)	-0.075	0.412**	1.100***	0.412**	*F* = 0.584
(std. err.)	(0.115)	(0.200)	(0.179)	(0.200)	[0.445]
diesel model (3)	-0.082	0.285	1.397***	0.800***	*F* = 0.342
(std. err.)	(0.197)	(0.204)	(0.386)	(0.253)	[0.559]
Models with 4 Lags	γ_4^+	γ_4^-	γ_2^+	γ_4^-/γ_2^-	**Symmetry**
gasoline model (3)	-0.184	0.292***	1.134***	0.292***	*F* = 2.634
(std. err.)	(0.199)	(0.113)	(0.190)	(0.113)	[0.105]
diesel model (3)	-0.097	0.146	1.414***	0.770***	*F* = 0.177
(std. err.)	(0.329)	(0.209)	(0.417)	(0.210)	[0.674]

Note: Three asterisks indicate statistical significance at the 1% significance level, two at the 5%, and one at the 10%. Test results in bold mean the rejection of symmetry. The probability values are in square brackets.
Source: Authors' calculations

As in the tables above, asterisks denote the statistical significance levels, and the probability values are square brackets. The significance of the parameters for lagged changes in the crude oil price indicates in the diesel price equation the dynamics of a maximum of two periods during both an increase and a price decrease, and in the gasoline price equation, two periods during a price increase, but up to four periods during a price decrease.

This result may indicate a certain temporal asymmetry of the reaction with gasoline. However, based on the comparison of the cumulated response, price symmetry was not rejected in any dynamic model either for the price of gasoline or for the diesel.

Error Correction Models

Table 4 shows the results of cointegration tests of models with a correction term (6) and models supplemented by the change in the price of the second fuel (7). Asterisks and square brackets denote the statistical significance and probability values. Engle-Granger (Engle et al., 1987) and Phillips-Ouliaris test (Phillips et al., 1990) in bold mean the non-rejection of cointegration. Both test procedures are based on tests of the unit root of the residuals of the cointegrating relationship. Engle and Granger use the augmented Dickey-Fuller test, and Phillips and Ouliaris use the Phillips-Perron statistics based on Newey-West estimate of standard error. In addition to these tests, the loading parameters $\lambda = \beta_1 - 1$ for (6) and (7) are estimated. In all models for gasoline prices, the variables are cointegrated. In contrast, in the models for diesel prices, cointegration was not rejected only in the trend model with no additional explanatory variables added.

Table 4
Error Correction Models of Croatia's Fuel Prices

ECM Gasoline	Engle-Granger	Phillips-Ouliaris	$\lambda = \beta_1 - 1$
model (6)	**τ = -4.429**	**τ = -4.163**	**-0.032*****
[p value]	[0.002]	[0.004]	(0.007)
model (6) + trend	**τ = -5.832**	**τ = -5.733**	**-0.072*****
[p value]	[0.000]	[0.000]	(0.000)
model (7)	**τ = -4.377**	**τ = -4.627**	**-0.079*****
[p value]	[0.002]	[0.001]	(0.017)
model (7) + trend	**τ = -4.662**	**τ = -4.839**	**-0.082*****
[p value]	[0.004]	[0.002]	(0.017)
ECM Diesel	**Engle-Granger**	**Phillips-Ouliaris**	$\lambda = \beta_1 - 1$
model (6)	τ = -2.991	τ = -2.586	-0.013
[p value]	[0.114]	[0.245]	(0.011)
model (6) + trend	**τ = -5.328**	**τ = -5.026**	**-0.063*****
[p value]	[0.000]	[0.001]	(0.021)
model (7)	τ = -2.939	τ = -3.342	-0.031*
[p value]	[0.270]	[0.130]	(0.017)
model (7) + trend	τ = -3.442	τ = -4.027	-0.066***
[p value]	[0.226]	[0.067]	(0.025)

Note: Three asterisks indicate statistical significance at the 1% significance level, two at the 5%, and one at the 10%. Engle-Granger (Engle et al., 1987) and Phillips-Ouliaris test (Phillips et al., 1990) in bold mean the non-rejection of cointegration. The probability values are in square brackets.
Source: Authors' calculations

Only models with unrejected cointegration are converted to asymmetric models (8) and (9). With their help, short-term and long-term symmetry is tested separately and then in a joint hypothesis. Table 5 shows the results of estimations and tests of asymmetric models with the correction terms (8) and (9). The test results in bold mean the rejection of long-run (LR), short-run (SR) or both symmetries.

Table 5
Asymmetric Error Correction Models of Croatia's Fuel Prices

A-ECM Gasoline	λ^+	λ^-	LR Symmetry	SR Symmetry	Both Sym.
model (8)	-0.044*	-0.016	*F* = 0.738	*F* = 0.537	*F* = 1.268
[p value]	(0.025)	(0.013)	[0.391]	[0.464]	[0.282]
model (8) + trend	-0.066***	-0.077***	*F* = 0.072	*F* = 1.711	*F* = 0.924
[p value]	(0.025)	(0.023)	[0.788]	[0.192]	[0.398]
model (9)	-0.071**	-0.077***	*F* = 0.017	*F* = 3.232	*F* = 1.629
[p value]	(0.028)	(0.025)	[0.895]	[0.073]	[0.197]
model (9) + trend	-0.075***	-0.079***	*F* = 0.006	*F* = 3.255	*F* = 1.704
[p value]	(0.028)	(0.025)	[0.939]	[0.072]	[0.183]
A-ECM Diesel	λ^+	λ^-	**LR Symmetry**	**SR Symmetry**	**Both Sym.**
model (8) + trend	-0.091**	-0.039	*F* = 0.926	***F* = 5.199**	*F* = 2.665
[p value]	(0.037)	(0.028)	[0.336]	[0.023]	[0.071]

Note: Three asterisks indicate statistical significance at the 1% significance level, two at the 5%, and one at the 10%. Test results in bold mean the rejection of long-run (LR), short-run (SR) or both symmetries. The probability values are in square brackets.
Source: Authors' calculations

Table 6 shows the key results of the cointegration test and the selection of the vector error correction model's deterministic scheme (11). The first part shows the values of Akaike information criteria (AIC) for the Johansen procedure's three most frequently used deterministic schemes. Based on the results of AIC and lambda trace statistics in the middle part, a model with two cointegrating vectors and a fourth deterministic scheme with a linear trend in their cointegrating relationships and an unrestricted constant was chosen.

Table 6
Vector Error Correction Model of Croatia's Fuel Prices

Akaike Information Criteria (AIC) by Rank and Model			
	Deterministic scheme		
Rank or No. of CEs	**2**	**3**	**4**
0	25.939	25.951	25.951
1	25.920	25.928	25.912
2	25.929	25.933	**25.896**
3	25.951	25.951	25.912
Cointegration Rank Trace Test - 4th deterministic scheme			
Hypoth. No. of CE(s)	**Eigenvalue**	**Trace Stat**	**p value**
None	0.0700	58.743	[0.001]
At most 1	0.0472	27.814	[0.028]
At most 2	0.0168	**7.224**	[0.322]
Test of restrictions:	**1.1260**	p value	[0.890]
Restricted estimate	**Gasoline** (i = 1)	**Diesel** (i = 2)	**Oil** (i = 3)
α_{i1}	-0.080***	0	0
(std. err.)	(0.016)	-	-
α_{i2}	0	-0.077***	0
(std. err.)	-	(0.016)	-

Note: Three asterisks indicate statistical significance at the 1% significance level. The probability values are in square brackets. The value of AIC in bold is the minimum of all AIC values.
Source: Authors' calculations

Furthermore, estimates of the loading matrix parameters supplemented after introducing acceptable restrictions have been included. The first cointegrating vector expressing the long-run relationship between the price of gasoline and oil affects only the equation of the price of gasoline. On the contrary, the second cointegrating vector expressing the long-run relationship between the price of diesel and oil affects only the equation of the price of diesel. This fact means that the parameters from (11) α_{21} and α_{31}, together with α_{12} and α_{32}, equal 0, which does not reject the restriction test equal to 1.126 with a p-value of 0.890 displayed at the bottom of Table 6.

Table 7 presents the results of testing short-run and long-run symmetry. Finally, both hypotheses are joined in the gasoline and diesel price equation of the estimated model from Table 6. Test results in bold mean the rejection of long-run (LR), short-run (SR), or both symmetries. The probability values are in square brackets. At the 5% significance level, only short-run symmetry in the gasoline equation is rejected.

Table 7
Asymmetric Irreversible Vector Error Correction Model of Croatia's Fuel Prices

SR symmetry	Gasoline	Diesel
statistics	**F = 5.470**	F = 2.811
[p value]	[0.019]	[0.094]
LR symmetry	**Gasoline**	**Diesel**
statistics	F = 0.824	F = 0.677
[p value]	[0.364]	[0.411]
Both symmetries	**Gasoline**	**Diesel**
statistics	F = 5.474	F = 2.952
[p value]	[0.065]	[0.229]

Note: Test results in bold mean the rejection of long-run (LR), short-run (SR) or both symmetries. The probability values are in square brackets. *Source:* Authors' calculations

Threshold Autoregressive Cointegration Models

Table 8 shows the results of symmetry testing in TAR and M-TAR models. The most appropriate model was selected using the Schwarz information criterion from models with differences lagged by up to 8 periods. Standard errors have been computed using the Newey-West procedure. The table presents estimates with one (crude oil prices) and two explanatory variables (crude oil and other fuel prices). The latter are marked as Cross TAR and Cross M-TAR.

