



**Faculty of Economics-Skopje**  
**Ss. Cyril and Methodius University in Skopje**

# **ECONOMIC AND BUSINESS TRENDS SHAPING THE FUTURE**

**International Scientific Conference, Skopje,**

**13-14 November 2025**

**6th International Scientific Conference  
Economic and Business Trends Shaping the Future  
November 13-14, 2025**

**Conference Proceedings**

Edited by  
Nikolina Palamidovska-Sterjadovska, PhD

Published by  
Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje  
Bul. Goce Delchev 9V, 1000 Skopje, Republic of North Macedonia  
[www.eccf.ukim.edu.mk](http://www.eccf.ukim.edu.mk)

Edition: 6th  
Publication type: e-book  
Available at: <https://repository.ukim.mk/handle/20.500.12188/34522>  
Published: Skopje, December 2025  
ISSN: 2671-3705

## **EDITORIAL NOTE**

These Conference Proceedings of the 6th International Scientific Conference “Economic and Business Trends Shaping the Future (EBTSF 2025)”, held on 13-14 November 2025 at the Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje includes a total of 39 contributions presented at the conference and approved for publication.

Submissions were received in accordance with the EBTSF Authors Guidelines. All submissions underwent a double-blind peer-review process managed through the EasyChair system, and only accepted submissions that were presented at the conference are included in these proceedings.

The volume is organized into four sections: Full Papers, Panel Papers, Extended Abstracts, and Doctoral Symposium Papers.

## SCIENTIFIC COMMITTEE (alphabetical order)

- Agnieszka Piekutowska, Faculty of Economics and Finance, University of Bialystok, Poland
- Aleksandar Petreski, Jönköping International Business School, Jönköping University, Sweden
- Anita Ciunova-Shuleska, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Anna Lasakova, Faculty of Management, Comenius University in Bratislava, Slovakia
- Atanasko Atanasovski, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Aykut Hamit Turan, Sakarya University, Turkey
- Boban Melovic, Faculty of Economics, University of Montenegro, Montenegro
- Bobi Ivanov, College of Communication and Information, University of Kentucky, USA
- Bojan Shimbov, University Jaume I, Spain
- Branislav Boricic, Faculty of Economics, University of Belgrade, Serbia
- Catalin Nicolae Albu, Bucharest Academy of Economic Studies, Faculty of Accounting and Management Information Systems, Romania
- Dejan Spasic, University of Nis, Faculty of Economics
- Dragan Gligoric, Faculty of Economics, University of Banjaluka, Bosna and Herzegovina
- Fatmir Memaj, Faculty of Economy, University of Tirana, Albania
- Francisco Simões, Iscte-University Institute of Lisbon, Portugal
- Goce Andrevski, Queen's University, Canada
- Goce Petreski, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, Macedonian Academy of Sciences and Arts, North Macedonia
- Goran Petrevski, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Guillermo Duran, University of Buenos Aires, Argentina
- Igor Ivanovski, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Josef Windsperger, Faculty of Business, Economics and Statistics
- Yoshiaki Jinnai, Emeritus Professor, Tokyo Keizai University, Japan
- Katerina Toshevska-Trpcevska, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Kimberly A. Parker, School of Journalism and Media, University of Kentucky, USA
- Ksenija Dencic Mihajlov, Faculty of Economics, University of Nis, Serbia
- Lidija Pulevska Ivanovska, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Ljupcho Eftimov, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Marija Grishin, Dublin City University, Ireland
- Marko Djogo, Faculty of Economics, University of East Sarajevo, Republic of Srpska, Bosnia and Herzegovina
- Matej Cerne, Ph.D., School of Economics and Business, University of Ljubljana, Slovenia
- Metka Tekavčič, School of Economics and Business, University of Ljubljana, Slovenia

- Mihai Mitascu, Faculty of Economics & Business Administration, West University of Timisoara
- Miroslav N. Jovanovic, University of Geneva, Switzerland
- Momchil Antov, Dimitar A. Tsenov, Academy of Economics, Svishtov, Bulgaria
- Muhammad Abdur Rahman Malik, Suleman Dawood School of Business, Lahore University of Management Sciences (LUMS), Pakistan
- Natalia Ribberink, Hamburg University of Applied Sciences, Faculty of Business & Social Sciences, Germany
- Neven Šerić, Faculty of Economics, University of Split, Croatia
- Nikola Milović, Faculty of Economics, University of Montenegro, Montenegro
- Nina Begicevic Redzep, Faculty of Organization and Informatics, University of Zagreb, Croatia
- Petar Veselinović, Ekonomski fakultet, Univerziteta u Kragujevcu, Serbia
- Predrag Bjelic, Faculty of Economics, University of Belgrade, Serbia
- Predrag Trpeski, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Rezarta Perri, Faculty of Economics, University of Tirana, Albania
- Sanda Renko, Faculty of Economics and Business, University of Zagreb, Croatia
- Sasho Arsov, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Simon Best, Middlesex University Business School, U.K. Simon Best, Middlesex University Business School, U.K.
- Sladjana Benkovic, University of Belgrade, Faculty of Organizational Sciences
- Stanko Stanic, Faculty of Economics, University of Banja Luka, Bosnia and Herzegovina
- Tadija Đukić, Faculty of Economics, University of Nis, Serbia
- Taki Fiti, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje and Macedonian Academy of Sciences and Arts, North Macedonia
- Tareq Rasul, Gulf University for Science & Technology (GUST), Kuwait
- Vincent Charles, School of Management, University of Bradford, U.K.
- Vladimir Filipovski, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia
- Zorica Bozinovska-Lazarevska, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, North Macedonia

## ORGANIZING COMMITTEE

- Nikolina Palamidovska-Sterjadovska, PhD, Chair of the Conference
- Irena Bogoevska-Gavrilova, PhD, Secretary of the Conference
- Marina Trpeska, PhD
- Mijalche Santa, PhD
- Atanasko Atanasovski, PhD
- Daniela Bojadjieva, PhD
- Nikola Levkov, PhD
- Marija Trpkova-Nestorovska, PhD
- Biljana Tashevska, PhD
- Violeta Cvetkoska, PhD
- Elena Markevska Disoska, PhD
- Suzana Makreshanska Mladenoska, PhD
- Bojan Kitanovikj, MSc

---

## TABLE OF CONTENTS

---

### FULL PAPERS

---

#### *Accounting & Finance*

---

**Modeling Earnings under Two Types of Risk.....10**  
*Daisuke Suzuki, Shoji Kamimura and Jumpei Yamada*

**From Basics to Intelligent Tools: Accounting Information and AI in Public-Sector Internal Auditing.....22**  
*Ivan Dionisijev, Zorica Bozhinovska Lazarevska and Todor Tocev*

**Dynamic Connectedness of Eastern European Stock Markets: An Extended Joint Connectedness Approach.....46**  
*Gjorgji Gockov and Goran Hristovski*

**Quantifying FDI's Effects on GDP and Unemployment: Evidence from North Macedonia.....61**  
*Bojan Malchev, Marina Trpeska and Zoran Minovski*

**Digital Financial Inclusion and the Level of Financial Development.....91**  
*Sasho Arsov, Pece Nedanovski and Aleksandar Naumoski*

---

#### *E-Business*

---

**Trust through Data: Managing Risk in the Digital Age.....102**  
*Mojca Kunšek, Barbara Miklič and Metka Tekavčič*

**Behavioral Archetypes of Gen Z in Gamified E-Government Apps.....114**  
*Kalina Trenevaska Blagoeva, Lidija Pulevska Ivanovska, Saso Josimovski, Marina Mijoska Belshoska and Martin Kiselicki*

**AI as a New Academic Ally: The Impact of Artificial Intelligence on Students' Learning Habits.....127**  
*Cvetkoska Violeta, Filip Peovski, Gjore Gakjev, Binela Karamaleska, Elena Avramovikj, Lina Taneska and Aleksandra Peshova*

---

#### *Economics*

---

**Market Competition in Public Procurement of North Macedonia (2021–2025): Economic Analysis of Bidders per Tender.....141**  
*Viktor Mitevski*

**The Macroeconomic Effects of Fiscal Consolidation after Crises: Evidence from a Small Open Economy.....154**  
*Faton Sulejmani and Dragan Tevdovski*

**Digital Technologies and ICT Tools as Enablers of Learning and Innovation in the SME Sector.....166**  
*Faruk Hadžić and Ema Radončić*

**Women, Work, and Birth Rates in Southeastern Europe: A Regional Panel Perspective.....180**  
*Marija Trpkova-Nestorovska*

**Engines and Constraints of Growth in the Western Balkans: Insights for EU Convergence.....194**  
*Marija Trpkova-Nestorovska and Katarina Gacova*

---

### *Management*

---

**A Bibliometric Insight into Machine Learning Applications for Decision-Making.....204**  
*Ivona Serafimoska, Bojan Kitanovikj and Filip Peovski*

**Stakeholders and Sustainability in the Balkan Wine Industry: A Mapping Approach to Governance.....218**  
*Tamara Kaftandjieva and Metka Tekavčič*

**The Long-Run Average Cost Curve: Evidence from the Bottled Water Industry.....231**  
*Bojan Mladenović*

**Shaping the New Business Changemakers: Toward a Novel Taxonomy of Social Entrepreneurship Education.....241**  
*Bojan Kitanovikj, Stojan Debarliev and Aleksandra Janeska-Iliev*

---

### **PANEL PAPERS**

#### *Public Opinion 2.0: Challenges in the Digital Age and Possible Research Directions*

---

**Rethinking Public Opinion Studies in the Digital Era: Expert Discourse on the Theoretical and Methodological Challenges of Public Opinion Research.....258**  
*Elena Negrea-Busuioc, Agnieszka Hess, Agnieszka Stepińska, Anita Ciunova-Shuleska and Anna Bączkowska*

**Expert Discourse on Key Theories and Approaches to Studying Public Opinion: A Concept Map Analysis.....261**  
*Agnieszka Hess, Agnieszka Stepińska, Elena Negrea-Busuioc, Anita Ciunova-Shuleska and Anna Bączkowska*

**Methodological Proposal for Public Opinion Research in the Era of Democratic Regression and Digital Transformation.....263**  
*Kinga Adamczewska, Roksana Gloc, Agnieszka Hess, Agnieszka Stepińska and Anna Bączkowska*

**The Use of VR Technology in Political Communication: Prospects and Challenges in Shaping Public Opinion.....267**  
*Małgorzata Durmaj and Anna Kulma*

**Public Opinion and the Construction of Self- and Other-Identity in Anti-Immigration Discourse Towards Ukrainians in Poland.....270**  
*Anna Bączkowska, Agnieszka Hess, Artur Lipiński, Arkadiusz Misztal, Aneta Dhutek and Joanna Tillack*

---

## **EXTENDED ABSTRACTS**

---

### *Economics*

---

**Effects of Labour Market Measures and Policies on Improving Labour Market Performance in Central and South Eastern European (CSEE) Countries.....273**  
*Kristijan Kozheski, Trajko Slaveski, Predrag Trpeski, Borce Trenovski and Gjunter Merdzan*

**How Do Firms Respond to Minimum Wage Increases in Macedonia?.....279**  
*Dragan Tevdovski, Branimir Jovanovikj and Viktor Stojkoski*

**The Impact of Economic Development on Environmental Quality through Testing the Environmental Kuznets Curve Hypothesis in the Western Balkan Countries.....283**  
*Dorđe Kotarac*

**Life Expectancy and Income Inequality in North Macedonia: An Empirical Analysis.....288**  
*Marijana Cvetanoska, Biljana Tashevska and Predrag Trpeski*

**The Impact of Remittances on Domestic Investment and Consumption Expenditures: The Case of North Macedonia.....293**  
*Bojan Shimbov, Oliver Morrissey and Maite Alguacil*

**The Impact of Political Institutions on Economic Growth in Post-Transition Europe.....299**  
*Gunter Merdzan, Predrag Trpeski and Daniela Bojadjieva*

**The Role of Empirical Evidence and Government Project Office in the Formation of the Slovenian Development Strategy.....305**  
*Aleksandar Kešeljević*

**Digital Infrastructures as the New Pipes of Global Capital.....308**  
*Jasna Tonovska and Predrag Trpeski*

---

### *Marketing, Management & E-Business*

---

**From Human Development to Innovation Outcomes: A Cross-Country and Income Group Analysis.....313**  
*Andrijana Bojadzievska Danevska*

**Neuroscience Techniques in Human Resource Management: An Overview.....316**  
*Bojan Kitanovikj and Ljupcho Eftimov*

**Triggers of Strategic Transition through Sensemaking in Nascent Entrepreneurship.....321**  
*Tea Josimovska, Daniel Nauck and Vladimir Trajkovik*

**Mapping Research on AI and Consumer Purchase Intention: Bibliometric Insights (2009–2025).....325**  
*Snezana Ristevska-Jovanovska, Ivona Serafimovska and Irena Bogoevska-Gavrilova*

**The Role of Message Credibility in Building Influencer Loyalty and Increasing Purchase Intentions.....332**  
*Irena Bogoevska-Gavrilova and Anita Ciunova-Shuleska*

---

## **DOCTORAL SYMPOSIUM PAPERS**

---

**Digital Transformation in Terms of Improving the Performance of Companies in North Macedonia.....338**  
*Bojana Antikj*

**Social Capital and the Role of Social Brokers in AI (Non) Adoption in Developing Countries.....342**  
*Blerton Zejneli and Mijalche Santa*

**Classification of a Foreign Exchange Difference from an Intragroup Monetary Liabilities and Assets in Multinational Telecommunication Companies under IFRS 18.....348**  
*Miroslav Serafimovski*

**Cooperative vs. Non-Cooperative Games: A Comparison of Solutions and Implications.....356**  
*Stefan Mladenoski*

## FULL PAPERS

This section includes full-length research papers presented at the conference that underwent a double-blind peer-review process and are organized into the following areas: Accounting & Finance, E-business, Economics, and Management.

### *Accounting & Finance Papers*

#### MODELING EARNINGS UNDER TWO TYPES OF RISK

**Daisuke Suzuki**

*Showa Women's University, Japan  
d-suzuki@swu.ac.jp*

**Shoji Kamimura**

*Kanagawa University, Japan  
kamimura@kanagawa-u.ac.jp*

**Jumpei Yamada**

*Meiji Gakuin University, Japan  
junyama@eco.meijigakuin.ac.jp*

#### **ABSTRACT**

*This paper develops a unified framework for earnings measurement by introducing two types of risk: inflow-related risk and outflow-related risk. Using the concept of risk tolerance—the acceptable level of stock-related risk—the framework spans the full range of earnings measures, from conservative cash-based accounting to forward-looking economic income. The analysis shows that (i) the asymmetry between inflows and outflows is central to earnings recognition, (ii) varying risk tolerance explains practices such as historical cost and depreciation, and (iii) realization, matching, and conservatism can be reconciled within a risk-based model. The contribution lies in formally linking risk tolerance to the stock–flow structure of accounting, providing a clearer representation of how uncertainty shapes earnings measurement. The framework offers implications for both theory and practice while also facing limitations, including its stylized two-period setting and simplified risk measure. These point to directions for future extensions and empirical validation.*

**Keywords:** *Earnings, Stock-flow congruence, Matching/realization, Risk tolerance, Uncertainty*

**JEL classification:** *M41*

## 1. INTRODUCTION

In recent years, accounting research has increasingly drawn on insights from information economics. At the same time, traditional accounting theory has long emphasized foundational concepts—such as realization, the matching principle, and conservatism—that explain how earnings are determined. While these concepts are fundamental to accounting practice, their reliance on verbal reasoning and interpretive frameworks has often hindered their incorporation

into formal theory and empirical analysis. As a result, these foundational ideas are often excluded from contemporary information-based models. The matching principle, in particular, is a prime example of a concept whose analytical treatment remains limited. To address this gap, this study seeks to develop a generalizable framework that encompasses all logically possible earnings measurement methods. In accounting, earnings are determined by allocating cash inflows and outflows across periods. To evaluate the optimal form of earnings measurement, it is first necessary to delineate the entire spectrum of logically possible methods. Only then can we assess which constraints or assumptions give rise to a normatively preferable method of earnings measurement.

This paper focuses on situations where the relationship between stocks and flows is not straightforward—specifically, cases where changes in stocks cannot be clearly observed. For example, while the disappearance of a physical asset clearly implies a flow (i.e., an expense), depreciation reflects a less transparent decline in value. In such cases, accounting requires a conceptual justification for expense recognition. To address these ambiguities, this study introduces the concept of risk. Specifically, it focuses on risks inherent in stocks arising from cash inflows and outflows, and explores how such risks shape the recognition of stocks and flows in accounting. By incorporating the concept of risk, we show that a wide spectrum of earnings measurement methods—ranging from strict cash-based accounting to economic income—can be expressed in terms of varying degrees of risk tolerance. At the same time, the model enables us to analytically revisit traditional accounting principles and articulate their roles more precisely. In doing so, the study offers a new theoretical justification for foundational concepts whose importance has long been recognized in practice but seldom formalized rigorously.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature, focusing on how risk has been treated in discussions of earnings measurement. Section 3 presents a formal analysis of the stock–flow relationship in accounting. Section 4 introduces the principle of congruence and shows how earnings arise from the intertemporal allocation of cash flows. Section 5 develops a model with two types of risk and demonstrates that earnings can be represented as a function of risk tolerance. The paper yields three key findings. First, under uncertainty, the central issue in earnings measurement lies in the stock–flow structure rather than in the traditional asset–liability versus revenue–expense dichotomy. Second, the proposed framework unifies diverse earnings measures by varying the degree of risk tolerance for inflows and outflows. Third, it shows that single-risk models fail to capture the logic of matching, while also explaining why traditional accounting principles remain institutionally resilient under uncertainty. A preliminary version of this study was previously presented as a poster under the same title, which has limited archival visibility. The present paper substantially extends the model and provides updated analyses and results.

## **2. PREVIOUS STUDIES**

The concept of risk in accounting can be broadly divided into two types (Sunder, 2015). The first is risk as hazard (downside risk), which focuses on the possibility of adverse outcomes such as losses or failures. This perspective is closely related to the traditional accounting principle of conservatism, which calls for early recognition of potential losses and delayed recognition of uncertain gains. Modern IFRS standards have redefined this principle under the concept of prudence, which advises caution in making judgments under conditions of uncertainty.

The second type is risk as dispersion of outcomes, which treats both upward and downward variations as sources of risk. This conception is central in finance and has become increasingly prominent in accounting standards. The IFRS Conceptual Framework refers to uncertainty in terms of existence (whether an asset or liability exists), outcome (the timing and amount of future inflows or outflows), and measurement (estimation uncertainty) (IASB, 2018, paras. 5.14, 6.61, 5.20). However, these references often acknowledge risk without specifying how accounting information should represent or convey it effectively to users (Barker & Penman, 2020).

The Accounting Standards Board of Japan (ASBJ) has addressed this issue in its Discussion Paper on *the Conceptual Framework for Financial Accounting*. It emphasizes that accounting information should help investors assess the extent to which expected investment outcomes have been realized, introducing the concept of "*release from investment risk*" (ASBJ, 2006, para. 3.23). This concept differentiates between business investments, where revenue is recognized upon realization of operating cash flows, and financial investments, where changes in fair value can be immediately recognized in earnings due to asset liquidity.

Penman (2016) further contrasts historical cost and fair value accounting in their treatment of risk. Historical cost accounting delays recognition until uncertainty is resolved, whereas fair value accounting incorporates forecasts of future cash flows, thereby making risk visible only retrospectively through earnings volatility. Penman (2016) argues that, despite its forward-looking nature, fair value accounting may mislead users by incorporating unrealized expectations into reported earnings. On the other hand, Barker and Penman (2020) classify earnings recognition into four types of uncertainty resolution: Type 1 (direct matching of revenues and expenses), Type 2 (ex ante matching, e.g., depreciation), Type 3 (ex post matching, e.g., fair value changes), and Type 4 (no matching, e.g., immediate expensing of R&D). While Types 3 and 4 challenge the matching principle, they highlight how accounting standards address different forms of uncertainty resolution.

In a model-based analysis, Wagenhofer (2003) identified conditions under which accrual accounting outperforms cash flow as a performance measure. Dutta and Zhang (2002) analyzed how different revenue recognition rules affect managerial decision-making and showed that desirable earnings tend to embody a conservative bias. However, these studies primarily focus on risks associated with cash inflows in revenue recognition, while treating the matching principle as a given assumption rather than a subject of analysis. In a different approach, Konstantinidi and Pope (2016) examined the risk inherent in earnings themselves. They showed that even a parsimonious model based on accrual earnings, cash flows, and special items can reasonably predict the distributional shape of earnings.

### **3. STOCK AND FLOW**

#### **3.1. Conceptual Foundations of Stocks and Flows**

In earnings measurement, a long-standing debate concerns whether it should be based on stocks or flows. Ultimately, the key issue is how the relationship between stocks and flows is conceptualized. In economics, a stock is defined as a quantity at a specific point in time, while a flow represents the change in that quantity over a period (Mankiw, 2016). We adopt the same view: a stock refers to the quantity of an item at a point in time, and a flow is the change in that item over a period. Whereas a stock can be defined independently, a flow is necessarily derived from changes in stocks.

A stock measure alone is often insufficient for decision-making under uncertainty. If changes in stocks followed a deterministic rule, flows could be directly inferred. In reality, however, future changes in stocks are uncertain. Understanding such changes requires decomposing past flows—for example, distinguishing between inflows and outflows—in order to identify their underlying causes. This decomposition is essential for forecasting. Therefore, both stock and flow measures—particularly the explanatory content of flows—are indispensable for providing decision-useful information under uncertainty.

### **3.2. Stock and Flow in Financial Statements**

Financial statements embody this distinction: the balance sheet reports stocks, and the income statement reports flows. The FASB (1976) distinguishes two approaches: the asset–liability approach, which emphasizes measuring assets and liabilities and defines earnings as changes in net resources; and the revenue–expense approach, which focuses on recognizing revenues and expenses and defines earnings as their difference. Yet the essential issue lies in the stock–flow relationship. Can equity be derived from previously defined revenues and expenses—in other words, can stocks be constructed from flows?

The IASB (1989) defines revenue as increases in equity from asset inflows or liability reductions, excluding capital transactions. This definition presupposes the prior concepts of assets, liabilities, and equity. Without them, revenues and expenses cannot be meaningfully defined. Moreover, if capital transactions are not clearly separated, classifying flows becomes problematic. Thus, accounting flows depend on prior stock definitions. Flows represent changes in stocks, and their definition depends on identifying the underlying stock items. However, this does not resolve the debate. Even when stocks are defined first, their scope may be adjusted to serve the purpose of flow measurement. The key question is not whether to prioritize stocks or flows, but rather how stocks are defined. The historical development of accounting may have blurred these distinctions. Framing the debate merely as a choice between the asset–liability and revenue–expense approaches yields limited insight.

Prior research (e.g., Kusano, 2012) links the asset–liability approach to fair value measurement, and the revenue–expense approach to historical cost. The former emphasizes valuing assets and liabilities, while the latter focuses on recognizing revenues and expenses. However, valuing assets and liabilities does not necessitate the exclusive use of fair value. Fair value typically reflects equilibrium market prices, which may diverge from a firm’s specific economic value. As Barker and Penman (2020) note, fair value involves substantial uncertainty, especially for long-term investments without active markets, which can mislead investors. Under uncertainty, identifying which elements are certain becomes critical. Although not market-based, historical cost indicates how capital has been invested and provides a conceptual justification. Even within the asset–liability approach, fair value is not mandated. The FASB (1976) explicitly states that neither approach is tied to a specific measurement basis. Both fair value and historical cost can, in principle, be valid under either framework. Asset and liability measurement is also closely tied to principles such as realization and matching, further complicating the simplistic link between measurement approach and basis.

## **4. THE PRINCIPLE OF CONGRUENCE**

As discussed in the previous section, accounting flows are defined as changes in stocks, especially in the context of measuring equity and earnings. Rather than treating stocks and flows as opposing concepts, we have argued that both serve as equally essential sources of information for financial reporting. In this section, we clarify the relationship between cash

flows and earnings—a relationship that underpins the measurement of equity. We then further examine the fundamental nature of stocks and flows in accounting. Specifically, we analyze how changes in cash during each accounting period are allocated over time and represented as accounting stocks and flows. This analysis lays the foundation for our subsequent focus on the risks embedded in cash flows and how these risks influence the structure of earnings measurement.

We assume a setting without uncertainty at any point in time  $t(= 0, \dots, T)$ , where  $T$  denotes the time of the firm's liquidation. Under this assumption, the firm's assets  $A_t$  are equal to the sum of its liabilities  $L_t$  and equity  $E_t$ , as expressed by the basic balance sheet identity:

$$A_t = L_t + E_t \quad (1)$$

While accounting classifications may be more detailed in practice, assets  $A_t$  can generally be divided into three categories: cash equivalents  $C_t$ , uncollected items representing future income  $FI_t$ , and capitalized items resulting from past cash outflows  $PO_t$ . These items are recorded as economic resources on the balance sheet. Similarly, liabilities  $L_t$  can be divided into past inflows  $PI_t$ , such as deferred revenues, and future outflows,  $FO_t$  including accrued expenses and obligations like borrowings. These classifications can be expressed as follows:

$$A_t = C_t + FI_t + PO_t \quad (2)$$

$$L_t = PI_t + FO_t \quad (3)$$

As shown in Equation (1), equity  $E_t$  on the balance sheet is defined as the difference between assets  $A_t$  and liabilities  $L_t$ . Given the linkage between the balance sheet and the income statement, equity can also be understood as the sum of contributed capital  $S_t$ , representing capital contributions from shareholders, and retained earnings  $R_t$ , which consist of past earnings not distributed as dividends. These relationships are expressed as:

$$E_t = S_t + R_t \quad (4)$$

Retained earnings  $R_t$ , as a component of equity, increases with current-period earnings  $e_t$  and decrease with dividends  $d_t$  paid during the same period. In other words, earnings can be viewed as a flow, calculated as the difference between revenues and expenses, or as a stock-based measure, derived from the change in retained earnings adjusted for dividends. This relationship holds regardless of any changes in contributed capital  $S_t$  and can be expressed as:

$$e_t = R_t - R_{t-1} + d_t \quad (5)$$

The relationship between earnings  $e_t$  and cash flows can be formally derived as follows. Substituting Equation (4) into Equation (5), we obtain:  $e_t = R_t - R_{t-1} + d_t = (E_t - S_t) - (E_{t-1} - S_{t-1}) + d_t$ . Next, substituting Equation (1) into this expression, and applying Equations (2) and (3), we derive the following. Here, let  $\Delta_t$  denote the change in variable between time  $t$ .

$$\begin{aligned} e_t &= (A_t - L_t - S_t) - (A_{t-1} - L_{t-1} - S_{t-1}) + d_t \\ &= (C_t + FI_t + PO_t - PI_t - FO_t - S_t) \\ &\quad - (C_{t-1} + FI_{t-1} + PO_{t-1} - PI_{t-1} - FO_{t-1} - S_{t-1}) + d_t \\ &= \Delta C_t + \Delta FI_t + \Delta PO_t - \Delta PI_t - \Delta FO_t - \Delta S_t + d_t \end{aligned} \quad (6)$$

As shown in the preceding discussion, Equation (6) demonstrates that earnings  $e_t$  are determined by the temporal allocation of cash flows and related accrual adjustments. Specifically, the second through fifth terms in Equation (6) represent changes in the following accrual-related items: (a) accrued revenues (i.e., cash inflows to be received in future periods), (b) deferred expenses, such as capital investments to be expensed in later periods, (c) deferred

revenues (i.e., cash inflows received in past periods but not yet recognized), and (d) accrued expenses (i.e., cash outflows to be paid in future periods). Over the firm's full life span, up to liquidation, these accruals eventually offset one another: accrued revenues and accrued expenses are realized as cash flows in future periods, while deferred revenues and expenses are recognized as revenues and expenses, respectively. As a result, their cumulative effect becomes zero when aggregated across all periods. Accordingly, the cumulative sum of earnings can be expressed as:

$$\sum e_t = \sum \Delta C_t - \sum \Delta S_t + \sum d_t \quad (7)$$

Equation (7) shows that the cumulative amount of earnings is equal to total cash inflows and outflows, excluding transactions related to capital contributions and distributions. In other words, regardless of the specific method used to calculate earnings, when accumulated over the firm's entire life up to liquidation, total earnings converge to net cash inflows minus outflows, adjusted for capital transactions. This fundamental property is known as the *principle of congruence*, which underscores the central role of cash flows in modeling and interpreting earnings over time.

## 5. EQUITY AND EARNINGS FROM THE PERSPECTIVE OF TWO RISKS

In the previous section, we examined how both equity (a stock) and earnings (a flow) arise from the allocation of cash inflows and outflows over time. Earnings, as increases in equity, represent the outcomes of investment activities once associated uncertainties have been resolved. In this sense, the concept of earnings is fundamentally linked to risk. Barker and Penman (2020) argue that the traditional dichotomy between the asset-liability and revenue-expense approaches is misleading. They advocate for a framework in which earnings are interpreted through the lens of uncertainty. Schmalenbach (1955) similarly emphasized that both stocks and flows are essential components of earnings. In this section, we revisit the concept of earnings by focusing on two distinct but interrelated types of risk: those associated with cash inflows and those with cash outflows.

### 5.1. Earnings Measurement from the Perspective of Risk

According to IASB (2018), uncertainty in accounting arises in two forms: outcome uncertainty, concerning variability in cash flows, and existence uncertainty, concerning whether an asset or liability exists. However, under the clean surplus assumption, stock and flow are interdependent, casting doubt on the utility of treating these uncertainties separately. Barker and Penman (2020) propose that uncertainty should instead be understood as a unified concept. This unified view aligns with ASBJ (2006), which defines revenues and expenses as the portions of asset and liability changes *released from investment risk*. Once invested resources contribute to generating cash flows, they are regarded as having been discharged from risk. This approach applies equally to assets, revenues, and expenses, thereby supporting a single, integrated risk perspective.

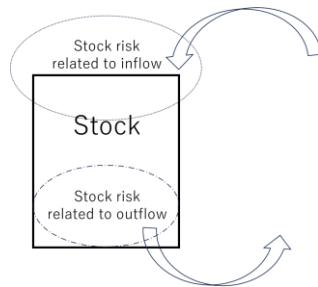
Sakurai (2007) interprets this notion of “released from investment risk” as a conceptual substitute for the realization principle. Because expenses are recognized concurrently with revenues, the matching principle is also implicitly embedded within this view. Moreover, since risk exposure is typically tied to investment activities, the notion of *control*—that is, whether a firm has authority over the use of resources—may also be relevant. Taken together, the unified risk perspective appears to encompass several foundational accounting principles. However, reducing these principles to a single framework risks oversimplifying their distinct roles and normative functions. Earnings are determined by allocating cash inflows and outflows

across time—two components that differ in both dynamics and informational content. Although earnings are ultimately observed as a single variable, a single-risk approach may conceal the inherent asymmetry between inflows and outflows. We therefore proceed to analyze risk in earnings measurement from both perspectives.