Table 8
TAR and M-TAR Models of Croatia's Fuel Prices Asymmetry

TAR Models	$\rho_1 = \rho_2 = 0$	$\rho_1 = \rho_2$	M-TAR Models	$\rho_1 = \rho_2 = 0$	$\rho_1 = \rho_2$
gasoline	**F = 7.010**	F = 0.462	gasoline	**F = 5.842**	F = 0.558
[p value]	[0.001]	[0.497]	[p value]	[0.003]	[0.455]
gasol. + trend	**F = 12.10**	F = 0.105	gasol. + trend	**F = 8.887**	F = 0.280
[p value]	[0.000]	[0.746]	[p value]	[0.000]	[0.597]
diesel	**F = 4.560**	F = 0.006	diesel	F = 1.896	F = 0.614
[p value]	[0.011]	[0.940]	[p value]	[0.152]	[0.434]
diesel + trend	**F = 5.349**	F = 0.055	diesel + trend	**F = 4.621**	F = 0.015
[p value]	[0.005]	[0.815]	[p value]	[0.010]	[0.903]
Cross TAR	**$\rho_1 = \rho_2 = 0$**	**$\rho_1 = \rho_2$**	**Cross M-TAR**	**$\rho_1 = \rho_2 = 0$**	**$\rho_1 = \rho_2$**
gasoline	**F = 5.679**	F = 0.001	**gasoline**	**F = 10.06**	**F = 8.255**
[p value]	[0.004]	[0.981]	[p value]	[0.000]	[0.004]
gasol. + trend	**F = 6.955**	F = 0.016	**gasol. + trend**	**F = 9.203**	**F = 6.423**
[p value]	[0.001]	[0.900]	[p value]	[0.000]	[0.012]
diesel	**F = 4.292**	F = 2.133	diesel	F = 3.010	F = 0.549
[p value]	[0.015]	[0.145]	[p value]	[0.051]	[0.459]
diesel + trend	**F = 3.737**	F = 0.098	diesel + trend	**F = 4.389**	F = 0.327
[p value]	[0.025]	[0.755]	[p value]	[0.013]	[0.568]

Note: F test result in bold represents the rejection of null hypothesis. The probability values are in square brackets. *Source:* Authors' calculations

F test result in bold represents the rejection of null hypothesis. The probability values are in square brackets. In the estimations of M-TAR models, in which the price of diesel appears as the second explanatory variable (Cross M-TAR), the symmetrical adjustment of gasoline prices is rejected (in the other TAR and M-TAR models, the symmetrical adjustment of fuel prices is not rejected).

Adjustment Cost Function in Linear-Exponential Form

Tables 9 and 10 show the estimation results using GMM. Table 9 shows the single-equation (20) estimates, and Table 10 shows the system estimate. The lagged first differences of retail fuel and crude oil prices (up to two lags) are used as instruments. Asterisks and square brackets denote the statistical significance and probability values. *J* test result in bold means the rejection of orthogonality. The probability values are in square brackets.

Table 9
Results of the Linex Adjustment Cost Function Approach – single equations GMM

GMM Models	k	γ	J	Bias
gasoline	1.579***	–0.001***	0.521	**0.3145**
(std.err.)	(0.186)	(0.0001)	[0.470]	
diesel	1.470***	–0.001***	0.00001	**0.2122**
(std.err.)	(0.392)	(0.0002)	[0.997]	

Note: Three asterisks indicate statistical significance at the 1% significance level. J test result in bold means the rejection of orthogonality. The probability values are in square brackets.
Source: Authors' calculations

The biases (21) in € per 1000 litres are computed in the tables' last column. Table 10 presents the test of the mutual correlation of retail fuel prices. According to the value of BP = 206.76, the noncorrelation of the stochastic terms in both equations are not reject.

Table 10
Results of the Linex Adjustment Cost Function Approach – system GMM

GMM Model	k	γ	J	Bias
gasoline	1.537***	–0.001***	0.004	**0.2944**
(std.err.)	(0.205)	(0.0001)	[0.998]	
diesel	1.352***	–0.001***		**0.1716**
(std.err.)	(0.433)	(0.0002)	BP = 206.76	

Note: Three asterisks indicate statistical significance at the 1% significance level. J test result in bold means the rejection of orthogonality. The probability values are in square brackets.
Source: Authors' calculations

The reaction function specification coefficients k and γ are statistically significant in all cases. As the hypothesis $\gamma < 0$ is not rejected, an asymmetry in the retail fuel price adjustment is indicated. According to the computed bias values in both estimates, the degree of asymmetry is estimated to be higher when gasoline prices are adjusted.

A challenge is to fill the international comparison results gap. A natural way is to compare the Croatian results with the European ones. However, there is no data on average fuel prices in the Eurozone or EU average. The possibility is to compute the average European retail fuel prices. The European Commission Weekly Oil Bulletin publishes the retail fuel prices for all European Union member states, but averaging this data without state weights does not respect the reality. The European Union is a heterogeneous group of economies with different business power and characteristics. The better way is to use a weighted average. Such weights could be the retail fuel consumption or GDP in constant prices for each economy. Unfortunately, the data exists only in quarter or annual frequencies. Due to the robustness of the results, this section uses several possibilities to compute the European average retail fuel prices: the simple average, median, average weighted by fuel consumption, and average weighted by GDP in constant prices.

The yearly fuel consumption and the quarterly GDP at constant market prices are gathered from the Eurostat database. In the weights' data series, the value for each week is the value for a corresponding year or quarter. Using the data, Brent crude oil prices, and the system GMM, the estimates of (20) follow. The results are summarised in Table 11.

Table 11
Results of the Linex Adjustment Cost Function Approach using Different Computations of the Average Retail Fuel Prices in the European Union – system GMM

EU average	***k***	γ	**Bias**	**EU median**	***k***	γ	**Bias**
gasoline	1.494***	–0.001***	**0.3666**	gasoline	1.734***	–0.001***	**0.5913**
diesel	1.497***	–0.001***	**0.4076**	diesel	1.638***	–0.001***	**0.5268**
EU average weighted by consumption	***k***	γ	**Bias**	**average weighted by GDP**	***k***	γ	**Bias**
gasoline	1.446***	–0.001***	**0.3174**	gasoline	1.549***	–0.001***	**0.3757**
diesel	1.469***	–0.001***	**0.3869**	diesel	1.631***	–0.001***	**0.4644**

Note: Three asterisks indicate statistical significance at the 1% significance level.
Source: Authors' calculations

The table is divided into four parts according to the computations of the European averages of the retail fuel prices. The reaction function specification coefficients *k* and γ are statistically significant in all cases. As the hypothesis $\gamma < 0$ is not rejected, an asymmetry in the retail fuel price adjustment is indicated. The estimated price biases (21) are in the last columns.

Discussion

Deltas et al. (2020) noted that different methods, data samples, and frequencies generate different results when examining the crude oil price pass-through to retail fuel prices. The same is true of this analysis.

The result of simple static models implies that gasoline pricing is asymmetric, whereas diesel pricing is not. However, according to the simple dynamic model, both commodity prices change symmetrically. Only in one ECM model is the short-run symmetric adjustment of retail diesel prices rejected (not the long-run) in the specification with a trend in which the gasoline price does not appear as another explanatory variable. On the contrary, vector error correction models reject the short-run asymmetric price adjustment of retail gasoline prices. In the estimations of M-TAR models, in which the price of diesel appears as the second explanatory variable (Cross M-TAR), the symmetrical adjustment of gasoline prices is rejected (in the other TAR and M-TAR models, the symmetrical adjustment of fuel prices is not rejected).

According to the Linex adjustment cost function approach, the retail fuel prices adjust asymmetrically. The corresponding price bias is slightly higher for gasoline. Price decreases are more sluggish than price increases. There are several theoretical explanations for this behaviour.

The first theoretical rationale, known as oligopolistic coordination theory, posits that firms in an oligopoly market structure attempt to maintain a tacit agreement with their competitors. They do this by responding asymmetrically to changes in oil prices, thereby signalling their commitment to the agreement.

The second explanation, rooted in the production and inventory adjustment cost, reveals a strategic decision-making process. Aware of the costliness of adjusting production and inventory levels, firms spread the adjustment over time.

The search theory, the third theoretical reason, sheds light on how firms exploit the high oil price volatility. This volatility allows them to leverage consumers' high search costs, temporarily increasing their margin after an oil price hike.