## 5.2. Risk and Tolerance in Cash Inflows and Outflows

We model equity and earnings as functions of two types of stock-related risk: (1) *inflow-related risk*, associated with assets that may generate future revenue, and (2) *outflow-related risk*, tied to assets that give rise to future expenses. These risks can be represented, for analytical purposes, by the variability (e.g., standard deviation) of relevant stock items. Because both assets and liabilities may embody elements of each risk type, stocks and flows must be treated in an integrated fashion.

A basic accounting assumption is that only identifiable and verifiable economic resources are recognized on the balance sheet. When the existence and use of a resource are certain—such as risk-free cash compensation—it can be recognized without explicit reference to risk. By contrast, when the existence of a resource is uncertain, the concept of risk becomes essential to determining whether it qualifies as stock.



Barker and Penman (2020) classify earnings measurement into four types. In Type 1, involving the direct matching of revenues and expenses, some elements may not involve explicit risk. For example, in some merchandise sales, inflow-related stock risk exists at the time of purchase, but revenue is recognized only once the sale occurs and the right to receive cash is certain. Here, revenue recognition corresponds to the resolution of risk. Expense recognition, in contrast, follows the disappearance of the asset upon delivery, rather than risk resolution per se.

For outflows, as illustrated by depreciation, the associated risk is often difficult to quantify and may not evolve linearly. Even as time passes, the exact extent of depreciation remains uncertain. Accounting practice therefore, seeks to exclude high-risk stock items from recognition, consistent with the principle that non-existent resources should not be reported. A central issue, then, is to determine whether a given item surpasses the threshold for recognition—effectively, evaluating the degree of risk attached to both inflows and outflows.

To formalize this idea, we define *risk tolerance* as the acceptable level of stock-related risk in calculating equity and earnings. For example, accounts receivables are generally recognized as revenues despite some collection risk, as they are presumed to be relatively low risk. Similarly, unrealized gains on marketable securities may be recognized despite uncertainty, due to the liquidity of their market value. The degree to which risk is incorporated into earnings depends on both the nature of each transaction and broader economic conditions. In practice, uncertainty can never be fully eliminated. Accordingly, earnings measurement should be understood as a process that reflects the degree of risk tolerance embedded in accounting recognition.

### 5.3. Modeling Risk Tolerance in Earnings Measurement

It is possible to represent all forms of earnings calculation methods—from strictly cash-based accounting to economic income—within a unified framework by adjusting the risk tolerances associated with inflows and outflows of stock. Such a framework enables comprehensive comparison of alternative approaches to earnings measurements and provides a basis for evaluating what may constitute an optimal earnings measure.

For cash inflows, when the tolerance for stock-related risk is extremely low, earnings are not recognized until the inflow becomes certain, and only cash is recognized as stock. Non-cash assets are excluded from recognition because they carry the risk of non-realization. Conversely, under high risk tolerance, expected future inflows may be recognized as stock, with increases recorded as earnings. Thus, various forms of revenue recognition can be represented by adjusting the level of tolerance for inflow risk. For cash outflows, when tolerance for stock-related risk is very low, all cash outflows—including those related to fixed assets—must be immediately expensed. Even if made with investment intent, such outflows are treated entirely as expenses because their value is expected to decline over time. By contrast, with high tolerance, the asset is retained on the balance sheet despite the associated risk, and the related expense is deferred until the decline in value becomes sufficiently certain.

According to this model, adjusting the levels of the two types of risk tolerance allows us to represent the entire spectrum of earnings measurement systems—from conservative, cash-based approaches to forward-looking concepts of economic income. In practice, accounting standards position themselves somewhere between these two extremes. If the entire range is made explicit within the model, it is reasonable to assume that an optimal earnings measure from a social perspective exists within it. Furthermore, by examining how current accounting standards incorporate these two forms of risk, we can concretely assess the gap between existing practices and theoretically optimal measures. To support such an analysis, we now introduce a stylized model based on a set of simplifying assumptions.

As an example, we propose a simplified model to illustrate how the calculation of revenues and expenses depends on the degree of risk tolerance, given the relationship between investment and cash flows. We consider a two-period model with time points  $t = 0, 1, \text{ and } 2$ . At time  $t = 0$ , a cash outflow of  $I_0$  occurs as an investment, and the investment is assumed to expire worthless at  $t = 2$ . A tilde ( $\sim$ ) indicates a stochastic variable. Let  $\tilde{I}_t$  denote the stochastic value of the investment, and let  $\tilde{C}_t$  denote the stochastic cash inflow generated by this project at time  $t$ . Their realizations are denoted by  $\tilde{I}_t$  and  $\tilde{C}_t$ , respectively. For simplicity, we assume no cash outflows occur at  $t = 1$  or  $t = 2$ .

Let  $E_t[\cdot]$  denote the conditional expectation given the information available at time  $t$ , and let  $\sigma_t[\cdot]$  denote the corresponding measure of risk. The notion of risk here may refer to any quantity derived from the conditional distribution of the stochastic variable; for example, it may be specified as the standard deviation. To classify the degree of risk, let  $\alpha$  and  $\beta$  be two positive real numbers with  $\alpha \leq \beta$ . These values serve as thresholds for evaluating risk:

- If  $\sigma_s[\tilde{C}_t] \leq \alpha$ , the cash flow  $\tilde{C}_t$  is considered low risk at time  $s$ .
- If  $\sigma_s[\tilde{C}_t] \geq \beta$ , the cash flow  $\tilde{C}_t$  is considered high risk.
- If  $\alpha < \sigma_s[\tilde{C}_t] < \beta$ , the risk is regarded as moderate.

We refer to the pair  $(\alpha, \beta)$  as the risk tolerance parameters. The same classification applies analogously to  $\sigma_s[\tilde{I}_t]$  as well.

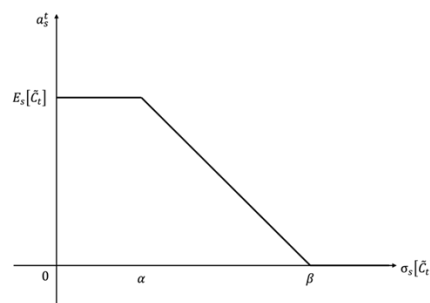
We now define earnings as follows. Let  $a_t$  and  $b_t$  denote the revenue and expense at time  $t$ , respectively. Then, the earnings  $e_t$  at time  $t$  is given by  $e_t = a_t - b_t$ . In this context, both  $a_t$  and  $b_t$  are specified in terms of the stochastic variable  $\tilde{I}_t$  and  $\tilde{C}_t$  with particular emphasis on the risks.

Let the risk tolerance for the cash inflow  $\tilde{C}_t$  (for  $t = 1, 2$ ) be given by the pair  $(\alpha, \beta)$ . Based on this, we define the contribution of  $\tilde{C}_t$  to the revenue  $a_s$  at time  $s$ , denoted by  $a_s^t$ , sequentially from  $s = 0$  as follows. First, if  $\sigma_0[\tilde{C}_t] \leq \alpha$ , that is, if the risk of  $\tilde{C}_t$  at time 0 is considered low, then  $a_0^t = E_0[\tilde{C}_t]$ . If  $\sigma_0[\tilde{C}_t] \geq \beta$ , meaning that the risk is considered high at time 0, then  $a_0^t = 0$ . For the intermediate case where  $\alpha < \sigma_0[\tilde{C}_t] < \beta$ , we define  $a_0^t$  as the value of a continuous function of  $\sigma_0[\tilde{C}_t]$  that passes through the two points  $(\alpha, E_0[\tilde{C}_t])$  and  $(\beta, 0)$ . For simplicity, we adopt the following linear specification for  $a_0^t$  when the risk of  $\tilde{C}_t$  lies between  $\alpha$  and  $\beta$ :

$$a_0^t = -\frac{E_0[\tilde{C}_t]}{\beta - \alpha} \sigma_0[\tilde{C}_t] + \frac{\beta E_0[\tilde{C}_t]}{\beta - \alpha}$$

Summarizing the definition of  $a_0^t$ , we have:

$$a_0^t = \begin{cases} E_0[\tilde{C}_t], & \sigma_0[\tilde{C}_t] \leq \alpha \\ \left(\frac{\beta - \sigma_0[\tilde{C}_t]}{\beta - \alpha}\right) E_0[\tilde{C}_t], & \alpha < \sigma_0[\tilde{C}_t] < \beta \\ 0, & \sigma_0[\tilde{C}_t] \geq \beta \end{cases}$$



Given the above definition of the contribution of each cash inflow  $\tilde{C}_t$  to revenue, the total revenue at time  $s = 0$  is expressed as  $a_0 = a_0^1 + a_0^2$

Expenses at time  $t$  are defined on the basis of the decline in asset value  $\tilde{I}_{t-1} - \tilde{I}_t$ . Let the risk tolerance for this reduction be given by the pair  $(\alpha', \beta')$  with  $(\alpha' < \beta')$ . The contribution of period  $t$  to expenses at time 0, denoted by  $b_0^t$ , is defined as follows:

$$b_0^t = \begin{cases} 0, \sigma_0[\tilde{I}_{t-1} - \tilde{I}_t] \leq \alpha' \\ \left( \frac{\sigma_0[\tilde{I}_{t-1} - \tilde{I}_t] - \alpha'}{\beta' - \alpha'} \right) E_0[\tilde{I}_{t-1} - \tilde{I}_t], \alpha' < \sigma_0[\tilde{I}_{t-1} - \tilde{I}_t] < \beta' \\ E_0[\tilde{I}_{t-1} - \tilde{I}_t], \sigma_0[\tilde{I}_{t-1} - \tilde{I}_t] \geq \beta' \end{cases}$$

By defining expenses in this manner, total expenses at time 0 are given by  $b_0 = b_0^1 + b_0^2$ . With an appropriately chosen pair  $(\alpha', \beta')$ , any expense  $b_0$  satisfying  $0 < b_0 < I_0$  can be represented. Accordingly, when combined with the definition of revenues, it follows that by specifying suitable levels of risk tolerance, any earnings  $e_0$  satisfying  $-B_0 < e_0 < E[\tilde{C}_1] + E[\tilde{C}_2]$  can be expressed.

Finally, we decompose the stochastic reduction in asset value into a component correlated with revenues and an orthogonal residual:  $\tilde{I}_{t-1} - \tilde{I}_t = \gamma_t \tilde{C}_t + \tilde{\epsilon}_t$ , where

$$\gamma_t = \frac{\text{Cov}[\tilde{I}_{t-1} - \tilde{I}_t, \tilde{C}_t]}{\text{Var}_0[\tilde{C}_t]}$$

is a constant, and  $\tilde{\epsilon}_t$  is orthogonal to  $\tilde{C}_t$ , i.e.  $\text{Cov}[\tilde{C}_t, \tilde{\epsilon}_t] = 0$ . The term  $\gamma_t \tilde{C}_t$  represents the portion of the asset reduction that moves together with revenues, while  $\tilde{\epsilon}_t$  represents the reduction independent of revenues. Since  $E_0[\tilde{I}_{t-1} - \tilde{I}_t] = \gamma_t E_0[\tilde{C}_t] + E_0[\tilde{\epsilon}_t]$ , the expected reduction  $\gamma_t E_0[\tilde{C}_t]$  can naturally be regarded as the expense corresponding to the expected revenue. Hence, if the recognized revenue is  $a_0^t$ , the corresponding expense is naturally defined as  $\gamma_t a_0^t$ .

For the residual component  $E_0[\tilde{\epsilon}_t]$ , expenses are recognized depending on risk tolerance. Let the risk tolerance for  $\tilde{\epsilon}_t$  be given by  $(\alpha', \beta')$ . Then, the contribution of period t to expenses at time 0 is defined as

$$b_0^t = \begin{cases} 0, \sigma_0[\tilde{\epsilon}_t] \leq \alpha' \\ \left( \frac{\sigma_0[\tilde{\epsilon}_t] - \alpha'}{\beta' - \alpha'} \right) E_0[\tilde{\epsilon}_t], \alpha' < \sigma_0[\tilde{\epsilon}_t] < \beta' \\ E_0[\tilde{\epsilon}_t], \sigma_0[\tilde{\epsilon}_t] \geq \beta' \end{cases}$$

Summarizing, the contribution of the period to expenses at time 0 is defined as  $\gamma_t a_0^t + b_0^t$ . Thus, the model provides a unified representation of revenues and expenses under different levels of risk tolerance, thereby reconciling the principles of realization, matching, and conservatism within a single analytical framework.

## 6. CONCLUSION

This paper has examined earnings measurement from the perspective of two types of risk—those associated with cash inflows and those associated with cash outflows. By introducing the concept of risk tolerance, we proposed a unified framework that encompasses approaches ranging from conservative, cash-basis accounting to forward-looking concepts of economic income. In this sense, we demonstrate that earnings measurement is ultimately a matter of how risk is treated within accounting.

Our analysis yields three key findings. First, the asymmetry between inflows and outflows lies at the core of earnings recognition. Second, varying levels of risk tolerance reproduce diverse accounting practices, such as historical cost and depreciation. Third, traditional principles of realization, matching, and conservatism can be formally reconciled within a risk-based model. Compared with prior literature, the main contribution of this study is to provide a formal representation of risk tolerance that links stock–flow structures with earnings measurement, thereby offering a theoretical framework that spans the entire spectrum of logically possible methods.

The proposed framework carries implications for both theory and practice. It provides a basis for integrating foundational accounting concepts with modern information-based approaches, clarifies why traditional principles remain resilient under uncertainty, and helps interpret how existing standards implicitly reflect assumptions about risk.

## REFERENCES

- Accounting Standards Board of Japan (ASBJ) (2006). *Discussion paper: Conceptual framework of financial accounting*, ASBJ.
- Barker, R., and S. Penman (2020). “Moving the conceptual framework forward: Accounting for uncertainty”, *Contemporary Accounting Research*, Vol. 37, pp. 322-357.
- Dichev, I. (2008). “On the balance sheet-based model for financial reporting”, *Accounting Horizons*, Vol.22, No. 4, pp. 453-470.
- Dutta, S., and Zhang, X.J. (2002). “Revenue recognition in a multiperiod agency setting”, *Journal of Accounting Research*, Vol. 40, No. 1, pp. 67–83. <https://doi.org/10.1111/1475-679X.00039>
- Financial Accounting Standards Board (FASB) (1976). *An analysis of issues related to conceptual framework for financial accounting and reporting: Elements of financial statements and their measurement, discussion memorandum*, Norwalk, CT: FASB.
- International Accounting Standard Board (IASB) (1989). *Framework for the preparation and presentation of financial statements*, IASB.
- International Accounting Standards Board (IASB). (2018). *Conceptual Framework for Financial Reporting*, London: IFRS Foundation.
- Konstantinidi, T., and Pope, P. F. (2016). “Forecasting risk in earnings”, *Contemporary Accounting Research*, Vol. 33, No. 2, pp. 487-525. <https://doi.org/10.1111/1911-3846.12149>
- Kusano, M. (2012). “Does the Balance Sheet Approach Improve the Usefulness of Accounting Information?”, *The Japanese Accounting Review*, Vol. 2, pp. 139-152.
- Mankiw, N. G. (2016). *Principles of Economics* (7th ed.). Cengage Learning.
- Nissim, D., and S. Penman (2008). *Principles for the application of fair value accounting, White Paper 2*, Center for Excellence in Accounting and Security Analysis, Columbia Business School.
- Paton, W. A., and A. C. Littleton (1940). *An Introduction to Corporate Accounting Standard, American Accounting Association Monograph 3*, American Accounting Association.
- Penman, S. H. (2016). “Valuation: Accounting for risk and the expected return”, *Abacus*, Vol. 52, No. 1, pp. 106-130. <https://doi.org/10.1111/abac.12077>
- Sakurai, H. (2007). Convergence of conceptual frameworks. In S. Saito (Ed.), *Tōgi shiryō: Zaimu kaikei no gainen framework [Discussion materials: Conceptual framework of financial accounting]* (2nd ed). Tokyo: Chuokeizai-sha. (in Japanese).
- Schmalenbach, E. (1955). *Dynamische Bilanz. 12Aufl.*, Westdeutscher Verlag.
- Sunder, S. (2015). “Risk in Accounting”, *Abacus*, Vol. 51 No.4, pp. 536-548.