The last explanation is the theory of strategic interactions. Price makers try not to pit against themselves a minority of consumers who tend to "punish" firms for unjustified or insufficiently explained price increases. There is a tendency to use the crude oil price increase to adjust retail fuel prices by additional unexpected costs or the effects of an increase in demand.

Due to frequent changes in trends and volatilities, some authors using traditional approaches to estimating the asymmetries in pricing tend to split data samples (Bagnai et al., 2018; Bumpass et al., 2019; Cipcic, 2021). On the other hand, the Linex approach uses the GMM estimator, the advantages of which will be demonstrated with sufficiently large numbers of observations. Therefore, splitting the data samples is not recommended when applying the approach.

Cipcic (2021) was one of the few authors investigating the asymmetric transition of crude oil prices to retail gasoline prices in Croatia. The author used the ECM model with a specification in which the oil price appeared as the only explanatory variable. She did not reject the symmetrical adjustment of gasoline prices in Croatia, even when she split the data sample into two subsamples. The same method reached the same conclusion (Table 4, model (8)).

The study estimates a weekly bias of 29 cents per 1000 litre (0.029 cents per litre) for gasoline and 17 cents per 1000 litre (0.017 cents per litre) for diesel (based on the two-equation system estimate). It should be noted that the observations include a highly volatile period since the outbreak of the COVID-19 pandemic.

The paper's authors have so far applied the Linex approach of verifying asymmetric pricing of retail fuels on weekly data from Slovakia and the USA. Both studies confirmed the asymmetric reactions of gasoline and diesel prices to changes in oil prices. Using Slovak data on fuel prices in the period 2009 to 2019 and the Linex approach, Szomolanyi et al. (2020) estimated the average weekly bias of gasoline prices in the value of approximately 0.13 – 0.16 euros per 1000 litres and the average weekly bias of diesel prices in the value of approximately 0.20 - 0.22 euros per 1000 litres.

Using the same approach, Szomolanyi et al. (2022b) note the asymmetric formation of retail fuel prices in the US in US regions and selected US states and cities. Interestingly, the asymmetric response of wholesale fuel prices traded in the three major US ports is not confirmed. However, on the other hand, research finds asymmetric responses of retail fuel prices to changes in New York and Gulf port fuel prices. These findings are similar to those of Gosinski et al. (2020). The authors found no asymmetry between crude oil and Polish wholesale fuel prices. In contrast, Polish retail fuel prices' reactions to wholesale price changes have been asymmetric between 2000 and 2016. The subject of the authors' further research is to examine the asymmetric reactions of retail fuel prices in all EU states.

As expected, the estimated price biases in the EU differ according to how retail fuel prices are averaged. Unlike in Croatia, in the European Union, average diesel price bias estimates are higher than gasoline in all cases. Comparing the GMM system estimates, the average price biases are higher in the EU average.

The theoretical reasons for asymmetric pricing (oligopolistic coordination, production and inventory adjustment cost, search theory, and strategic interactions) are described in more detail in the Introduction. The theory of strategic interactions between a firm and its consumers is one of three theories explaining the price stickiness hypothesis used in New Keynesian monetary models.

Douglas et al. (2010) argue that the evidence of asymmetric reactions of retail fuel prices predicts that this theory is the main reason for the price stickiness hypothesis. The hypothesis can be formulated mathematically by adjusting the profit functions of companies by adjusting costs in the form of Linex. Adjustment and estimated parameter values can be applied in dynamic stochastic general equilibrium models (DSGE), which analyse the effects of monetary and fiscal policy and other shocks in Croatia.

Conclusion

The paper focused on the possibilities of quantifying the asymmetric transition of crude oil prices to Croatian retail fuel prices. The traditional, well-known procedures and the Linex approach formulating a non-linear adjustment cost function were used. This approach using the GMM estimator demands a large data set, so the data sample was not divided into more subsamples. Although the studied time series are subject to frequent changes in trends and volatilities, the Linex approach on Croatian data detects systematic price bias caused by asymmetric price reactions in the sense of rockets' and feathers' effects.

Among all the approaches, Linex is preferred because, as the methodological part demonstrated, the theoretical justifications for the asymmetric price transition correspond to the non-linear formulation of the adjustment cost function. Another advantage of the Linex approach is that it can be used to directly estimate the average price bias, which allows comparison across different periods and regions.

As Deltas et al. (2020) point out, traditional methods of investigating asymmetric pricing depend on the choice of frequency and range of data and the method used. According to these results and the results of the authors' studies (Szomolanyi et al., 2020, 2022a, 2022b), the Linex approach always confirmed price asymmetries. The differences were only in the size of the bias.

A comparison with the results from Cipcic (2021) also corresponds to this. One of the approaches applied in this study uses the same methodology as the author. It also does not reject the symmetrical adjustment of fuel prices in Croatia. The result obtained by the Linex approach is the opposite.

Asymmetric price reactions do not have to occur only in the fuel market but also in other markets, such as agricultural, energy, and others, for the analysis of which the presented approaches, including Linex, can be used. Considering the advantages of the Linex approach, the price asymmetries for different European fuel prices will be compared in future research.

Limitations of the research

The transition from crude oil prices to retail fuel prices is complex. Several entities are involved in the supplier-customer chain. However, the available data reflect prices only at the input and output of the entire chain. The study's limitation is that the originator of the asymmetries cannot be identified.

In addition, the econometric test specification is non-linear, which is associated with other limitations of the paper. The results of software iterative methods for calculating parameter estimates are sensitive to the choice of starting values. The results of the instrumental estimation of the specification depend on the choice of instruments. The choice is applied by testing the instruments' weakness, orthogonality, and regressors' endogeneity. Tests of weakness of instruments and endogeneity of regressors are commonly applied only in linear econometric models.

References

1. Adda, J., & Cooper, R. W. (2003). Dynamic economics: Quantitative methods and applications. Cambridge, MA: The MIT Press
2. Asane-Otoo, E., & Dannemann, B. C. (2022). Rockets and Feathers Revisited: Asymmetric Retail Gasoline Pricing in the Era of Market Transparency. The Energy Journal, 43(6), 103–122. https://doi.org/10.5547/01956574.43.6.easa
3. Bacon, R. W. (1991). Rockets and feathers: The asymmetric speed of adjustment of UK retail gasoline prices to cost changes. Energy Economics, 13(3), 211-218. https://doi.org/10.1016/0140-9883(91)90022-R
4. Bagnai, A., & Mongeau Ospina, C. A. (2018). Asymmetries, outliers and structural stability in the US gasoline market. Energy Economics, 69, 250–260. https://doi.org/10.1016/j.eneco.2017.11.014
5. Balke, N. S., Brown, S. P. A., Yucel, M. K. (1998). Crude oil and gasoline prices: an asymmetric relationship? Federal Reserve Bank of Dallas, Economic and Financial Review, Q1, 2–11.
6. Borenstein, S., Cameron, A. C., & Gilbert, R. (1997). Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes? The Quarterly Journal of Economics, 112(1), 305–339. https://doi.org/10.1162/003355397555118
7. Borenstein, S., & Shepard, A. (2002). Sticky Prices, Inventories, and Market Power in Wholesale Gasoline Markets. The RAND Journal of Economics, 33(1), 116–139. https://doi.org/10.2307/2696378
8. Bragoudakis, Z., & Sideris, D. (2021). Asymmetric price adjustment and the effects of structural reforms and low demand in the gasoline market: The case of Greece. Journal of Applied Economics, 24(1), 504–522. https://doi.org/10.1080/15140326.2021.1922188
9. Breusch, T. S., & Pagan, A. R. (1980). The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. The Review of Economic Studies, 47(1), 239–253. https://doi.org/10.2307/2297111
10. Brown, S. P. A., & Yücel, M. K. (2000). Gasoline and crude oil prices: Why the asymmetry? Economic and Financial Policy Review, Q3, 23–29. Retrieved from https://fraser.stlouisfed.org/title/5730/item/583808/toc/527860
11. Bumpass, D., Douglas, C., Ginn, V., & Tuttle, M. H. (2019). Testing for short and long-run asymmetric responses and structural breaks in the retail gasoline supply chain. Energy Economics, 83, 311–318. https://doi.org/10.1016/j.eneco.2019.07.021
12. Cipcic, M. L. (2021). The Impact of Oil Prices on Petroleum Product Prices in Central East European Markets. In M. Simic & B. Crnkovic (Ed.), 10th International Scientific Symposium Region Entrepreneurship Development (RED 2021),174–188, Univ Osijek.
13. Deltas, G., & Polemis, M. (2020). Estimating retail gasoline price dynamics: The effects of sample characteristics and research design. Energy Economics, 92, 104976. https://doi.org/10.1016/j.eneco.2020.104976
14. Dickey, D. A., & Fuller, W. A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. Econometrica, 49(4), 1057–1072. https://doi.org/10.2307/1912517
15. Douglas, C., & Herrera, A. M. (2010). Why are gasoline prices sticky? A test of alternative models of price adjustment. Journal of Applied Econometrics, 25(6), 903–928. https://doi.org/10.1002/jae.1115
16. Duffy-Deno, K. T. (1996). Retail price asymmetries in local gasoline markets. Energy Economics, 18(1–2), 81–92. https://doi.org/10.1016/0140-9883(95)00051-8
17. ecb.europa.eu (2023), European Central Bank. Retrieved from https://ecb.europa.eu
18. eia.gov, (2023), US Energy Information Administration website. Retrieved from https://eia.gov