- Suzuki, D., Yamada, J., and Kamimura, S. (2024), “The perspective of cash flow risks on earnings measurement”, *ASEAN Journal on Science and Technology for Development*, Vol. 41 No. 2, pp. 150-155. <https://doi.org/10.61931/2224-9028.1578>
- Wagenhofer, A. (2003). “Accrual-based compensation, depreciation and investment decisions”, *European Accounting Review*, Vol. 12, No. 2, pp. 287-309.

## FROM BASICS TO INTELLIGENT TOOLS: ACCOUNTING INFORMATION AND AI IN PUBLIC-SECTOR INTERNAL AUDITING

**Ivan Dionisijev**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
ivan.d@eccf.ukim.edu.mk*

**Zorica Bozhinovska Lazarevska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
zoricab@eccf.ukim.edu.mk*

**Todor Tocev**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
todor.tocev@eccf.ukim.edu.mk*

### **ABSTRACT**

*This study examines the integration of artificial intelligence (AI) in public sector internal auditing, focusing on the extent of AI adoption, the types of AI tools used, and the challenges faced by auditors in implementation. The research employs both descriptive statistical analysis and inferential techniques, including Spearman's correlation and ANOVA, to assess the relationships between AI adoption and institutional factors. Findings indicate that while data analytics is the most commonly used AI tool, a significant proportion of respondents do not utilize AI in their auditing practices. The primary barriers to AI adoption include a lack of training, high costs, and concerns regarding data privacy. The study further reveals that AI usage varies depending on the type of institution in which they work. These insights contribute to the ongoing discussion on digital transformation in auditing, emphasizing the need for enhanced training programs and strategic investments to facilitate AI integration.*

**Keywords:** *Internal auditing, Accounting information, Artificial Intelligence (AI), Public sector*

**JEL classification:** *M42, H83*

### **1. INTRODUCTION**

The increasing complexity of public sector financial management requires effective internal auditing practices to ensure accountability, transparency, and governance. Internal auditors play a crucial role in interpreting and applying accounting information to assess fiscal health, mitigate risks, and support evidence-based decision-making. In recent years, digital transformation has introduced new tools and technologies that are reshaping traditional audit functions, with artificial intelligence (AI) emerging as a key driver of change.

AI has the potential to revolutionize public sector internal auditing by automating complex tasks, improving fraud detection, and enhancing data analysis accuracy. Recent studies emphasize that AI-driven technologies enable auditors to process vast amounts of data efficiently, identify patterns of financial anomalies, and strengthen risk assessment mechanisms (Aldemir and Uysal, 2024). However, while AI presents significant opportunities, its integration also raises concerns regarding data privacy, ethical risks, and the potential

diminishing role of human judgment in auditing decisions (Fedyk *et al.*, 2022). The balance between leveraging AI for efficiency and maintaining the core principles of accountability and transparency remains a critical challenge for auditors in the public sector (Setyaningrum *et al.*, 2022).

Given the rapid advancements in AI adoption, it is essential to assess how public sector internal auditors are adapting to these changes and whether they possess the necessary competencies to effectively utilize AI-driven tools. The Institute of Internal Auditors (IIA) has recognized this need by outlining a framework for AI competencies, emphasizing the importance of continuous professional development (IIA, 2017a). However, gaps remain in understanding how internal auditors in different national contexts, such as North Macedonia, integrate AI into their audit processes and whether current training initiatives adequately prepare them for an AI-enhanced environment.

This paper aims to explore how internal auditors in the public sector currently use accounting information in their auditing activities. Understanding their reliance on traditional accounting records versus AI-generated financial data will provide insights into the dynamic role of accounting information in public sector audits. Additionally, the research seeks to evaluate both the perceived and actual impacts of AI integration on audit efficiency, data accuracy, risk assessment, and decision-making processes. Furthermore, this research analyzes the benefits and challenges internal auditors face when adopting AI technologies. At the same time, the study assesses potential barriers to AI adoption, including issues related to ethical concerns, regulatory compliance, and organizational resistance.

## **2. LITERATURE REVIEW**

AI has significantly influenced internal auditing by enhancing automation, improving fraud detection, and optimizing decision-making processes (Fedyk *et al.*, 2022). AI-driven auditing techniques, including machine learning and data analytics, enable auditors to efficiently analyze large datasets, identify anomalies, and streamline financial assessments (Allami, 2022). De Sousa *et al.* (2019) emphasize that AI's role in auditing is expanding beyond automation, providing auditors with predictive capabilities and deeper insights into risk assessment and compliance. Babina *et al.* (2020) suggest that AI adoption in auditing correlates with improved firm performance and innovation, particularly when leveraged for advanced data-driven decision-making. However, while existing literature explores AI's impact on private sector auditing, there remains a gap in understanding how AI is transforming internal auditing in the public sector, particularly in countries like North Macedonia.

The integration of AI requires auditors to develop specific competencies that bridge traditional auditing skills with technological advancements (Aldemir and Uysal, 2024). The Institute of Internal Auditors (IIA, 2017a) highlights that while auditors do not need to become data scientists, they must develop a foundational understanding of AI tools and their application in audit functions. Setyaningrum *et al.* (2022) argue that competency gaps among auditors remain a challenge, as many professionals lack adequate training in AI-related technologies, impacting their ability to conduct efficient audits. Issa *et al.* (2016) further emphasize the need for universities and professional organizations to incorporate AI-related auditing courses in their curricula to bridge this knowledge gap.

AI enhances fraud detection capabilities in auditing by using predictive analytics and anomaly detection to identify fraudulent transactions (Rehman and Hashim, 2022). Fedyk *et al.* (2022) further illustrate how AI-powered algorithms can improve risk assessment by identifying

patterns of financial misstatements and fraudulent activities that might be overlooked by traditional methods. Raji and Buolamwini (2022) explore how AI audits can expose biases in algorithmic decision-making, ensuring transparency and accountability in automated auditing systems. The use of AI in fraud detection is particularly significant in the public sector, where financial mismanagement and corruption remain pressing concerns (Wirtz *et al.*, 2019).

Recent empirical research demonstrates that AI investments lead to improvements in audit quality and efficiency (Fedyk *et al.*, 2022). AI adoption has been shown to reduce audit restatements by 5%, lower audit fees, and enhance the overall reliability of financial reporting (Bughin *et al.*, 2017; Babina *et al.*, 2020). The ability of AI to process vast datasets and apply predictive analytics allows auditors to detect anomalies more effectively and focus on high-risk areas (Rock, 2020). Frey and Osborne (2017) argue that auditing is one of the professions most exposed to technological advancements, making AI integration essential for maintaining audit integrity and effectiveness.

While AI offers various benefits in internal auditing, it also presents ethical challenges, particularly in data privacy, bias, and algorithmic transparency (Wirtz *et al.*, 2019). The OECD (2018) emphasizes the importance of ethical AI adoption, calling for robust regulatory frameworks to ensure fairness and accountability. Minkkinen *et al.* (2022) discuss the need for continuous auditing of AI applications to prevent biases and inaccuracies in financial reporting. The IIA (2017b) underscores that AI should be used as an enhancement tool rather than a replacement for human auditors, ensuring that ethical considerations remain at the forefront of AI integration. Burton and Rezaee (1994) advocate for ongoing ethics training for auditors to address these emerging challenges. Despite these discussions, research has not yet provided sufficient insights into how public sector internal auditors perceive and manage AI-related ethical dilemmas, particularly in countries like North Macedonia, where regulatory frameworks for AI auditing may still be evolving.

As AI continues to develop, its role in internal auditing is expected to grow, necessitating continuous professional development for auditors (Burton and Rezaee, 1994). Issa *et al.* (2016) propose that AI-driven auditing will become the standard practice in the future, requiring internal auditors to adopt new methodologies and tools for enhanced financial governance. The IIA (2021) suggests that organizations should invest in AI training programs to equip internal auditors with the necessary skills to navigate this dynamic environment. Further, the integration of blockchain with AI is emerging as a future trend that could enhance the transparency and security of audit processes (Frey and Osborne, 2017). However, while the theoretical benefits of AI in auditing are widely acknowledged, research gaps remain regarding how internal auditors in the public sector can best leverage AI to improve transparency and accountability, and what barriers - whether technological, organizational, or regulatory - hinder full adoption.

### **3. METHODOLOGY**

#### **3.1. Research design**

This study employs a quantitative research design using a structured survey to collect data from internal auditors working in public sector institutions in North Macedonia. The survey aims to assess how auditors utilize accounting information, their perceptions and use of AI in internal auditing, and the benefits and challenges associated with AI adoption in the public sector. The collected data is analyzed using statistical techniques to identify trends, correlations, and comparative differences based on the internal auditors' institutional affiliations.

### 3.2. Survey structure

The survey is divided into six distinct sections, each designed to capture specific aspects of the research objectives (see Table 1).

Table 1: Survey's structure

Section	Description	Question	Type of answers
General Characteristics of Respondents	Demographic and professional background information of respondents	Job title	Open-ended response
		Years of experience in internal auditing	Categorical response: Less than 2 years, 2–5 years, 6–10 years, More than 10 years
		Educational background	Accounting, Audit, Finance, Information Technology, Other
		Type of organization	Local government, Central government, Government agency, Public enterprise, Other
Use of Accounting Information in Internal Auditing	Role of accounting information in public sector audits	1 question for use of paper-based vs. AI-generated accounting records	5-point Likert scale: Never to Always
		6 questions for agreement with statements on the importance of accounting information in auditing	5-point Likert scale: Strongly Disagree to Strongly Agree
Perception and Use of AI in Internal Audit	Auditors' familiarity and use of AI tools	1 question for self-reported familiarity with AI in auditing	5-point Likert scale: Not familiar at all to Very familiar
		1 question for frequency of AI tool usage (e.g., data analytics, anomaly detection)	5-point Likert scale: Never to Always
		8 questions for perceptions of AI's impact on audit accuracy, efficiency, fraud detection, and workload	5-point Likert scale: Strongly Disagree to Strongly Agree
		1 question for current AI tools used in auditing	e.g., machine learning, predictive modeling, process automation (Multiple-choice)
Benefits of AI in Internal Audit	Auditors' perceptions of AI's effectiveness in various audit aspects	6 questions for the usefulness of AI in efficiency, fraud detection, and decision-making	5-point Likert scale: Not useful to Very useful
		1 question for the extent to which AI has transformed the approach to accounting information in auditing	5-point Likert scale: Not at all to Completely
Challenges and Concerns with AI in Internal Audit	Barriers to AI adoption	6 questions for concerns regarding data privacy, ethics, complexity, and costs	5-point Likert scale: Strongly Disagree to Strongly Agree
		1 question for the perceived difficulty of integrating AI into audit practices	5-point Likert scale: Not a challenge at all to Extremely challenging

		1 question for key barriers to AI implementation	Multiple-choice: Lack of training, High cost, Privacy concerns, complexity, Resistance to change, Lack of support from management
Future Perspectives on AI in Public Sector Internal Audit	Auditors' expectations regarding AI's future role	6 questions for agreement with statements on AI's growing importance, evolving auditor roles, and the necessity for continuous learning	5-point Likert scale: Strongly Disagree to Strongly Agree
		Overall optimism about AI's impact on public sector auditing	Open question

(Source: Authors' research)

### 3.3. Data collection

The data for this study was collected through an online survey (Appendix 4) distributed to internal auditors in the public sector of North Macedonia. The Ministry of Finance provides a publicly accessible list of internal auditors, including their names, institutional affiliations, and contact details (email addresses) - [LINK](#) (internal auditors at the central level) and [LINK](#) (internal auditors at the local level). This list was last updated in December 2024, ensuring the accuracy and completeness of the sampling frame.

The study population initially consisted of 189 internal auditors employed at both central and local government levels. However, due to the unavailability of email addresses for some individuals, the survey was electronically distributed to 173 internal auditors for whom valid contact information was available. Following the email distribution, 18 messages were returned as undeliverable due to inactive or non-existent email addresses. Consequently, the final sample size for the study comprised 155 internal auditors.

Participation in the survey was entirely voluntary and anonymous, ensuring adherence to ethical research standards and maintaining data confidentiality. A total of 48 responses were received, yielding a response rate of 30.97%. This response rate provides valuable insights into internal auditors' perspectives on the utilization of accounting information and the integration of AI in public sector internal auditing. This rate aligns with findings from Nulty (2008), who suggests that response rates for online surveys can be lower than those for paper-based surveys yet still yield meaningful data for analysis. Furthermore, a study by Fincham (2008) indicates that response rates between 30% and 40% are considered acceptable in social science research, providing sufficient data for reliable statistical analysis.

### 3.4. Data analysis

The collected data is analyzed using SPSS software, employing a combination of descriptive and inferential statistical techniques:

- *Descriptive statistics* were used to summarize and present key characteristics of the dataset, providing insights into central tendencies, variability, and distribution patterns of the variables.
- *ANOVA (Appendix 1)* was applied to determine whether there were statistically significant differences between groups based on institutional affiliation (local government vs. central government agencies), making it a suitable method for comparing means across these categories and identifying potential variations in responses.

- Spearman's rank correlation coefficient (Appendix 2) was used to assess the strength and direction of the relationship between familiarity with AI auditing and the frequency of AI use, as it is a non-parametric measure suitable for examining associations between ordinal or non-normally distributed variables. This method allows for identifying monotonic relationships without assuming linearity or homoscedasticity in the data.
- Descriptive analysis of multiple-choice responses.

#### 4. RESULTS AND DISCUSSION

Before any further analysis of the results, a reliability analysis was carried out. The reliability of the survey instrument was assessed using Cronbach's alpha, which yielded a value of 0.719 for the 36 items. According to established guidelines (Taber, 2018; Tavakol and Dennick, 2011), a Cronbach's alpha value of 0.70 or higher is considered acceptable, indicating good internal consistency among the items. Therefore, the result suggests that the items in the survey demonstrate strong reliability in measuring the intended constructs, providing confidence that the responses are consistent and the instrument is effective in capturing the relevant dimensions of the study.

The demographic characteristics of the survey respondents are presented in the following table.