19. energy.ec.europa.eu, (2023), European Commission Weekly Oil Bulletin Retrieved from https://energy.ec.europa.eu
20. Enders, W., & Granger, C. W. J. (1998). Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates. Journal of Business & Economic Statistics, 16(3), 304–311. https://doi.org/10.2307/1392506
21. Enders, W., & Siklos, P. L. (2001). Cointegration and Threshold Adjustment. Journal of Business & Economic Statistics, 19(2), 166–176. https://doi.org/10.1198/073500101316970395
22. 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
23. Genakos, C., & Pagliero, M. (2022). Competition and Pass-Through: Evidence from Isolated Markets. American Economic Journal: Applied Economics, 14(4), 35–57. https://doi.org/10.1257/app.20200863
24. Gosinska, E., Leszkiewicz-Kedzior, K., & Welfe, A. (2020). Who is responsible for asymmetric fuel price adjustments? An application of the threshold cointegrated VAR model. Baltic Journal of Economics, 20(1), 59–73. https://doi.org/10.1080/1406099X.2020.1746114
25. Granger, C. W. J., & Lee, T. H. (1989). Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models. Journal of Applied Econometrics, 4(S1), S145–S159. https://doi.org/10.1002/jae.3950040508
26. Johansen, S. (1988). Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control, 12(2–3), 231–254. https://doi.org/10.1016/0165-1889(88)90041-3
27. Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. Econometrica, 59(6), 1551–1580. https://doi.org/10.2307/2938278
28. Kang, W., de Gracia, F. P., & Ratti, R. A. (2019). The asymmetric response of gasoline prices to oil price shocks and policy uncertainty. Energy Economics, 77, 66–79. https://doi.org/10.1016/j.eneco.2018.09.007
29. Karrenbrock, J. D. (1991). The Behavior of Retail Gasoline Prices: Symmetric or Not? Federal Reserve Bank of St. Louis Review, 73, 19–29. https://doi.org/10.20955/r.73.19-29
30. Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica, 55(3), 703–708. https://doi.org/10.2307/1913610
31. Okun, A. M. (1981). Prices and quantities: A macroeconomic analysis. Washington, DC: Brookings Institution Press
32. Perdiguero-Garcia, J. (2013). Symmetric or asymmetric oil prices? A meta-analysis approach. Energy Policy, 57, 389–397. https://doi.org/10.1016/j.enpol.2013.02.008
33. Pfann, G. A., & Palm, F. C. (1993). Asymmetric Adjustment Costs in Non-linear labour Demand Models for the Netherlands and U.K. Manufacturing Sectors. The Review of Economic Studies, 60(2), 397–412. https://doi.org/10.2307/2298064
34. Phillips, P. C. B., & Ouliaris, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration. Econometrica, 58(1), 165–193. https://doi.org/10.2307/2938339
35. Radchenko, S. (2005). Oil price volatility and the asymmetric response of gasoline prices to oil price increases and decreases. Energy Economics, 27(5), 708–730. https://doi.org/10.1016/j.eneco.2005.06.001
36. Rotemberg, J. J. (2011). Fair Pricing. Journal of the European Economic Association, 9(5), 952–981. https://doi.org/10.1111/j.1542-4774.2011.01036.x
37. Surico, P. (2007a). The Fed's monetary policy rule and U.S. inflation: The case of asymmetric preferences. Journal of Economic Dynamics and Control, 31(1), 305–324. https://doi.org/10.1016/j.jedc.2005.11.001
38. Surico, P. (2007b). The Monetary Policy of the European Central Bank. The Scandinavian Journal of Economics, 109(1), 115–135. https://doi.org/10.1111/j.1467-9442.2007.00485.x
39. Surico, P. (2008). Measuring the Time Inconsistency of US Monetary Policy. Economica, 75(297), 22–38. https://doi.org/10.1111/j.1468-0335.2007.00590.x

40. Szomolanyi, K., Lukacik, M., & Lukacikova, A. (2020). Asymmetric Retail Gasoline and Diesel Price Reactions in Slovak Market. Ekonomicky Casopis, 68(2), 115–133.
41. Szomolanyi, K., Lukacik, M., & Lukacikova, A. (2022a). Asymmetric Reactions of Retail Gasoline Prices on the Changes in Crude Oil Prices in Chosen US Cities. ENTRENOVA - ENTerprise REsearch InNOVAtion, 8(1), 9–15. https://doi.org/10.54820/entrenova-2022-0002
42. Szomolanyi, K., Lukacik, M., & Lukacikova, A. (2022b). Estimation of asymmetric responses of U.S. retail fuel prices to changes in input prices based on a linear exponential adjustment cost approach. Central European Journal of Operations Research, 30(2), 757–779. https://doi.org/10.1007/s10100-021-00783-0
43. Torrado, M., & Escribano, A. (2020). European gasoline markets: Price transmission asymmetries in mean and variance. Applied Economics, 52(42), 4621–4638. https://doi.org/10.1080/00036846.2020.1739224
44. Tweeten, L. G., & Quance, C. L. (1969). Positivistic Measures of Aggregate Supply Elasticities: Some New Approaches. American Journal of Agricultural Economics, 51(2), 342-352. https://doi.org/10.2307/1237584

About the authors

Karol Szomolányi is an associate professor of operations research and econometrics at the University of Economics in Bratislava, Faculty of Economic Informatics, Department of Operations Research and Econometrics. He is a secretary of the Slovak Society for Operations Research. He deals with macroeconomic analysis and econometrics. His areas of interest are real business cycle models and dynamic stochastic macroeconomic models and their use for developing countries of Eastern Europe. The author can be contacted at **karol.szomolanyi@euba.sk**.

Martin Lukáčik is a full professor of econometrics at the University of Economics in Bratislava, Faculty of Economic Informatics, Department of Operations Research and Econometrics. He is an ex-president of the Slovak Society for Operations Research. He deals with econometric methods in macroeconomic frameworks and time series econometrics. His areas of interest are VAR models used in macroeconomics and forecasting methods. The author can be contacted at **martin.lukacik@euba.sk**.

Adriana Lukáčiková is an assistant professor of operations research and econometrics at the University of Economics in Bratislava, Faculty of Economic Informatics, Department of Operations Research and Econometrics. She is a member of the Slovak Society for Operations Research. She deals with the computable general equilibrium models and micro-econometrics. Her area of interest is the econometric modelling. The author can be contacted at **adriana.lukacikova@euba.sk**.

Decision-Making Model to Support Agricultural Policies in Realizing Economic and Social Sustainability

Jaka Žgajnar, *Lidija Zadnik Stirn*
Biotechnical faculty, University of Ljubljana, Slovenia

Abstract

Background: Achieving economic and social sustainability is the goal of any policy when defining measures. We focus on the beef sector, where many challenges have arisen due to its structural characteristics, such as an unfavourable scale structure, high costs, low efficiency, and a low environmental footprint. This paper presents an example of the support provided by a mathematical programming model in the development of a Common Agricultural Policy Strategic Plan for the period 2023-2027. **Methods/approach:** It is a model based on linear programming that allows such an ex-ante analysis by calculating production plans at the farm level and aggregating the results at the sector level. **Objectives:** When defining the interventions, the question arose as to what the reform of the Common Agricultural Policy will bring and to what extent the sector should be supported in meeting these challenges. These were the concerns of agricultural policy that we sought to support by modelling different scenarios. **Results:** The results show that the situation of the sector will worsen, especially for larger farms, but they also show the great importance of production-related payments to mitigate the negative trend. **Conclusions:** The applied approach proves to be suitable for supporting the design of agricultural policy and achieving greater economic and social sustainability in the sector.

Keywords: decision-making model; farm model; mathematical programming; agricultural policies; CAP reform; beef sector

JEL classification: C00, C02, C54
Paper type: Research article

Received: Feb 12, 2024
Accepted: Jun 22, 2024

Citation: Žgajnar, J., & Zadnik Stirn, L. (2024). Decision-Making Model to Support Agricultural Policies in Realizing Economic and Social Sustainability. *Business Systems Research, 15*(2), 177-190.
DOI: doi.org/10.2478/bsrj-2024-0023

Introduction

The Common Agricultural Policy (CAP) aims to support the EU agricultural sector in addressing local and global challenges and to drive the development towards a smart, sustainable, competitive, resilient, and diversified agricultural sector to ensure long-term food security (European Commission, 2023). The new programming period 2023-2027 therefore faces many challenges. In addition to the issue of income and a fairer distribution of direct payments among farms, climate preferences, etc., the environment plays a particularly important role (e.g. Avasiloaei, 2022). As various studies show, one of the ways to reduce the environmental footprint of EU agriculture can be achieved at the price of lower production and higher food prices (Petsakos et al., 2022), which has a significant impact on farms. The latter is even more pronounced in sectors that have been more strongly supported by production-related measures in the past, such as the beef sector. To achieve these goals, which are reflected in nine specific objectives, the CAP provides tailored support through national CAP Strategic Plans (CSPs). These documents are drawn up by each Member State and cover all agricultural policy measures for the next programming period, 2023-2027. The interventions and CSPs are, therefore, tailor-made solutions adapted to the conditions and needs of each Member State in terms of resource allocation and priorities related to agricultural structure and environmental, economic, and social challenges (European Commission, 2023).

When drafting the CSP, many questions arose as to how the general and specific objectives could be achieved most efficiently and how the instruments could be made as effective and financially sustainable as possible. Animal production poses a particular challenge, as it is a rather complex sector that accounts for 44% of the total value of agricultural production in the EU and provides many ecosystem services (van der Linden et al., 2020). Here, for example, the question arises as to which livestock farming model we want, whether it is economically sustainable, whether financial support through interventions is necessary and for whom and to what extent. On the other hand, there is also the question of how the efficiency of coupled income support (CIS) can be improved and how the target sectors can be made more profitable and less dependent on subsidies. Many of these solutions require the support of experts and tools, including the use of modelling. A review of the literature revealed a variety of models that have been developed to support decision making.

The use of various methods to support policymakers has a long history (e.g. Žibert et al., 2020, Kocjančić et al., 2016; Quendler, 2019). In reviewing the literature, we find various examples of their use, particularly for ex-ante analysis, ranging from macro-sectoral analysis (models based primarily on the partial and general equilibrium approach), which were particularly prevalent in the first phase of analysis by agricultural economists, to the microsimulation models that emerged in the latter phase (Langrell et al., 2013). These are a type of microsimulation model commonly referred to as farm models. Van der Linden et al. (2020) emphasise that operational models, once operational, enable relatively fast and cost-effective analyses, even though their development is often tedious and costly.

Farm models are often used to assess the economic situation of farms and to model the effects of various policy and market changes (Reidsma et al., 2018). Such models enable a better understanding of decision-making and management at the farm level and, on the other hand, give policymakers a better insight into what is happening on individual farms, enabling them to make better evidence-based decisions and thus achieve greater targeting accuracy. The need for farm-level models became more apparent after the 2013 CAP reform, which introduced greening as an additional level of conditionality for farm-specific obligations to receive direct payments (Kremmydas et al., 2022).

In most cases, these are models based on the optimisation potential of mathematical programming (Reidsma et al., 2018). These include, for example, the IFM-CAP model used by the European Commission in the EU. IFM-CAP is based on the Positive Mathematical Programming (PMP) approach and enables the assessment of different policy impacts on existing aggregates and groups of farms (Louhichi et al., 2015). Its main purpose is to assess and analyse the different impacts of the CAP on the economic and environmental performance of farms. Its main advantage is its EU-wide coverage. An analysis of the impact of the CAP at the farm level for the post-2020 period was also carried out with IFM-CAP (Petsakos et al., 2022). This model is based on FADN data. Van der Linden et al. (2020) mention some reviews of existing agricultural models that have also been used for policy analysis and support. In recent years, agent-based models (ABM) have also gained popularity to model the impact of policy measures (Huber et al., 2018). In addition, Britz et al. (2021) also mention life cycle assessments and agri-environmental simulations as examples of policy impact assessment models. These models naturally differ both in terms of the input data and the (accuracy of) the modelling assumptions.

In Slovenia, we also have a microsimulation tool based on mathematical programming, which was used for this analysis. It is the Slovenian model of typical farms (SiTFarm), which enables various analyses at the level of the agricultural production plan, whereby the results can also be aggregated at the sector level (Žgajnar et al., 2022). The main purpose of SiTFarm is to enable analyses from the perspective of income sustainability at the level of typical farms that are representative of a certain number of real farms. This model does not require FADN data and, therefore, allows a more detailed analysis also for smaller farms that would otherwise not be included in the sample. The model calculates different economic indicators and allows the inclusion of different CAP interventions at different levels and under different conditions (socio-economic context of the analysis). From this point of view, the model has also been used to support CSPs in Slovenia (e.g. Žgajnar et al., 2023).

In this study, we show an example of the use of SiTFarm to support the design of a CSP using the example of cattle fattening. In this paper, we focus on the cattle fattening sector, as it is of great importance in cattle farming from an economic, social, and environmental point of view. It is the second most important sector in Slovenian agriculture after dairy cow farming (Žgajnar et al., 2023). In addition, the beef sector is characterised by persistent economic challenges and has consistently received coupled income support (CIS) in all CAP periods since the introduction of the CIS mechanism, making it a sector with persistent problems. In the case of cattle fattening, the actual question in the CAP reform was in what form and at what level coupled support should be granted to the sector and whether it makes a significant contribution to improving the economic situation of the different beef farms (size, production intensity, breeding technology, etc.). In this way, we have modelled the impact of the CAP reform after 2023 on the selected economic indicators of typical cattle farms in an ex-ante approach.

The published models mentioned above were developed in different programming environments or based on different programming languages. Van der Linden et al. (2020) state in their study that the most used programs are MS Excel, GAMS and ModelMaker. MS Excel is often used in combination with other add-in tools such as @Risk. In terms of programming languages, models are most often developed in R, Visual Basic for Applications, C or C++, Fortran, Pascal, Python and Stella.

In the following, we briefly present the SiTFarm tool used, the types of farms for which the analysis was carried out, and the assumptions of the scenario. We then present the most important results and conclude the paper with the key findings.

Methodology

Farm model and aggregate analysis

The SiTFarm used in this study is a tool based on a mathematical programming approach that allows analyses of the impact of policy measures at the level of the agricultural production plan and the level of the aggregated sector. The model was developed in MS Excel in combination with Visual Basic for applications. The methodological background allows the use of different techniques to solve the production plan problem, which is the basic level of analysis in our example. It is a tool that follows modern trends in agricultural economic analysis in this field and enables analysis at the level of the Typical Agricultural Holding (TAH) (Žgajnar et al., 2022). TAHs are static models of the agricultural system that enable the simulation and analysis of various factors at the level of a farm's production plan. In this way, the production plan shows the situation in a specific type of farm, which is a typical representative of a larger number of farms in practice. It is, therefore, not a specific farm but a typical representative of a certain group of farms that can be identified with it.

The optimisation potential of mathematical programming is used to calculate the baseline (BL) and to simulate various scenario sequences at the level of a farm's production plan. In the present version of SiTFarm, deterministic LP is used. Although LP modelling has inherent weaknesses due to factors such as the assumed maximisation behaviour and the explicitly linear technology (constant input-output coefficients), the models provide a fairly accurate simulation of both the revenue and the production and cost structures of the farms.

For each TAH, a matrix of production possibilities is developed, which is an example of production planning where we focus on finding the optimal combination of production activities given production constraints to maximise the expected gross margin (EGM). GM is calculated as the difference between total revenue, including subsidies at the farm level, and variable costs. A separate matrix is created for each farm. In addition to production activities, the matrix also includes marketing activities, the combination of technological activities, decoupled payments and CAP interventions that apply to a specific scenario.

Mathematically, the LP model used in SiTFarm could be defined as in equations (1) to (4). The objective function is to maximise the EGM (1). However, the main purpose of using LP is not to optimise the overall production plan but mainly to reconstruct the baseline production plan (BL) and balance it according to the key information we had for each typical cattle farm. This includes not only the production activities but also the nutrient balances and other input flows at the farm level. We, therefore, refer to this process as partial optimisation. Partial optimisation means that we specify a certain part of the activities (x_f) and require the solver to include them in the optimal solution (3). In our example, for example, these are the number of cattle at the farm level. These are the variables that define the type of farm. So, the basic idea is to estimate or calculate the missing data – variables (x_j) using a linear programme, maximising the EGM. To obtain the optimal solution for the LP model, the Analytic Solver V2021 (21.0.0.0) from FrontlineSolvers® was used.

$$axEGM = \sum_{j=1}^{n} c_j x_j + \sum_{f=1}^{n} c_f x_f \qquad (1)$$

so that

$\sum_{j=1;f=1}^{n;r} a_{ij}x_j + a_{if}x_f \leq b_i$ for all $i = 1$ to m (2)

$x_f = b_f$ for all $f = 1$ to r (3)

$x_j \geq 0$ for all j (4)

Where: n – number of activities included in the production plan of the analyzed farm; m – number of constraints taken into account when drawing up the farm's production plan; r – total number of binding constraints on the inclusion of production activities in the production plan; those defining the type of farm - e.g. number of bulls, cows, etc.; j – sequential production activity, marketing activity, technological activity; i – sequential constraint in the model; b_i – limitation of a single resource (e.g. ha of arable land); b_f – e.g. the number of bulls in the herd.