Table 2: Demographic characteristics of the respondents

Feature	Description	Number	Percent
Years of Experience in Internal Auditing	0-2 years	4	8.3%
	3-5 years	11	22.9%
	6-10 years	7	14.6%
	more than 10 years	26	54.2%
	<b>Total</b>	<b>48</b>	<b>100%</b>
Educational Background	Accounting	12	25%
	Auditing	6	12.5%
	Law	4	8.3%
	Finance	20	41.7%
	Other	6	12.5%
	<b>Total</b>	<b>48</b>	<b>100.00%</b>
Institution Type	Local Government	14	29.2%
	Central Government	16	33.3%
	Government Agency	9	18.8%
	State or local owned enterprise	9	18.8%
	<b>Total</b>	<b>48</b>	<b>100%</b>

(Source: Author's calculation)

The demographic characteristics of the respondents indicate that the majority have more than 10 years of experience in internal auditing (54.2%), followed by those with 6-10 years (14.6%), 3-5 years (22.9%), and 0-2 years (8.3%), suggesting that most participants are seasoned professionals. In terms of educational background, Finance (41.7%) and Accounting (25.0%) are the most common fields, while others come from Auditing (12.5%), Law (8.3%), and other disciplines (12.5%), reflecting a diverse academic profile with a strong financial focus. Regarding institutional affiliation, respondents primarily work in Central Government (33.3%) and Local Government (29.2%), with additional representation from Government Agencies (18.8%) and State or Locally Owned Enterprises (18.8%), ensuring a broad perspective on auditing practices across different public sector entities.

#### 4.1. Descriptive statistics

Below in Table 3 are the results of the descriptive statistics of the questions asked on a Likert scale and Yes/No questions.

Table 3: Descriptive statistics

	N	Mean	Std. Deviation
<b>Section 2: Use of Accounting Information in Auditing</b>			
Frequency Accounting Info Use	48	3.625	1.2653
Accounting Essential Audit	48	3.104	1.1893
Accounting Improves Transparency	48	4.229	.6270
Audit Reports Rely Accounting	48	4.271	.7363
Training Accounting Data	48	3.958	.9444
Accounting Data Accuracy	48	3.646	.9563
Access Accounting Info	48	3.979	.9563
<b>Section 3: Perception and Use of AI in Auditing</b>			
Familiarity AI Auditing	48	2.771	1.0766
Frequency AI Use	48	2.500	1.0106
AI Improves Accuracy	48	1.792	.8742
AI Detects Fraud	48	3.354	.6681
AI Increases Efficiency	48	3.396	.6760
AI Reduces Workload	48	3.563	.7118
Confidence AI Use	48	3.521	.9891
Org Provides AI Resources	48	3.250	1.0000
Training AI Audit	48	1.833	1.0383
<b>Section 4: Benefits of AI in Internal Auditing</b>			
AI Increases Efficiency	48	1.604	.8440
AI Improves Data Accuracy	48	3.688	.7192
AI Enhances Fraud Detection	48	3.542	.6510
AI Supports Decision Making	48	3.583	.6790
AI Reduces Human Error	48	3.563	.7693
AI Assists Data Analysis	48	3.625	.7033
AI Transforms Accounting Info Use	48	3.833	.7244
<b>Section 5: Challenges and Concerns with AI in Auditing</b>			
AI Data Privacy Concern	48	2.438	1.0897
AI Ethical Concern	48	3.500	.9453
AI Lack Guidelines	48	3.417	.7945
AI Tool Complexity	48	4.396	.8184
AI Cost Barrier	48	3.667	.9528
AI Reduces Human Auditors	48	3.458	1.0306
AI Integration Challenge	48	2.875	1.2484
<b>Section 6: Future Perspectives on AI in Public Sector Auditing</b>			
AI Essential Next 5 Years	48	3.646	1.1011
AI Benefits Outweigh Drawbacks	48	3.271	1.0051
AI Creates New Roles	48	3.417	.7672
AI Requires Continuous Learning	48	3.479	.8503
Optimism AI Public Auditing	48	4.083	.6469
AI Impact Public Auditing	48	3.625	.9368

(Source: Author's calculations)

### ***Section 2: Use of Accounting Information in Auditing***

The findings indicate that internal auditors frequently use accounting information in their auditing activities, with a mean score of 3.625. This suggests that accounting data is an integral part of the auditing process, though some variation in frequency exists. Additionally, the perception of accounting information as essential for auditing recorded a mean score of (3.104), indicating that while many auditors recognize its importance, some may not view it as indispensable.

The role of accounting information in improving transparency was strongly affirmed, with a high mean score of 4.229. Similarly, the reliance of audit reports on accounting information was emphasized, with a mean of (4.271), demonstrating the fundamental role of financial data in reporting processes. In terms of training, auditors reported a moderate level of training in the use of accounting data, with a mean score of 3.958.

When assessing the accuracy of accounting data, the responses yielded a mean score of 3.646, indicating a general trust in data reliability but also pointing to possible concerns about inconsistencies or errors. Moreover, access to accounting information was reported with a mean score of 3.979, reflecting relatively good availability of financial records within public sector auditing.

### ***Section 3: Perception and Use of AI in Auditing***

The level of familiarity with AI in auditing was relatively low, with a mean score of 2.771, indicating that many auditors are still in the early stages of understanding AI's applications. Furthermore, the frequency of AI use in auditing was reported at 2.500, reinforcing the notion that AI adoption remains limited in practice.

Auditors expressed skepticism regarding AI's role in improving accuracy, with a mean score of 1.792. In contrast, AI's potential to detect fraud was acknowledged, with a mean of (3.354). The perceived ability of AI to increase efficiency received a mean score of 3.396, while its potential to reduce workload scored slightly higher at 3.563.

Confidence in using AI tools was moderate, with a mean score of 3.521, suggesting that while some auditors feel comfortable with AI applications, others may still lack confidence. However, organizational provision of AI-related resources was reported at (3.250), indicating that while some institutions provide support, there is still room for improvement. Additionally, the availability of training on AI in auditing received a low mean score of 1.833, highlighting the need for more education and skill-building initiatives.

### ***Section 4: Benefits of AI in Internal Auditing***

The perception of AI as an efficiency-enhancing tool was low, with a mean score of 1.604, suggesting that many auditors are not yet convinced of its operational benefits. However, AI was seen as improving data accuracy, with a mean score of (3.688), and as enhancing fraud detection, with a mean of (3.542).

AI's role in supporting decision-making received a mean score of 3.583, while its potential to reduce human error was acknowledged with a mean of 3.563. The ability of AI to assist with data analysis was also rated positively, with a mean score of 3.625. Additionally, AI's transformative impact on the use of accounting information was rated relatively high, with a mean of 3.833, indicating that many auditors recognize its potential to redefine traditional processes.

### ***Section 5: Challenges and Concerns with AI in Auditing***

Concerns about AI's impact on data privacy were notable, with a mean score of 2.438, reflecting apprehensions about security risks. Ethical concerns surrounding AI usage were moderately high, with a mean of 3.500, indicating awareness of potential biases and ethical dilemmas.

A lack of clear guidelines for AI adoption was identified as a significant challenge, with a mean score of 3.417. Additionally, the complexity of AI tools emerged as a major barrier, receiving the highest mean score in this section at 4.396, suggesting that many auditors find AI systems difficult to implement and use.

The cost of AI adoption was also seen as a barrier, with a mean score of 3.667, reinforcing the notion that financial constraints hinder widespread AI integration. Concerns about AI replacing human auditors were moderately high, with a mean score of 3.458, suggesting that some auditors fear job displacement. Lastly, challenges related to AI integration received a mean score of 2.875, indicating that while some progress has been made, integrating AI into existing audit processes remains a hurdle.

### ***Section 6: Future Perspectives on AI in Public Sector Auditing***

Internal auditors generally perceived AI as essential in the next five years, with a mean score of 3.646, signaling an expectation of increased reliance on AI. However, the belief that AI's benefits outweigh its drawbacks received a moderate mean score of 3.271, indicating some lingering skepticism.

The potential of AI to create new roles within auditing was recognized, with a mean score of 3.417, highlighting expectations of evolving job functions. The necessity for continuous learning in AI auditing received a mean score of 3.479, emphasizing the need for ongoing education and skill development. Optimism regarding AI's role in public sector auditing was relatively high, with a mean score of 4.083, reflecting positive attitudes toward future advancements. Lastly, the perceived impact of AI on public auditing was recorded at 3.625, demonstrating a general consensus that AI will play a transformative role in the field.

## **4.2. ANOVA**

The ANOVA analysis (Appendix 1) reveals several statistically significant differences in responses based on the type of institution where internal auditors are employed.

In the domain of accounting information use, significant differences were observed for the perception that accounting improves transparency ( $p=.004$ ), the necessity of training in accounting data ( $p = .016$ ), and the perceived accuracy of accounting data ( $p=.004$ ). These findings suggest that auditors from different institutional backgrounds may have varying experiences and levels of confidence in the reliability and role of accounting information in auditing.

Regarding AI-related training and resources, a statistically significant difference was found in access to AI training ( $p = .002$ ), indicating that some institutions provide more structured AI-related training opportunities than others. Additionally, perceptions of AI's ethical concerns ( $p = .028$ ) and the lack of clear guidelines for AI use ( $p = .011$ ) were significantly different among

auditors from different institutions. These results highlight the potential discrepancies in institutional approaches to AI governance and ethical considerations in auditing.

Furthermore, the belief that AI requires continuous learning ( $p = .003$ ) also varied significantly among respondents, implying that institutional differences might shape attitudes toward AI adoption and skill development in auditing.

### 4.3. Spearman's rank correlation coefficient

Spearman's rank correlation coefficient between familiarity with AI auditing and the frequency of AI use revealed a significant positive correlation ( $r=.307, p=.034$ ). This suggests that individuals who are more familiar with AI in auditing tend to use it more frequently. Additionally, significant correlations were observed with other variables. For instance, the frequency of AI use showed a strong positive correlation with the perception that AI improves accuracy ( $r=.597, p<.001$ ) and with the provision of AI-related resources by organizations ( $r=.457, p=.001$ ). Conversely, familiarity with AI auditing was negatively correlated with the perceived complexity of AI tools ( $r=-.411, p=.004$ ), indicating that those who find AI tools more complex tend to be less familiar with AI auditing. These findings highlight the importance of training and organizational support in fostering both AI familiarity and its adoption in auditing practices.

### 4.4. Descriptive analysis of multiple-choice questions

Finally, Table 4 presents the results of the descriptive analysis of the two multiple-choice questions.

Table 4: Descriptive analysis of multiple-choice questions

Which AI tools do you currently use in auditing?	Responses	Percent
<input type="checkbox"/> ChatGPT	2	4.17%
<input type="checkbox"/> Data Analytics	14	29.17%
<input type="checkbox"/> Anomaly detection	3	6.25%
<input type="checkbox"/> None	32	66.67%
What are the main barriers to AI implementation in your work?	Responses	Percent
<input type="checkbox"/> Lack of training	29	60.42%
<input type="checkbox"/> High cost	16	33.33%
<input type="checkbox"/> Data privacy concerns	14	29.17%
<input type="checkbox"/> Complexity of AI tools	11	22.92%
<input type="checkbox"/> Resistance to change	6	12.50%
<input type="checkbox"/> Lack of support from management	17	35.42%

(Source: Author's calculations)

The responses indicate that AI adoption in auditing remains limited, with the majority of respondents (66.7%) reporting that they do not use any AI tools. Among those who do, data analytics tools (29.2%) are the most commonly utilized, followed by anomaly detection (6.3%) and ChatGPT (4.2%), suggesting that AI applications are primarily used for data-driven insights rather than language-based support.

Regarding the barriers to AI implementation, the lack of training (60.4%) is the most significant challenge, highlighting the need for skill development in AI technologies. Other major obstacles include lack of support from management (35.4%), high costs (33.3%), and data privacy concerns (29.2%), which reflect both organizational and financial constraints. Additionally, the complexity of AI tools (22.9%) and resistance to change (12.5%) further hinder adoption, indicating that both technical and cultural factors influence AI integration in auditing practices.

## **5. CONCLUSION**

This study provides valuable insights into the current use of accounting information and AI in internal auditing within the public sector. The findings confirm that accounting information remains a cornerstone of auditing practice, especially in enhancing transparency and informing audit reports. However, perceptions about its indispensability vary, and there is a moderate level of training and trust in the reliability of accounting data.

In contrast, the use and understanding of AI in auditing are still in their infancy. Although auditors recognize AI's potential in areas such as fraud detection, decision-making, and reducing human error, actual adoption remains limited. The low frequency of AI use and training, combined with concerns about complexity, cost, and ethical considerations, underscores significant barriers to integration. Importantly, auditors who are more familiar with AI tend to use it more frequently, suggesting that targeted training and institutional support could accelerate adoption.

The analysis also reveals notable differences between institutions regarding access to training, ethical concerns, and perceptions of AI's impact - highlighting the need for standardized guidelines and broader capacity-building efforts. Most auditors believe that AI will become essential in the near future, yet their optimism is tempered by the challenges they currently face.

Based on the current state identified in this study, i.e., low familiarity with AI tools, infrequent training, and barriers related to complexity, cost, and uneven institutional support, internal audit units should: (i) start with low-risk pilots using currently familiar analytics (e.g., exception and anomaly detection) on high-volume, high-risk processes; (ii) adopt a basic AI governance pack (usage policy, privacy checklist, bias/ethics notes, documentation templates) and designate an "AI champion" per unit; (iii) deliver short, role-based training linked to live audit tasks and pair it with supervised practice; (iv) improve data readiness through standardized data extracts and secure access protocols; (v) pool procurement and share models/dashboards across peer institutions to reduce costs and duplication; and (vi) monitor adoption with simple KPIs (share of audits using analytics, exception rates, rework/clarification cycles) and review lessons learned.

The results are derived from a cross-sectional, self-reported survey of public-sector internal auditors in North Macedonia with a modest sample size, which may limit generalizability. The constructs emphasize perceptions rather than observed behavior, and the design lacks longitudinal tracking of AI adoption over time. Future work should broaden the sample, integrate qualitative evidence, and follow AI-enabled audit workflows longitudinally to capture capability maturation and impact.

Ultimately, for AI to fulfill its transformative promise in public sector auditing, investments in education, clearer policies, and supportive organizational cultures will be essential. Bridging

the gap between current practices and future potential will require a strategic, inclusive, and well-resourced approach to digital innovation in audit processes.