As the main source of economic (c_j and c_f) and technological data (technological coefficients - a_{ij}) for the individual production activities, the model uses the budget calculations (model calculations) of the Agricultural Institute of Slovenia (AIS, 2023). According to Jones et al. (2017), they can be defined as component models. They include cropping models and livestock models and enable real-time adjustment of individual budget calculations in terms of technology, intensity, and price-cost relations to the conditions on the analysed farm (TAHx). This makes the system of model calculations an important reference source for analytical and economic data at the level of the individual production activities of the analysed farms. Thus, the production costs of individual agricultural products depend on the production technology, intensity, size of the plot, and some other technological parameters, which enable an adjustment to the analysed conditions of the individual TAHs.

In terms of time resolution, SiTFarm makes it possible to carry out analyses for different time scales (monthly, annual average, average of several years, etc.). When calculating the economic indicators, we included the average prices for the three years 2018-2020 in the analysis. In this way, we have reduced the impact of inter-annual fluctuations, which can otherwise have an important impact on (market) revenues, costs and, above all, the EGM. These are the three main indicators we use to measure the impact of changes in our analysis.

Typical beef farms

The analysis for the beef sector was carried out based on 12 typical beef farms. These are typical representatives of the beef industry in Slovenia and are representative of a different number of farms in each size group in Slovenia (Table 1). They were identified based on an in-depth analysis of available statistical data, standard output analysis and other sources at workshops with various experts (Žgajnar et al., 2022). According to national data, 3,630 farms in Slovenia predominantly raise beef, excluding those that also raise suckler cows and excluding the part of fattening that is carried out on dairy farms.

This is a rather heterogeneous group of farms (Table 1), both in terms of size (number of cattle), natural resources (available area and proportion of arable and permanent grassland), intensity and quality of feed produced, and intensity of breeding (with daily weight gains ranging from 850 g/day up to 1,400 g/day). Most of them (97 %) are small farms where part-time labour is required (< 0.5 FTE). This also has an important impact on decision-making and management. As Huber et al. (2018) emphasise, farmers' decisions in such cases are often influenced by non-agricultural activities, as most of these farms are both a household and a farm/business unit.

Except for the last farm (TAH12), which also produces hops in addition to fattening cattle, all other farms are typical fattening farms. As can be seen from Table 1, the farming technology varies considerably between the farms. First, in addition to the breeding technology and the scope of breeding, the breed structure of the herd also varies. Most small farms mainly breed Brown Swiss and Simmental cattle. On the larger farms, mainly beef breeds and a combination of Simmental and Charolais are bred. These farms usually also fatten heavier calves (240 kg) that come from suckler cow herds. On farms with larger herds and more precise breeding techniques, beef calves are also fattened to higher final weights and usually achieve higher daily weight gains (on average up to 1,400 kg/day).

In addition to the intensity of breeding, there are also differences in the areas under cultivation and, therefore, in the quantity and quality of the feed produced. Farms that grow most of their feed on permanent grassland also achieve lower average daily weight gains. However, the latter also depends on the number of mowing operations and thus on the quality of the forage produced on grassland (on the field) and permanent grassland (Table 1).

Table 1

Typical agricultural holdings specialise in beef farming in Slovenia.

TAHs		TAH1	TAH2	TAH3	TAH4	TAH5	TAH6	TAH7	TAH8	TAH9	TAH10	TAH11	TAH12[e]
Farms	(No)	600	600	600	400	400	450	250	250	30	30	18	2
Beef	(No)	1	2	3	6	8	12	17	25	60	75	150	150
Breed		Bro	Sim	Sim	Bro	Sim	SimBro	Sim	Lim	SimCha	SimCha	SimCha	SimCha
Beginning of fat.	(kg)	120	120	120	240	240	240	240	240	240	240	240	240
End of fat.	(kg)	680	700	700	680	700	680	730	730	750	750	750	750
FTE (1,800h)		0.13	0.15	0.17	0.20	0.22	0.24	0.32	0.41	0.54	0.82	1.33	1.85
Arable land	(ha)				1.27	2.38	3.49	5.29	6.91	6.13	19.54	42.00	42.00
Grass/Lucerne mixtures	(ha)				0.25[a]	0.48[a]	0.70[a]	1.06[a]	1.38[b]		3.91[b]	8.40[b]	8.40[b]
Barley	(ha)				0.25					2.45	4.88	6.57	6.57
Corn	(ha)				0.76	1.90	2.79	4.23	5.53	3.68	10.75	27.03	27.03
Permanent grass	(ha)	1.0[c]	1.54[c]	2.02[c]	1.84[c]	0.92[c]	0.92[c]	0.92[c]	1.38[d]	9.90[d]	3.68[d]	5.52[d]	5.52[d]
Average plot size	(ha)	0.5	0.5	0.5	0.5	0.6	0.8	0.8	1	1.5	1.5	1.5	1.5
Distance from the farm	(km)	1	1	1.5	2	2.5	3	4	5	8	5	5	5
Average slope	(%)	7	7	8	10	3	5	7	6	5	2	2	2
Machine line capacity	(1-3)	1	1	1	1	1	2	2	2	3	3	3	3
Entitlements	(€/ha)	153	182	192	257	303	332	350	345	242	387	382	378

Note: *[a]Three-cut silage-bale, [b]Four-cut silage-silo and bale [c]Three-cut grass (silage bale, hay bale), [d]Four-cut grass (silage bale & silo, hay bale), [e]Includes also 5 ha of hops production, [f]eligible only in BL 2024-2022. Bro – Brown cattle, Sim – Simmental cattle; Lim – Limousine cattle, SimCha – a mixed herd of Simmental and Charolais cattle; FTE – full-time equivalent;*
Source: Authors' work

Profitability is also influenced by the average distance and size of individual plots, which are larger on larger farms. As shown in Table 1, larger farms are in flat areas (lower slope), while smaller farms (with less than 25 cattle) can also be located in a hillier area. However, in most cases, fattening does not take place in hilly mountainous areas, as could be the case with a certain share of suckler cow husbandry. The profitability of land management and, above all, feed preparation for the animals is also influenced by the farm's equipment and investment in machinery. Here, "1" means poor equipment and "3" means excellent, modern, and powerful equipment (Table 1).

Scenario analysis

In the analysis, we simulated the expected effects of changes in CAP interventions for cattle farms in Slovenia. First, we simulated the situation of the baseline (BL) before the reform (MAFF, 2015). We considered Pillar 1 measures and LFA payments. In contrast, the inclusion in voluntary farm environmental measures (eco-schemes) was simplified and modelled based on the data available at the time of the analysis (MAFF, 2023). We only considered two eco-schemes (Table 2) that are of interest to cattle farms operating on permanent grassland in eligible areas. We have assumed that certain farms within the group opt for it and others do not. As a result, the proportion of support considered may vary depending on the TAH. Thus, not all farms (real farms) within a particular type (TAH) may be included in a particular eco-measure.

For this study, we conducted an additional impact assessment for the CIS for beef (annual payment per animal). The aim was to help policymakers determine what kind of impact direct payments have and whether it is justified to support the beef sector as a sector with certain difficulties. Therefore, in Scenario 1 (S1), we have considered all expected payments to which the farm would be entitled, while in Scenario 2 (S2), we have excluded coupled income support for beef. In doing so, we analysed the impact of the CIS for beef on the economic indicators of the cattle TAHs.

Table 2

Considered CAP interventions in baseline and CAP strategic plan scenarios.

Scenario		BL	S1	S2
Period		2014-2022		2023-2027
Coupled income support				
Cereals	EUR/ha	126.4	0.0	0.0
Beef	EUR/LU	51.4	56.65[b]	0.0
Protein crops	EUR/ha	0	175.37	175.37
Decoupled income support				
Entitlements (A+B)				
A - Basic payment[d]	EUR/ha	161.3	0.0	0.0
B - Greening[d]	EUR/ha	91.4	0.0	0.0
Basic income support for sustainability		0.0	184.2	184.2
Redistributive payment (8.2 ha[a])	EUR/ha	0.0	23.16	23.16
Eco-scheme on grasslands				
Extensive grassland	EUR/ha	0.0	30.0	30.0
Traditional use of grassland	EUR/ha	0.0	30.0	30.0

Note: [a]*Farms receive payment only for the first 8,2 ha.* [b]*In order for the farm to be eligible, it must raise at least two beef.* [c]*Eco-scheme for the climate and the environment;* [d]*It is the average amount of payment. There are differences between individual TAHs due to historical payments and internal convergence. It is part of the value of the entitlement that the farm gets paid per ha of cultivated area. In the case of the analysed farms, this means that the amounts range from €152 to €387/ha. For more details per farm, look at Table 1; BL – baseline; S1 – scenario one, coupled income support for beef is included; S2 – scenario two, coupled income support for beef is not included.*

Source: Authors' work

The basic assumption of the scenario analysis was that the production plan at the farm level could only be partially changed (xj) and to the extent that does not change the production activities that determine the farm type (e.g. the number of cattle) (fx). This means that due to certain (favourable/unfavourable) conditions that an individual scenario entails, the production plan can only be partially changed. These are either a slight change in the distribution of production resources (labour, land, capital), an increased implementation of certain market activities or a shift in the share of home-grown fodder and other harvesting methods.

In the final step, we extrapolated the results at the TAH level to the sector level. This is done in such a way that each farm is given its weight according to the number of farms (Table 1) and their economic and social importance in agriculture. In this way, changes at the farm level also affect the sectoral level and thus enrich the analysis at the sectoral level. However, it should be noted that SiTFarm does not allow analyses from the point of view of structural change or the effects of structural change. The analysis is static in this respect.

Results and discussion

The core economic results of the scenario analysis are presented below. Beef is the main production activity on about 7% of farms in Slovenia (Fig. 1). In the SiTFarm, the whole sector, represented by 12 TAHs, contributes 4.4% to the total income of the Agriculture sector.

As shown in Table 3, 98% of cattle farms are smaller than the average Slovenian farm in terms of available land. Small herds predominate. Therefore, poor economic results (BL) can be expected for these farms.

As shown in Table 3, only farms with more than 25 cattle achieve an EGM of more than 10 €/hour. The exception is TAH9, where the results are less favourable. As can be seen from Table 1, this farm relies predominantly on forage from grassland, where the cost per unit of production is higher than when the majority of the forage comes from arable land. Therefore, the GM/ha is also much lower and among the worst in the sector.

Very small farms (accounting for 84% of Slovenian cattle farms) with less than 6 cattle usually achieve less than 4€ per working hour. According to the results, the last farm (TAH12), which also produces hops, stands out in all economic indicators due to hop production. In this case, hops also represent an important part of the farm's total income, although beef production is still the main agricultural activity. However, this type of farm is typical of only one region in Slovenia.

The other farms can be found all over Slovenia and could be categorised as farming models at the regional level in terms of spatial resolution. As explained in the methodological part of the paper, we have required in the modelling that the production plan should not change between the scenarios (BL, S1 and S2). At constant prices (2018-2020), the CAP measures are, therefore, an important factor for change.

Table 3
Typical agricultural holdings specialised in beef farming in Slovenia and selected economic indicators.

TAHs	Farms	Beef	FTE	Field	Permanent grass	Budgetary payments	GM	GM/h	GM/h	GM/h
	(No)	(No)	(1 800 h)	(ha)	(ha)	(EUR)	(EUR)	(EUR)	(EUR)	(EUR)
				Baseline (BL)					S1	S2
TAH1	600	1	0.13	0.00	1.00	309	262	1.1	1.1	1.1
TAH2	600	2	0.15	0.00	1.54	537	883	3.2	3.3	3.0
TAH3	600	3	0.17	0.00	2.02	809	1,344	4.5	4.5	4.1
TAH4	400	6	0.20	1.27	1.84	1,478	1,115	3.1	2.8	2.1
TAH5	400	8	0.22	2.38	0.92	1,833	1,721	4.4	4.2	3.3
TAH6	450	12	0.24	3.49	0.92	2,520	2,546	5.9	5.1	3.9
TAH7	250	17	0.32	5.29	0.92	3,826	4,933	8.5	7.9	6.6
TAH8	250	25	0.41	6.91	1.38	5,258	8,136	11.1	9.7	8.3
TAH9	30	60	0.54	6.13	9.90	7,779	5,973	6.1	5.6	3.0
TAH10	30	75	0.82	19.54	3.68	14,143	17,810	12.1	10.2	8.5
TAH11	18	150	1.33	42.00	5.52	27,987	33,504	14.0	11.2	9.2
TAH12	2	150	1.85	42.00	5.52	30,088	75,401	22.7	20.8	19.4
Total	**3 630**	**32 145**	**796**	**7 689**	**5 341**					

GM – gross margin, FTE – full-time equivalent, TAH – typical agricultural holding, BL – baseline; S1 – scenario one, coupled income support for beef is included; S2 – scenario two, coupled income support for beef is not included.
Source: Authors' work

Subsidies are particularly important in the beef sector. Budget payments account, on average, for more than 80 % of the GM and, in many small farms, even exceed the GM realised. Therefore, the sector can be expected to be very sensitive to changes in CAP measures. As the results show, the reform of CAP interventions will lead to a deterioration (Fig. 1), especially for medium and large farms (GM/hour), and even more so if the CIS for beef is not available (S2) (Table 3). The positive effect of production-related support is recognisable in all TAHs (S1). If we compare the hourly rate (GM/h) in BL and after the reform of CAP measures in both scenarios (S1 and S2), the effect of CIS payments is obvious. On average, the farms achieve 15 % more GM per effective working hour (S1) if a CIS is available (S1). In the individual farms, however, this effect can range from a few % (in small farms and hopefully diversified farms) to over 40 %. This also raises the further challenge of whether it is necessary to differentiate the level of the premium for different herd sizes or breeding technologies. However, the expected retention of the CIS (S1) for this sector will only partially mitigate the negative effects of the total CAP reform (Table 4). Without it, there will be an even greater deterioration, as can be clearly seen in Figure 1 and Table 3. This is also primarily a consequence of the abolition of entitlements (Table 1), which comprised an important proportion of historical payments from previous CAP periods. It is, therefore, also to be expected that the result of this abolition, together with the possible simultaneous abolition (S2) of the CIS for beef, will further worsen the economic situation and, thus, the social sustainability of these farms. In the modelling, we have taken into account that these farms can only participate in ECO programs on permanent grassland (Table 1). Although we do not provide detailed results here, it appears that ECO schemes are not an important factor and, therefore, do not contribute significantly to improving the performance of cattle fattening farms. Since they are linked to permanent grassland, the effect is expectedly greater on farms whose fodder is mainly produced on permanent grassland, which is particularly the case on small farms.

Table 4.
Total budgetary payments per TAH and sector level.

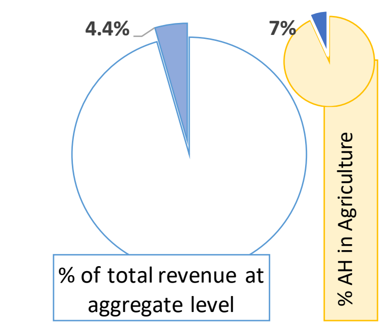
TAHs	Budgetary payments per TAH (EUR)		
	BL	S1	S2
TAH1	309	305	305
TAH2	537	559	474
TAH3	809	818	690
TAH4	1,478	1,359	1,104
TAH5	1,833	1,742	1,402
TAH6	2,520	2,162	1,652
TAH7	3,826	3,451	2,729
TAH8	5,258	4,233	3,171
TAH9	7,779	7,232	4,683
TAH10	14,143	8,759	6,529
TAH11	27,987	18,358	13,897
TAH12	30,088	19,691	15,230
Total per sector	**6,943,569**	**5,992,143**	**4,718,748**

BL – baseline; S1 – scenario one, coupled income support for beef is included; S2 – scenario two, coupled income support for beef is not included.
Source: Authors' work

Of course, the social sustainability of this sector is very important. It is closely linked to the milk production sector (mostly medium-sized farms) and suckler cow husbandry through the purchase and further fattening of calves, but nevertheless represents an important social aspect, as it employs almost 800 FTEs of the total effective labour force in the agricultural sector and accounts for 7% of all farms in Slovenia (Fig. 1).

Figure 1 shows that the CAP reform will bring only minor changes in revenue, which will be much more pronounced on the budgetary payments side. Since budgetary payments play an important role in the GM, it is to be expected that they will have a considerable influence on the GM achieved, given the production costs. Regardless, it is also clear from the results that simply monitoring the loss of budgetary payments does not give the same picture as monitoring the GM. This is also an important conclusion of our work. At the aggregate level, we estimate that GM has fallen by 10.3% (S1). If policymakers did not adopt a production-linked payment for beef (CIS), the deterioration at the aggregate level compared to BL would be up to 25 (S2) (Table 4). The overall decrease in budget payments would be 13.7% (S1) compared to the baseline of EUR 6.9 million, while production-related payments would amount to almost EUR 1.3 million (S2).

Figure 1
Summary of selected indicators for the beef sector in BL, S1 and S2.



	Gross margin (EUR)	Gross margin per FTE (EUR)	Gross margin per working hour (EUR/h)	Scenario	Gross margin per working hour (EUR/h)	Gross margin	Revenues	Budgetary payments	LFA
	Baseline (BL)			After the CAP reform		Share based on baseline			
Minimum	262	1,978	1.1	S1	1.1	0.758	0.948	0.619	0.712
				S2	1.1	0.482	0.930	0.462	0.712
Average	2,344	10,692	5.9	S1	5.3	0.897	0.981	0.863	0.942
				S2	4.5	0.747	0.956	0.680	0.942
Maximum	75,401	40,802	22.7	S1	20.8	1.024	1.006	1.040	1.267
				S2	19.4	0.986	0.998	0.988	1.267

LFA - Less-Favoured Area; AH – agricultural holding, BL – baseline; S1 – scenario one, coupled income support for beef is included; S2 – scenario two, coupled income support for beef is not included.
Source: Authors' work

Conclusion

Based on the analysis, we can conclude that the modelling approach used has proven to be effective in providing various business insights (indicators) in the scenario analysis, both at the farm level and at the sector level. The use of mathematical programming techniques allows us to balance the material balances. It flows at the farm level in a relatively simple way so that the production plan is technologically consistent and balanced.