## REFERENCES

- Aldemir, C. and Uysal, T. U. (2024), "AI competencies for internal auditors in the public sector", *The EDP Audit, Control, and Security Newsletter*, Vol. 69, No. 1, pp. 3-21.
- Allami, F. A. J. (2022), "The use of external auditor to data mining as an artificial intelligence technology to examine the internal control systems in an electronic business environment", *Czech Journal of Multidisciplinary Innovations*, Vol. 9, pp. 1-13.
- Babina, T., Fedyk, A., He, A. and Hodson, J. (2020), "Artificial intelligence, firm growth, and product innovation", *Review of Financial Studies*, Vol. 33, No. 8, pp. 3309-3347.
- Bughin, J., Seong, J., Manyika, J., Chui, M. and Joshi, R. (2017), "AI adoption advances, but foundational barriers remain", McKinsey Global Institute, available at: <https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain> (accessed 3 August, 2025)
- Burton, E. J. and Rezaee, Z. (1994), "University-based lifelong continued professional development (CPD) for internal auditors", *Managerial Auditing Journal*, Vol. 9, No. 2, pp. 32-36.
- Chen, H. and Vasarhelyi, M. A. (2019), "The impact of AI on auditing", *Journal of Emerging Technologies in Accounting*, Vol. 16, No. 1, pp. 71-92.
- De Sousa, W. G., de Melo, E. R. P., Bermejo, P. H. D. S., Farias, R. A. S. and Gomes, A. O. (2019), "How and where is artificial intelligence in the public sector going?", *Government Information Quarterly*, Vol. 36, No. 4.
- Fedyk, A., Hodson, J., Khimich, N. and Fedyk, T. (2022), "Is artificial intelligence improving the audit process?", *Review of Accounting Studies*, Vol. 27, No. 3, pp. 938-985.
- Fincham, J. E. (2008), "Response rates and responsiveness for surveys, standards, and the journal", *American Journal of Pharmaceutical Education*, Vol. 72, No. 2, 43.
- Frey, C. B. and Osborne, M. A. (2017), "The future of employment: How susceptible are jobs to computerization?", *Technological Forecasting and Social Change*, Vol. 114, pp. 254-280.
- Institute of Internal Auditors (IIA). (2017a), "Global perspectives and insights: Artificial intelligence – Considerations for the profession of internal auditing", available at: <https://www.theiia.org/globalassets/site/content/articles/global-perspectives-and-insights/2023/q3-gpi-ai.pdf> (accessed 22 July, 2025).
- Institute of Internal Auditors (IIA). (2017b), "The IIA's artificial intelligence auditing framework: Practical applications", available at: <https://www.theiia.org/en/resources/internal-audit-competency-framework/> (accessed 20 July, 2025).
- Institute of Internal Auditors (IIA). (2021), "Assessing internal audit competency: Minding the gaps to maximize insights", available at: <https://www.theiia.org/en/content/research/foundation/2021/assessing-internal-audit-competency-minding-the-gaps-to-maximize-insights/> (accessed 20 July, 2025).
- Issa, H., Sun, T. and Vasarhelyi, M. A. (2016), "Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation", *Journal of Emerging Technologies in Accounting*, Vol. 13, No.2, pp. 1-20.
- Minkkinen, M., Laine, J. and Mäntymäki, M. (2022), "Continuous auditing of artificial intelligence: A conceptualization and assessment of tools and frameworks", *Digital Society*, Vol. 1, No. 3, 21.

- Nulty, D. D. (2008), "The adequacy of response rates to online and paper surveys: What can be done?", *Assessment & Evaluation in Higher Education*, Vol. 33, No. 3, pp. 301-314.
- OECD. (2018), "The future of education and skills: Education 2030", available at: [https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf) (accessed 25 July, 2025)
- Raji, I. D. and Buolamwini, J. (2022), "Actionable auditing revisited: Investigating the impact of publicly naming biased performance results of commercial AI products", *Communications of the ACM*, Vol. 66, No. 1, pp. 101-108.
- Rehman, A. and Hashim, F. (2022), "Can internal audit function impact artificial intelligence?" in *The AIP Conference Proceedings of the 5th Innovation and Analytics Conference & Exhibition (IACE) in Kedah, Malaysia, 2021*.
- Rock, D. (2020), "AI and the future of work", *Harvard Business Review*, Vol. 98, No. 3, pp. 125-134.
- Setyaningrum, D., Mita, A. F. and Rosdini, D. (2022), "Internal auditor competency gap: Perception of students, academics, and practitioners", *Jurnal Reviu Akuntansi dan Keuangan*, Vol. 12, No. 3, pp. 696-710.
- Taber, K. S. (2018), "The use of Cronbach's alpha when developing and reporting research instruments in science education", *Research in Science Education*, Vol. 48, No. 6, pp. 1273-1296.
- Tavakol, M. and Dennick, R. (2011), "Making sense of Cronbach's alpha", *International Journal of Medical Education*, Vol. 2, pp. 53-55.
- Wang, Y. and Cuthbertson, R. (2022), "AI auditing in practice", *Journal of Accounting and Public Policy*, Vol. 41, No. 2, pp. 276-298.
- Wirtz, B. W., Weyerer, J. C. and Geyer, C. (2019), "Artificial intelligence and the public sector - Applications and challenges", *International Journal of Public Administration*, Vol. 42, No. 7, pp. 596-615.

**Appendix 1: ANOVA**

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Frequency_Accounting_Info_Use	Between Groups	11.678	3	3.893	2.694	.057
	Within Groups	63.572	44	1.445		
	Total	75.250	47			
Accounting_Essential_Audit	Between Groups	7.891	3	2.630	1.975	.132
	Within Groups	58.588	44	1.332		
	Total	66.479	47			
Accounting_Improves_Transparency	Between Groups	4.724	3	1.575	5.037	.004
	Within Groups	13.755	44	.313		
	Total	18.479	47			
Audit_Reports_Rely_Accounting	Between Groups	3.597	3	1.199	2.411	.080
	Within Groups	21.882	44	.497		
	Total	25.479	47			
Training_Accounting_Data	Between Groups	8.662	3	2.887	3.820	.016
	Within Groups	33.255	44	.756		
	Total	41.917	47			
Accounting_Data_Accuracy	Between Groups	11.233	3	3.744	5.190	.004
	Within Groups	31.746	44	.722		
	Total	42.979	47			
Access_Accounting_Info	Between Groups	2.566	3	.855	.931	.434
	Within Groups	40.413	44	.918		
	Total	42.979	47			
Familiarity_AI_Auditing	Between Groups	6.761	3	2.254	2.078	.117
	Within Groups	47.718	44	1.085		
	Total	54.479	47			
Frequency_AI_Use	Between Groups	4.460	3	1.487	1.502	.227
	Within Groups	43.540	44	.990		
	Total	48.000	47			
AI_Improves_Accuracy	Between Groups	1.238	3	.413	.524	.668
	Within Groups	34.679	44	.788		
	Total	35.917	47			
AI_Detects_Fraud	Between Groups	3.058	3	1.019	2.502	.072
	Within Groups	17.922	44	.407		
	Total	20.979	47			
AI_Increases_Efficiency	Between Groups	2.007	3	.669	1.512	.225
	Within Groups	19.472	44	.443		
	Total	21.479	47			
AI_Reduces_Workload	Between Groups	3.376	3	1.125	2.423	.078
	Within Groups	20.437	44	.464		
	Total	23.813	47			
Confidence_AI_Use	Between Groups	5.327	3	1.776	1.922	.140
	Within Groups	40.652	44	.924		

	Total	45.979	47			
Org_Provides_AI_Resources	Between Groups	4.421	3	1.474	1.523	.222
	Within Groups	42.579	44	.968		
	Total	47.000	47			
Training_AI_Audit	Between Groups	14.365	3	4.788	5.804	.002
	Within Groups	36.302	44	.825		
	Total	50.667	47			
AI_Increases_Efficiency	Between Groups	3.129	3	1.043	1.512	.225
	Within Groups	30.350	44	.690		
	Total	33.479	47			
AI_Improves_Data_Accuracy	Between Groups	1.066	3	.355	.673	.573
	Within Groups	23.246	44	.528		
	Total	24.313	47			
AI_Enhances_Fraud_Detection	Between Groups	2.480	3	.827	2.086	.116
	Within Groups	17.437	44	.396		
	Total	19.917	47			
AI_Supports_Decision_Making	Between Groups	2.480	3	.827	1.896	.144
	Within Groups	19.187	44	.436		
	Total	21.667	47			
AI_Reduces_Human_Error	Between Groups	2.955	3	.985	1.744	.172
	Within Groups	24.857	44	.565		
	Total	27.813	47			
AI_Assists_Data_Analysis	Between Groups	1.750	3	.583	1.194	.323
	Within Groups	21.500	44	.489		
	Total	23.250	47			
AI_Transforms_Accounting_Info_Use	Between Groups	3.301	3	1.100	2.266	.094
	Within Groups	21.366	44	.486		
	Total	24.667	47			
AI_Data_Privacy_Concern	Between Groups	2.407	3	.802	.661	.580
	Within Groups	53.406	44	1.214		
	Total	55.813	47			
AI_Ethical_Concern	Between Groups	7.745	3	2.582	3.316	.028
	Within Groups	34.255	44	.779		
	Total	42.000	47			
AI_Lack_Guidelines	Between Groups	6.570	3	2.190	4.172	.011
	Within Groups	23.096	44	.525		
	Total	29.667	47			
AI_Tool_Complexity	Between Groups	2.301	3	.767	1.156	.337
	Within Groups	29.179	44	.663		
	Total	31.479	47			
AI_Cost_Barrier	Between Groups	3.801	3	1.267	1.434	.246
	Within Groups	38.866	44	.883		
	Total	42.667	47			
AI_Reduces_Human_Auditors	Between Groups	4.794	3	1.598	1.558	.213
	Within Groups	45.123	44	1.026		
	Total	49.917	47			

AI_Integration_Challenge	Between Groups	6.063	3	2.021	1.324	.279
	Within Groups	67.187	44	1.527		
	Total	73.250	47			
AI_Essential_Next_5_Years	Between Groups	7.090	3	2.363	2.084	.116
	Within Groups	49.889	44	1.134		
	Total	56.979	47			
AI_Benefits_Outweigh_Drawbacks	Between Groups	4.153	3	1.384	1.406	.254
	Within Groups	43.326	44	.985		
	Total	47.479	47			
AI_Creates_New_Roles	Between Groups	2.837	3	.946	1.676	.186
	Within Groups	24.829	44	.564		
	Total	27.667	47			
AI_Requires_Continuous_Learning	Between Groups	9.224	3	3.075	5.465	.003
	Within Groups	24.755	44	.563		
	Total	33.979	47			
Optimism_AI_Public_Auditing	Between Groups	1.877	3	.626	1.547	.216
	Within Groups	17.790	44	.404		
	Total	19.667	47			
AI_Impact_Public_Auditing	Between Groups	6.571	3	2.190	2.779	.052
	Within Groups	34.679	44	.788		
	Total	41.250	47			

**Appendix 2: Spearman's rank correlation coefficient**

		<b>Familiarity_AI_Auditing</b>	<b>Frequency_AI_Use</b>
<b>Frequency_Accounting_Info_Use</b>	Correlation Coefficient	-.150	-.422**
	Sig. (2-tailed)	.310	.003
	N	48	48
<b>Accounting_Essential_Audit</b>	Correlation Coefficient	.416**	.388**
	Sig. (2-tailed)	.003	.006
	N	48	48
<b>Accounting_Improves_Transparency</b>	Correlation Coefficient	.169	-.001
	Sig. (2-tailed)	.252	.997
	N	48	48
<b>Audit_Reports_Rely_Accounting</b>	Correlation Coefficient	.246	-.055
	Sig. (2-tailed)	.092	.709
	N	48	48
<b>Training_Accounting_Data</b>	Correlation Coefficient	.472**	.320*
	Sig. (2-tailed)	.001	.027
	N	48	48
<b>Accounting_Data_Accuracy</b>	Correlation Coefficient	.206	.061
	Sig. (2-tailed)	.160	.680
	N	48	48
<b>Access_Accounting_Info</b>	Correlation Coefficient	.471**	.154
	Sig. (2-tailed)	.001	.296
	N	48	48
<b>Familiarity_AI_Auditing</b>	Correlation Coefficient	1.000	.307*
	Sig. (2-tailed)		.034
	N	48	48
<b>Frequency_AI_Use</b>	Correlation Coefficient	.307*	1.000
	Sig. (2-tailed)	.034	
	N	48	48
<b>AI_Improves_Accuracy</b>	Correlation Coefficient	.273	.597**
	Sig. (2-tailed)	.061	.000
	N	48	48
<b>AI_Detects_Fraud</b>	Correlation Coefficient	-.185	.016
	Sig. (2-tailed)	.208	.912
	N	48	48
<b>AI_Increases_Efficiency</b>	Correlation Coefficient	.007	.169
	Sig. (2-tailed)	.961	.252
	N	48	48
<b>AI_Reduces_Workload</b>	Correlation Coefficient	.095	.038
	Sig. (2-tailed)	.522	.797
	N	48	48
<b>Confidence_AI_Use</b>	Correlation Coefficient	-.029	-.093
	Sig. (2-tailed)	.847	.528
	N	48	48
<b>Org_Provides_AI_Resources</b>	Correlation Coefficient	.002	.457**
	Sig. (2-tailed)	.991	.001
	N	48	48

<b>Training_AI_Audit</b>	Correlation Coefficient	-.078	.017
	Sig. (2-tailed)	.599	.906
	N	48	48
<b>AI_Increases_Efficiency</b>	Correlation Coefficient	.260	.278
	Sig. (2-tailed)	.074	.055
	N	48	48
AI_Improves_Data_Accuracy	Correlation Coefficient	.047	-.222
	Sig. (2-tailed)	.752	.130
	N	48	48
<b>AI_Enhances_Fraud_Detection</b>	Correlation Coefficient	-.291*	-.264
	Sig. (2-tailed)	.045	.070
	N	48	48
<b>AI_Supports_Decision_Making</b>	Correlation Coefficient	-.119	-.106
	Sig. (2-tailed)	.422	.473
	N	48	48
<b>AI_Reduces_Human_Error</b>	Correlation Coefficient	-.141	-.043
	Sig. (2-tailed)	.339	.771
	N	48	48
<b>AI_Assists_Data_Analysis</b>	Correlation Coefficient	-.062	-.233
	Sig. (2-tailed)	.678	.112
	N	48	48
<b>AI_Transforms_Accounting_Info_Use</b>	Correlation Coefficient	-.025	.146
	Sig. (2-tailed)	.868	.322
	N	48	48
<b>AI_Data_Privacy_Concern</b>	Correlation Coefficient	.084	.296*
	Sig. (2-tailed)	.570	.041
	N	48	48
<b>AI_Ethical_Concern</b>	Correlation Coefficient	-.276	-.040
	Sig. (2-tailed)	.058	.785
	N	48	48
<b>AI_Lack_Guidelines</b>	Correlation Coefficient	-.217	.066
	Sig. (2-tailed)	.138	.657
	N	48	48
<b>AI_Tool_Complexity</b>	Correlation Coefficient	-.411**	-.164
	Sig. (2-tailed)	.004	.266
	N	48	48
<b>AI_Cost_Barrier</b>	Correlation Coefficient	-.188	-.137
	Sig. (2-tailed)	.200	.352
	N	48	48
<b>AI_Reduces_Human_Auditors</b>	Correlation Coefficient	.039	-.142
	Sig. (2-tailed)	.792	.336
	N	48	48
<b>AI_Integration_Challenge</b>	Correlation Coefficient	-.162	-.269
	Sig. (2-tailed)	.271	.064
	N	48	48
<b>AI_Essential_Next_5_Years</b>	Correlation Coefficient	.077	-.114
	Sig. (2-tailed)	.603	.440
	N	48	48
<b>AI_Benefits_Outweigh_Drawbacks</b>	Correlation Coefficient	-.095	-.208
	Sig. (2-tailed)	.522	.156

	N	48	48
<b>AI_Creates_New_Roles</b>	Correlation Coefficient	.029	-.157
	Sig. (2-tailed)	.843	.286
	N	48	48
<b>AI_Requires_Continuous_Learning</b>	Correlation Coefficient	.008	.067
	Sig. (2-tailed)	.956	.649
	N	48	48
<b>Optimism_AI_Public_Auditing</b>	Correlation Coefficient	-.188	-.068
	Sig. (2-tailed)	.201	.645
	N	48	48
<b>AI_Impact_Public_Auditing</b>	Correlation Coefficient	-.181	.083
	Sig. (2-tailed)	.219	.574
	N	48	48