However, it has been shown that the sensitivity of the LP model can be problematic when simulating different CAP measures, especially in marginal cases. This is indeed a problem of LP, where, in some cases, small (as well as larger) changes in conditions (CAP measures) can lead to completely different solutions. We circumvent the problem by including additional conditions in the model that make the model more static and consequently do not allow us to analyse possible structural changes.

Since the model is also static from the perspective of production technologies, as it does not include possible changes in production technologies in the modelling process, a limitation arises, namely that in this way, we can only analyse a certain part of CAP interventions that do not interfere with production and breeding technology. However, some interventions attempt to directly influence the change in a particular technology or management practice to reduce the environmental footprint, but this cannot be modelled due to the mentioned capabilities of the tool. In some cases, such dynamics may also mean that a different type of TAH occurs, which we also did not anticipate in this analysis.

While the results make an important contribution to the policy debate and the farm-level effects have generally been overlooked, an additional analysis at the farm level is needed. This is especially true in the development of CSPs, which should consider the national and local characteristics of Member States. CIS support for beef increases the resilience of the sector by supporting viable farm incomes. We have found that CIS is important for the beef sector anyway. On average, farms achieve 15% more GM per effective working hour when a CIS is in place. However, on individual farms, this effect can range from a few per cent to more than 40%. At an aggregate level, therefore, we estimate that total GM falls by 10%. If policymakers do not adopt a production-based payment for beef (CIS), the deterioration at the sector level would be up to 25% of total GM compared to BL (CIS payments thus represent this difference). The latter would certainly contribute to a deterioration in social sustainability, which is already bad for these types of farms.

An important conclusion of our work is also that it is not sufficient for policy impact assessment to look only at the level of budgetary payments and changes at this level, as is usually done by stakeholders, but that production costs and total revenues should also be considered. In such a case, we could obtain a significantly different picture.

References

1. AIS. (2023). Model calculations. Agricultural Institute of Slovenia https://www.kis.si/en/Model_Calculations_OEK/
2. Avasiloaei, D. (2022). Reinventing the Business Model in the Tourism Industry based on Sustainable Circular Economy and Innovation. *ENTRENOVA-ENTerprise REsearch InNOVAtion*, *8*(1), 435-448. https://doi.org/10.54820/entrenova-2022-0036
3. Britz, W., Ciaian, P., Gocht, A., Kanellopoulos, A., Kremmydas, D., Müller, M., ... & Reidsma, P. (2021). A design for a generic and modular bio-economic farm model. *Agricultural Systems*, *191*, 103133. https://doi.org/10.1016/j.agsy.2021.103133
4. European Commission. (2023). Approved 28 CAP Strategic Plans (2023–2027). Summary overview for 27 Member States. Facts and figures.
5. Huber, R., Bakker, M., Balmann, A., Berger, T., Bithell, M., Brown, C., ... & Finger, R. (2018). Representation of decision-making in European agricultural agent-based models. *Agricultural systems*, *167*, 143-160. https://doi.org/10.1016/j.agsy.2018.09.007
6. Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I., ... & Wheeler, T. R. (2017). Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. *Agricultural systems*, *155*, 269-288. https://doi.org/10.1016/j.agsy.2016.09.021
7. Kocjančič, T., Žgajnar, J., & Juvančič, L. (2016). Multiple-perspective reorganisation of the dairy sector: mathematical programming approach. *Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy*, *7*(2), 35-48. https://doi.org/10.1515/bsrj-2016-0011
8. Kremmydas, D., Petsakos, A., Ciaian, P., Baldoni, E., & Tillie, P. (2022). The EU-Wide Individual Farm Model for Common Agricultural Policy Analysis (IFM-CAP v.2): Manual of the model, EUR 30949 EN, Publications Office of the European Commission, Luxembourg, ISBN 978-92-76-46336-8, https://doi.org/10.2760/248136, JRC127014
9. Langrell, S., Ciaian, P., Espinosa, M., Gomez Paloma, S., Heckelei, T., Louhichi, K., Sckokai, P., Thomas, A., & Vard, T. (2013). Farm level modelling of CAP: a methodological overview. Joint Research Centre: Scientific and Policy Reports. Luxembourg: European Union.
10. Louhichi, K., Ciaian, P., Espinosa, M., Colen, L., Perni, A., & Gomez y Paloma, S. (2015). An EU-Wide Individual Farm Model for Common Agricultural Policy Analysis (IFM-CAP). JRC Scientific and policy reports. Luxembourg: Publication Office of the European Union.
11. MAFF (2015) Rural development programme of the Republic of Slovenia 2014-2020 (CCI 2014SI06RDNP001). Version 1.3. Ministry of Agriculture, Forestry and Food of the Republic of Slovenia; European Commission, Ljubljana; Brussels Available from: https://www.program-podezelja.si/en/rural-development-programme-2014-2020.
12. MAFF. (2023) Common Agricultural Policy Strategic Plan 2023-2027 for Slovenia. Version 3.0. Ministry of Agriculture, Forestry and Food of the Republic of Slovenia; European Commission, Ljubljana; Brussels Available from: https://skp.si/skupna-kmetijska-politika-2023-2027.
13. Petsakos, A., Ciaian, P., Espinosa, M., Perni, A., & Kremmydas, D. (2023). Farm-level impacts of the CAP post-2020 reform: A scenario-based analysis. *Applied Economic Perspectives and Policy*, *45*(2), 1168-1188. https://doi.org/10.1002/aepp.13257

14. Quendler, E. (2019). The digital transformation of the classroom–the what, how, and why in austrian secondary colleges of agriculture and forestry. *ENTRENOVA-ENTerprise REsearch InNOVAtion, 5*(1), 200-216.
15. Reidsma, P., Janssen, S., Jansen, J., & van Ittersum, M. K. (2018). On the development and use of farm models for policy impact assessment in the European Union–A review. *Agricultural systems*, *159*, 111-125. https://doi.org/10.1016/j.agsy.2017.10.012
16. van der Linden, A., de Olde, E. M., Mostert, P. F., & de Boer, I. J. (2020). A review of European models to assess the sustainability performance of livestock production systems. *Agricultural Systems*, *182*, 102842. https://doi.org/10.1016/j.agsy.2020.102842
17. Žibert, M., Rozman, Č., Škraba, A., & Prevolšek, B. (2020). A System Dynamics Approach to Decision-making Tools in Farm Tourism Development. *Business Systems Research: International Journal of the Society for Advancing Innovation and Research in Economy, 11*(2), 132-148. https://doi.org/10.2478/bsrj-2020-0020
18. Žgajnar, J., Kavčič, S., Tomšič, M., Zagorc, B., Brečko, J. (2023). Analysis of the impact of CAP changes between 2014-2022 and 2023-2027 on climate change, the environment and the economy of farms using the SiTFarm for Slovenia. Ljubljana: Biotechnical faculty, 617 p (in Slovene)
19. Žgajnar, J., Kavčič, S., Tomšič, M., Zagorc, B., Brečko, J., Hiti Dvoršak, A., Moljk, B., Verbič, J., Bedrač, M., Kožar, M., Cunder, T., Jerič, D. (2022). Development of a model for systematic monitoring of the economic situation and analysis of the impact of agricultural policy at the level of typical agricultural: final report. Ljubljana: Biotechnical faculty, 437 p. (in Slovene).

About the authors

Jaka Žgajnar, PhD is an Associate professor employed at the Biotechnical Faculty, University of Ljubljana (Slovenia). He teaches a course on farm management, animal health economics, and the utilisation of mathematical programming techniques in agriculture. In the field of research, he deals mainly with using various mathematical programming techniques in the analysis of decision-making in agriculture. He deals with farm-level modelling and conducting impact assessment analyses of Common agricultural policy. He is actively involved in several research projects. Since 2018, he has been a Council member of the International Farm Management Association (IFMA). The author can be contacted at **jaka.zgajnar@bf.uni-lj.si**.

Lidija Zadnik Stirn, PhD is a Professor Emeritus of Operations Research at the University of Ljubljana, Biotechnical Faculty, Slovenia, and an Associate Professor at the University of Tennessee, Knoxville, USA. Her research focuses mainly on deterministic, stochastic, dynamic, multi-objective and group decision-making methods for socioeconomic and environmental problems in natural resource management. Her work has resulted in more than 600 publications. She has been President of the Slovenian Society for Operations Research since 1997, Vice-President of the Slovenian Society INFORMATIKA since 2002, Slovenian representative in IFORS, EURO and IFIP TC7 and coordinator of the IUFRO group Managerial Economics and Accounting. The author can be contacted at **Lidija.Zadnik@bf.uni-lj.si.**