**Appendix 3: Case summaries**

		Frequency_Accounting_Info_Use	Accounting_Essential_Audit (Likert 1-5)	Accounting_Improves_Transparency (Likert 1-5)	Audit_Reports_Rely_Accounting (Likert 1-5)	Training_Accounting_Data (Likert 1-5)	Accounting_Data_Accuracy (Likert 1-5)	Access_Accounting_Info (Likert 1-5)	Familiarity_AI_Auditing	Frequency_AI_Use	AI_Improves_Accuracy (Likert 1-5)	AI_Detects_Fraud (Likert 1-5)	AI_Increases_Efficiency (Likert 1-5)	
	Total	N	14	14	14	14	14	14	14	14	14	14	14	
		Mean	4.214	3.071	4.571	4.500	4.071	3.714	4.286	3.286	2.571	1.929	3.429	3.500
		Std. Deviation	1.2514	1.2067	.5136	.7596	.7300	.9945	.8254	.9945	1.2839	.6157	.5136	.5189
	Total	N	16	16	16	16	16	16	16	16	16	16	16	
		Mean	3.688	3.313	4.313	4.438	4.313	4.000	4.000	2.625	2.750	1.625	3.063	3.125
		Std. Deviation	1.2500	1.0145	.4787	.5123	.4787	.8165	.8165	1.0247	.7746	.7188	.6801	.7188
	Total	N	9	9	9	9	9	9	9	9	9	9	9	
		Mean	2.778	3.556	3.667	4.111	4.000	3.889	3.778	2.778	2.556	2.000	3.333	3.444
		Std. Deviation	1.0929	1.3333	.8660	.7817	1.0000	.7817	.9718	1.2019	1.0138	1.0000	.7071	.5270
	Total	N	9	9	9	9	9	9	9	9	9	9	9	
		Mean	3.444	2.333	4.111	3.778	3.111	2.667	3.667	2.222	1.889	1.667	3.778	3.667
		Std. Deviation	1.1304	1.1180	.3333	.8333	1.3642	.7071	1.3229	.9718	.7817	1.3229	.6667	.8660
	Total	N	48	48	48	48	48	48	48	48	48	48	48	
		Mean	3.625	3.104	4.229	4.271	3.958	3.646	3.979	2.771	2.500	1.792	3.354	3.396
		Std. Deviation	1.2653	1.1893	.6270	.7363	.9444	.9563	.9563	1.0766	1.0106	.8742	.6681	.6760

		AI_Reduces_Workload (Likert 1-5)	Confidence_AI_Use (Likert 1-5)	Org_Provide_AI_Resources (Likert 1-5)	Training_AI_Audit (Likert 1-5)	AI_Increases_Efficiency (Likert 1-5)	AI_Improves_Data_Accuracy (Likert 1-5)	AI_Enhances_Fraud_Detection (Likert 1-5)	AI_Supports_Decision_Making (Likert 1-5)	AI_Reduces_Human_Error (Likert 1-5)	AI_Assists_Data_Analysis (Likert 1-5)	AI_Transforms_Accounting_Info_Use (Likert 1-5)	AI_Data_Privacy_Concern (Likert 1-5)
	Total	N	14	14	14	14	14	14	14	14	14	14	14
		Mean	3.643	3.357	2.786	1.286	1.286	3.786	3.357	3.357	3.286	3.500	4.071

	Total	Std. Deviation	.4972	.4972	1.1217	.4688	.4688	.4258	.4972	.6333	.7263	.5189	.7300	.8254	
		N	16	16	16	16	16	16	16	16	16	16	16	16	16
		Mean	3.250	3.688	3.500	1.750	1.563	3.500	3.500	3.625	3.500	3.500	3.813	2.688	
	Total	Std. Deviation	.9309	1.1383	.6325	.8563	.7274	.8165	.8165	.8062	.8165	.8165	.8342	.8732	
		N	9	9	9	9	9	9	9	9	9	9	9	9	
		Mean	3.556	3.000	3.444	2.889	2.000	3.667	3.444	3.444	3.667	3.667	3.333	2.556	
	Total	Std. Deviation	.5270	1.4142	1.0138	1.4530	1.3229	.8660	.5270	.5270	.8660	.8660	.5000	1.1304	
		N	9	9	9	9	9	9	9	9	9	9	9	9	
		Mean	4.000	4.000	3.333	1.778	1.778	3.889	4.000	4.000	4.000	4.000	4.000	2.111	
	Total	Std. Deviation	.5000	.5000	1.2247	.8333	.8333	.7817	.5000	.5000	.5000	.5000	.5000	1.6915	
		N	9	9	9	9	9	9	9	9	9	9	9	9	
		Mean	4.000	4.000	3.333	1.778	1.778	3.889	4.000	4.000	4.000	4.000	4.000	2.111	
	Total	N	48	48	48	48	48	48	48	48	48	48	48	48	
		Mean	3.563	3.521	3.250	1.833	1.604	3.688	3.542	3.583	3.563	3.625	3.833	2.438	
		Std. Deviation	.7118	.9891	1.0000	1.0383	.8440	.7192	.6510	.6790	.7693	.7033	.7244	1.0897	

		AI_Ethical_Concern (Likert 1-5)	AI_Lack_Guidelines (Likert 1-5)	AI_Tool_Complexity (Likert 1-5)	AI_Cost_BARRIER (Likert 1-5)	AI_Reduces_Human_Auditors (Likert 1-5)	AI_Integration_Challenge	AI_Essential_Next_5_Years (Likert 1-5)	AI_Benefits_Outweigh_Drawbacks (Likert 1-5)	AI_Creates_New_Roles (Likert 1-5)	AI_Requires_Continuous_Learning (Likert 1-5)	Optimism_AI_Public_Auditing (Likert 1-5)	AI_Impact_Public_Auditing	
	Total	N	14	14	14	14	14	14	14	14	14	14	14	
		Mean	3.071	3.143	4.429	3.929	3.929	2.643	4.000	3.500	3.714	3.929	4.071	3.929
		Std. Deviation	1.0716	.6630	.7559	.8287	.8287	.7449	.6794	1.0190	.6112	.6157	.4746	.8287
	Total	N	16	16	16	16	16	16	16	16	16	16	16	
		Mean	3.813	3.938	4.625	3.438	3.375	3.375	3.250	2.938	3.125	2.938	3.875	3.125
		Std. Deviation	.8342	.9287	.7188	1.0935	1.2042	1.5438	1.2383	.9287	.8851	.9287	.8062	.8851

	T o t a l	N	9	9	9	9	9	9	9	9	9	9	9	
		Mean	4.000	3.222	4.000	3.333	3.222	2.667	4.111	3.111	3.556	3.333	4.444	3.667
	T o t a l	Std. Deviation	1.0000	.4410	1.0000	.7071	.9718	1.3229	1.3642	1.1667	.5270	.5000	.5270	1.0000
		N	9	9	9	9	9	9	9	9	9	9	9	9
	T o t a l	Mean	3.111	3.111	4.333	4.000	3.111	2.556	3.333	3.667	3.333	3.889	4.111	4.000
		Std. Deviation	.3333	.6009	.8660	1.0000	.9280	1.1304	.8660	.8660	.8660	.7817	.6009	.8660
T o t a l		N	48	48	48	48	48	48	48	48	48	48	48	48
		Mean	3.500	3.417	4.396	3.667	3.458	2.875	3.646	3.271	3.417	3.479	4.083	3.625
		Std. Deviation	.9453	.7945	.8184	.9528	1.0306	1.2484	1.1011	1.0051	.7672	.8503	.6469	.9368

## **Appendix 4: Survey: The Role of Accounting Information and AI in Public Sector Internal Auditing**

### **Section 1: General Characteristics of Respondents**

1. Job Title
  - [Open Text]
2. Years of Experience in Internal Auditing
  - Less than 2 years
  - 2–5 years
  - 6–10 years
  - More than 10 years
3. Educational Background
  - Accounting
  - Auditing
  - Finance
  - IT
  - Other (please specify): \_\_\_\_\_
4. Organization Type
  - Local government
  - National government
  - Government agency
  - State-owned enterprise
  - Other (please specify): \_\_\_\_\_

### **Section 2: Use of Accounting Information in Auditing**

5. How frequently do you use accounting information in your internal audit work? (1 = Never, 5 = Always)
6. To what extent do you agree with the following statements? (1 = Strongly Disagree, 5 = Strongly Agree)
  - Accounting information is essential for assessing financial health in public sector audits.
  - Accounting information helps improve accountability and transparency in my audit work.
  - My audit reports rely heavily on accounting information for evidence-based findings.
  - I receive adequate training in interpreting accounting data for audits.
  - Accounting information provided by my organization is accurate and reliable.
  - Access to relevant accounting information is sufficient for my auditing needs.

### **Section 3: Perception and Use of Artificial Intelligence in Auditing**

7. How familiar are you with AI tools in auditing? (1 = Not familiar, 5 = Very familiar)
8. How often do you use AI tools (e.g., data analytics, anomaly detection) in your auditing work? (1 = Never, 5 = Always)
9. To what extent do you agree with the following statements about AI in auditing? (1 = Strongly Disagree, 5 = Strongly Agree)
  - AI tools improve the accuracy of audit findings.
  - AI helps in detecting anomalies and fraud more effectively than traditional methods.
  - AI increases the efficiency of the audit process.
  - AI reduces the workload of internal auditors.
  - I am confident in my ability to use AI tools in audits.
  - My organization provides adequate resources for AI tools in auditing.
  - I receive sufficient training on how to apply AI in my audit work.
10. Which AI tools do you currently use in auditing? (Select all that apply)
  - Data analytics
  - Machine learning
  - Predictive modeling
  - Anomaly detection
  - Process automation

- None
- Other (please specify): \_\_\_\_\_

#### **Section 4: Benefits of AI in Internal Auditing**

11. How beneficial do you find AI in each of the following aspects of auditing? (1 = Not Beneficial, 5 = Highly Beneficial)
- Increasing audit efficiency
  - Improving data accuracy and reliability
  - Enhancing fraud detection
  - Supporting evidence-based decision-making
  - Reducing human error
  - Assisting with complex data analysis
12. To what extent has AI transformed your approach to accounting information in auditing? (1 = Not at All, 5 = Completely)

#### **Section 5: Challenges and Concerns with AI in Auditing**

13. To what extent do you agree with the following statements about challenges with AI in auditing? (1 = Strongly Disagree, 5 = Strongly Agree)
- I am concerned about data privacy and security when using AI tools.
  - The use of AI may compromise ethical standards in auditing.
  - There are insufficient guidelines for AI use in public sector auditing.
  - AI tools are too complex for effective use without extensive training.
  - The cost of AI technology is a barrier for its implementation in auditing.
  - AI might eventually reduce the role of human auditors.
14. How challenging do you find it to integrate AI into your audit practices? (1= Not challenging, 5= Extremely challenging)
15. What are the main barriers to AI implementation in your work? (Select all that apply)
- Lack of training
  - High cost
  - Data privacy concerns
  - Complexity of AI tools
  - Resistance to change
  - Lack of support from management
  - Other (please specify): \_\_\_\_\_

#### **Section 6: Future Perspectives on AI in Public Sector Auditing**

16. To what extent do you agree with the following statements about the future of AI in auditing? (1 = Strongly Disagree, 5 = Strongly Agree)
- AI will become essential in public sector auditing within the next five years.
  - The benefits of AI in auditing outweigh its potential drawbacks.
  - AI will create new roles and responsibilities for internal auditors.
  - AI will require continuous learning and adaptation by auditors.
  - I am optimistic about the role of AI in improving public sector auditing practices.
17. In your opinion, how likely is it that AI will significantly change the public sector auditing profession? (1 = Very unlikely, 5 = Very likely)
18. What additional support or resources would you find helpful for integrating AI into your auditing work?
- [Open Text]

## **DYNAMIC CONNECTEDNESS OF EASTERN EUROPEAN STOCK MARKETS: AN EXTENDED JOINT CONNECTEDNESS APPROACH**

**Gjorgji Gockov**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
gjorgji.gockov@eccf.ukim.edu.mk*

**Goran Hristovski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
goran.hristovski@eccf.ukim.edu.mk*

### **ABSTRACT**

*This paper examines the dynamic return spillovers among ten Eastern European stock markets using an extended joint connectedness approach. We analyze daily log returns from 2010 to 2024 for equity indices of Serbia, Bosnia and Herzegovina, Croatia, Slovenia, North Macedonia, Romania, the Czech Republic, Hungary, Poland, and Ukraine. We employ the Diebold-Yilmaz connectedness framework, augmented with a Time-Varying Parameter VAR (TVP-VAR) model to capture evolving relationships, and incorporate the extended joint spillover methodology to address bias from normalization. The Connectedness Approach R package is used for implementation. Our results show that these markets exhibit a moderate degree of interconnectedness on average - the Total Connectedness Index is around 25%, indicating that roughly one-quarter of forecast variance is due to cross-market shocks. Spillovers are highly time-varying, with pronounced surges during major crises such as the European sovereign debt crisis, the COVID-19 pandemic, and the Russia-Ukraine conflict. Directional spillover analysis reveals that larger markets (e.g., Romania, Poland, Czech Republic) tend to be net transmitters of shocks, whereas smaller frontier markets (e.g., Serbia, Bosnia, North Macedonia) are net recipients. The extended joint connectedness measures largely confirm the traditional spillover estimates while providing a more theoretically grounded aggregate index. These findings shed light on the evolving integration of Eastern European stock markets, offering insights for portfolio diversification and financial stability monitoring.*

**Keywords:** *Dynamic connectedness, Eastern European stock markets, Spillovers, Extended Joint connectedness*

**JEL classification:** *C32, G11, G15*

### **1. INTRODUCTION**

The increasing integration of global financial markets has intensified interest in spillovers and interconnectedness among equity markets. Spillovers refer to the transmission of shocks or volatility from one market to another, which is crucial for understanding contagion risks and potential diversification benefits. This topic is especially important for emerging and less-developed markets such as those in Eastern Europe, where financial systems are smaller and potentially more vulnerable to external shocks. Investors and policymakers need to know whether these markets move in tandem or remain segmented; high connectedness can signal contagion risk, whereas low connectedness might imply scope for diversification. Eastern

European stock markets – including those in the Balkans and Central-Eastern Europe - have historically been less developed and less liquid than their Western counterparts. However, the past decade has seen these markets gradually integrate with global financial systems (e.g., through increased economic linkages and, for some countries, European Union accession). Analyzing their dynamic connectedness can reveal how shocks propagate within the region and whether these frontier markets are becoming more intertwined with each other or still offer insulation from global turmoil.

This study focuses on ten Eastern European equity indices from Serbia, Bosnia and Herzegovina, Croatia, Slovenia, North Macedonia, Romania, the Czech Republic, Hungary, Poland, and Ukraine to assess their return spillovers from 2010 through 2024. By examining a long post-2009 period that encompasses the European sovereign debt crisis, periods of relative stability, the COVID-19 shock in 2020, and recent geopolitical tensions (such as the 2022 Russia-Ukraine conflict), we provide a comprehensive view of regional market interconnectedness. Previous studies have examined spillovers in Eastern European markets using approaches such as DCC-GARCH and extensions of the Diebold–Yilmaz VAR framework (Demiralay & Bayraci, 2015; Škrinjarić, 2020, 2021; Syllignakis & Kouretas, 2011) but their evidence mostly covers earlier periods, offering limited insight into COVID-19 or the Russia–Ukraine conflict. More recent work, such as Škrinjarić (2022), incorporates the COVID-19 period, yet little is known about how connectedness in emerging and less-liquid European markets has evolved beyond the pandemic, particularly under subsequent shocks such as global inflation and the Russia–Ukraine war. Our aim is to determine the level and evolution of return spillovers in this region and to identify which markets act as shock transmitters versus receivers in the network. These insights are valuable for portfolio management (e.g., understanding the limits of regional diversification) and for regulators concerned with financial stability in emerging European markets.

To investigate these questions, we employ the Diebold-Yilmaz spillover index framework, which has become a standard approach for quantifying systemic connectedness in finance (Diebold and Yilmaz, 2009, 2012). We address the limitations of the original framework by using a time-varying parameter VAR model (Antonakakis *et al.*, 2018) instead of a rolling window, allowing the connectedness measures to evolve smoothly over time. Furthermore, we incorporate the extended joint connectedness approach recently developed by (Lastrapes and Wiesen, 2021) and extended by (Balcilar *et al.*, 2021), which provides a more theoretically grounded measure of total spillovers and enables the computation of directional spillovers within that joint framework. Leveraging an extended, recent daily dataset for Eastern Europe, we assess whether smaller frontier stock exchanges are integrating with regional peers and global financial cycles, or whether they remain sufficiently segmented to offer diversification

In summary, our contribution is twofold. First, we provide up-to-date empirical evidence on how ten Eastern European stock markets are interconnected, how those connections change over time, and how global/regional shocks influence the network. Second, we demonstrate the use of an improved connectedness approach - combining the Diebold-Yilmaz methodology, TVP-VAR modeling, and the joint spillover index - which yields more robust insights than traditional methods. The remainder of the paper is structured as follows. Section 2 reviews the relevant literature on connectedness measurement and previous findings. Section 3 outlines the methodology, including the connectedness frameworks and estimation procedure. Section 4 describes the data. Section 5 presents and discusses the empirical results. Section 6 concludes with a summary of findings and implications.

## 2. LITERATURE REVIEW

Understanding the dynamics of cross-market return spillovers and connectedness has become increasingly crucial for investors, policymakers, and financial analysts, particularly in the context of emerging and frontier markets where vulnerabilities to external shocks are often amplified. The transmission of shocks or volatility from one market or asset class to another has been a central theme in international finance over the past two decades, reflecting the growing integration and susceptibility of markets to systemic events.

The foundational framework of (Diebold and Yilmaz, 2009, 2012; Diebold and Yilmaz, 2014) transformed the empirical study of connectedness by introducing a VAR-based methodology grounded in forecast error variance decompositions (FEVD). This approach enabled a systematic quantification of total, directional (to/from), and net spillovers across markets, offering a unified language for describing interdependencies. Despite its influence, the standard Diebold-Yilmaz (DY) framework has notable drawbacks: its reliance on arbitrarily chosen rolling windows makes estimates sensitive to window length, and the row-normalization procedure can introduce bias that affects the identification of net transmitters and receivers (Caloia *et al.*, 2019).

To address these limitations, (Antonakakis *et al.*, 2018) and (Antonakakis *et al.*, 2020) introduced the time-varying parameter VAR (TVP-VAR) approach, which models interdependencies as smoothly evolving processes without the need for fixed subsamples, thereby producing more stable and responsive spillover estimates. Building on this, (Lastrapes and Wiesen, 2021) proposed the Joint Connectedness Index (JCI), which corrects the normalization bias and yields a theoretically consistent measure of systemic connectedness. (Balciyar *et al.*, 2021) extended this to the Extended Joint Connectedness Index (EJCI), which retains the advantages of JCI while preserving bilateral net spillover measures, an essential feature for understanding directional dominance between markets. Further methodological advances have expanded the analytical toolkit: frequency-domain decompositions (Baruník and Křehlík, 2018) disentangle short- and long-term spillover components; quantile-based frameworks (Chatziantoniou *et al.*, 2021) capture asymmetries across market states; and DCC-GARCH-based  $R^2$  measures (Cocca *et al.*, 2024) reveal nonlinear and time-varying dependencies in high-frequency data.

Empirical applications to global and developed markets have yielded important insights into the nature and drivers of financial interconnectedness. (Diebold and Yilmaz, 2012) reported that total return spillovers among major global equity markets averaged around 65%, rising sharply during the 2008 Global Financial Crisis. (Baruník and Křehlík, 2018) showed that short-term spillovers dominate in turbulent periods, while long-term spillovers are more relevant during stable market phases. (Gabauer and Gupta, 2018) found that volatility spillovers among the G7 equity markets are highly time-varying, with the United States persistently acting as the dominant net transmitter of shocks. (Antonakakis *et al.*, 2020) documented that global financial uncertainty strongly amplifies connectedness across asset classes, particularly between equities and commodities. (Bouri *et al.*, 2021) demonstrated that during the COVID-19 pandemic, technology and healthcare sectors in developed markets emerged as net transmitters, while energy and travel-related sectors became net receivers. Collectively, these studies highlight that connectedness intensifies during crises, that transmission roles vary across sectors and regions, and that ignoring time variation and bias correction can distort conclusions about systemic risk and diversification opportunities.

Additionally, empirical research on spillover dynamics in emerging and frontier markets is expanding, with Eastern Europe providing a particularly active strand of the literature. (Syllignakis and Kouretas, 2011) analyze financial contagion across Central and Eastern European (CEE) equity markets using a dynamic conditional correlation (DCC-GARCH) framework, showing that connectedness with developed markets rises sharply during crises, thereby eroding diversification opportunities for emerging markets. Similarly, (Demiralay and Bayraci, 2015) study selected CEE stock markets volatility spillover and find that volatility transmission intensifies in turbulent periods, underscoring the structural vulnerability of emerging European markets to external shocks. Within the Diebold–Yilmaz (DY) framework, (Dumitrescu, 2015) shows that Eastern European equities were disproportionately affected by volatility and illiquidity spillovers during the Eurozone debt crisis, while (Louzis, 2013) finds that Euro area peripheral bonds shifted from receivers of return spillovers to transmitters of shocks under stress periods. Complementary evidence comes from (Cevik *et al.*, 2017) who apply asymmetric causality tests and reveal that negative shocks generate stronger spillovers in CEE markets, and (Özer *et al.*, 2020), who use frequency-domain causality to show that SEE markets experience both short- and long-term spillovers, largely driven by transmissions from advanced economies. Building on these foundations, (Škrinjarić, 2020) applies the generalized DY VAR framework to SEE and CEE equities and finds that these markets function largely as net receivers of shocks, with spillovers intensifying during crises. Subsequent work expands the framework: (Škrinjarić, 2021) incorporates asymmetric decompositions to show that downside spillovers dominate and that penalizing exposure to downside transmitters improves risk-return trade-offs.; (Škrinjarić *et al.*, 2021) apply realized semivariances to demonstrate that negative shocks drive stronger transmission and alter systemic risk assessments; and (Škrinjarić, 2022) higher-order moments (volatility, skewness, kurtosis), finding that these distributional features materially influence regional interconnectedness over time and are especially pronounced during crises. At the broader level, (Gamba-Santamaria *et al.*, 2019) employ a DCC-GARCH extension of the DY methodology to show that volatility spillovers surge during crises and originate mainly from advanced markets, while (Fountas *et al.*, 2024) demonstrate that uncertainty spillovers amplify return spillovers across stocks, bonds, and foreign exchange, with heterogeneous effects across both developed and emerging markets, and that European countries were particularly vulnerable during COVID-19.

While the literature on spillovers in Eastern European markets is substantial, two gaps remain. First, few studies provide a unified, daily-frequency analysis that jointly spans both advanced markets, such as Poland, Hungary, and Czechia and frontier markets, such as Serbia, Bosnia and Herzegovina, and North Macedonia, across major stress episodes (the European debt crisis, COVID-19, and the Russia–Ukraine conflict), which hinders inference on how roles and magnitudes evolve in crises. Second, prior applications for the region predominantly rely on rolling-window implementations; to our knowledge, no study applies a time-varying-parameter (TVP) model coupled with an Extended Joint Connectedness framework to this regional set, which is crucial for capturing smoothly evolving spillovers during multiple crises. To address this, we apply the Extended Joint Connectedness framework within a TVP-VAR to daily equity returns from ten Eastern European markets. This approach enables a granular assessment of the magnitude, direction, and evolution of return spillovers, while correcting for normalization bias and preserving bilateral net dynamics. By conditioning explicitly on the key systemic stress episodes: the European debt crisis, the COVID-19 pandemic, and the Russia-Ukraine conflict, we deliver a time-varying characterization of systemic interdependence in Eastern Europe that is informative for both policy and portfolio design.

### 3. METHODOLOGY

We employ the connectedness framework within a time-varying parameter VAR (TVP-VAR) setting, extended through the bias-corrected Joint Connectedness methodology. The framework builds on Diebold and Yilmaz (2009, 2012), who introduced forecast error variance decompositions (FEVD) from vector autoregressions (VAR) to measure total, directional, and net spillovers across markets. Following Antonakakis et al. (2018, 2020), we estimate a TVP-VAR model via the Kalman filter, which enables both coefficients and volatilities to evolve smoothly over time and avoids the limitations of fixed rolling windows. In the TVP-VAR(1) specification, the system is defined as:

$$\gamma_t = \beta_t \gamma_{t-1} + \varepsilon_t$$

where  $\gamma_t$  is a  $K \times 1$  vector of returns,  $\beta_t$  is the time-varying coefficient matrix, and  $\varepsilon_t \sim N(0, \Sigma_t)$  are innovations with a time-varying variance-covariance matrix  $\Sigma_t$ .

Applying the Wold representation to the TVP-VAR yields:

$$\gamma_t = \sum_{h=0}^{\infty} A_{h,t} \varepsilon_{t-h},$$

where  $A_{0,t} = I_k$  and  $A_{h,t} = \beta_t A_{h-1,t-1}$  for  $h \geq 1$ . This representation enables computation of the generalized FEVD without orthogonalization (Koop et al., 1996; Pesaran & Shin, 1998), capturing the fraction of forecast error variance in one market explained by shocks in another.

The EJC approach adjusts the generalized spillover table  $gSOT_{ij,t}$  so that each row's sum matches the bias-corrected joint connectedness share  $S_{from,i,t}^{jnt}$ . The scaling factor is:

$$\lambda_i = \frac{S_{from,i,t}^{jnt}}{S_{from,i,t}^{gen}}$$

where  $S_{from,i,t}^{gen}$  is the row-normalized share in the generalized framework. The bias-adjusted, direction-preserving bilateral spillovers are then:

$$\begin{aligned} jSOT_{ij,t} &= \lambda_i \times gSOT_{ij,t}, & i \neq j, \\ jSOT_{ii,t} &= 1 - S_{from,i,t}^{jnt}. \end{aligned}$$

This ensures that:

$$\sum_{j \neq i} jSOT_{ij,t} = S_{from,i,t}^{jnt}, \quad \sum_{j \neq i} jSOT_{ij,t} = S_{to,i,t}^{jnt}.$$

By construction, the VAR-TVP-EJC framework captures time-varying, bias-corrected, and bilateral connectedness measures without losing directional interpretation, making it well-suited for studying interconnected equity markets.

### 4. DATA

We utilize daily data for ten Eastern European stock market indices over the period January 1, 2010, to December 31, 2024. The sample includes the BELEX15 from the Belgrade Stock Exchange in Serbia, the BIRS from the Banja Luka Stock Exchange in Bosnia and Herzegovina, the CROBEX from the Zagreb Stock Exchange in Croatia, the SBITOP from the Ljubljana Stock Exchange in Slovenia, the MBI10 from the Macedonian Stock Exchange in North Macedonia, the BET from the Bucharest Stock Exchange in Romania, the PX from

the Prague Stock Exchange in the Czech Republic, the BUX from the Budapest Stock Exchange in Hungary, the WIG20 from the Warsaw Stock Exchange in Poland, and the PFTS from the PFTS Stock Exchange in Ukraine.

The selection of these indices ensures broad coverage of Eastern European markets. It encompasses EU member states such as Poland, Hungary, the Czech Republic, Croatia, Slovenia, and Romania, as well as EU candidate or neighboring markets, including Serbia, North Macedonia, and Bosnia and Herzegovina, together with the distinct case of Ukraine. This composition allows us to capture both relatively developed emerging markets and smaller frontier markets within the region.

We use daily log returns to ensure stationarity and interpret connectedness as return spillovers:  $r_{i,t} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$ . ADF and Phillips-Perron tests confirm stationarity by rejecting unit roots for all return series (Table 1). The mean returns are close to zero, with the highest values observed for MBI10 (0.036%) and BUX (0.035%). Daily volatility ranges between 0.7% and 1.3%, being relatively higher in Poland and Hungary (WIG20 and BUX) and lower in Bosnia and North Macedonia (BIRS and MBI10). The return distributions exhibit left skewness and heavy tails, as indicated by the Jarque-Bera statistic ( $p < 0.01$ ). Ljung-Box Q(20) tests on both returns and squared returns are significant, suggesting serial correlation and volatility clustering. Accordingly, we estimate connectedness in the return domain using a VAR specification.

Table 1: Descriptive statistics

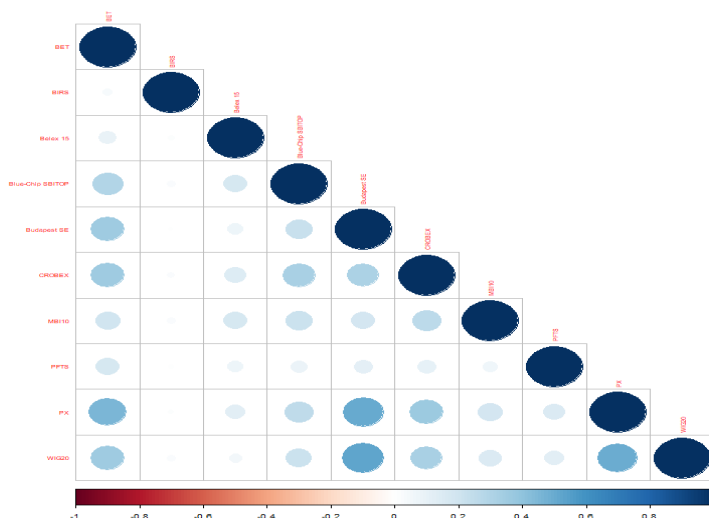
Asset	Mean	SD	Min	Max	Skewness	JBtest	ADFtest	PPtest	Q20	Q2_20
BET	0.00033	0.01039	-0.11892	0.09088	-1.293	57445.75***	-39.74**	-3463.86**	72.23***	799.33***
BIRS	-0.00002	0.00858	-0.11219	0.0945	0.262	120985.2***	-42.03**	-3489.61**	74.74***	947.66***
Belex 15	0.00014	0.00747	-0.07408	0.08229	-0.369	24488.4***	-38.95**	-3892.57**	86.33***	1549.87***
Blue-Chip SBITOP	0.00014	0.00874	-0.09383	0.05959	-1.008	17002.42***	-40.27**	-3786.76**	66.15***	1357.03***
Budapest SE	0.00035	0.01283	-0.12268	0.10674	-0.689	12827.47***	-44.65**	-3629.7**	44.91***	1384.86***
CROBEX	0.00012	0.00711	-0.10732	0.08563	-1.805	224300.71***	-40.29**	-4392.35**	149.2***	1392.65***
MBI10	0.00036	0.00797	-0.09848	0.06791	-1.1	99289***	-39.2**	-2930.39**	229.65***	2917.59***
PFTS	-0.00004	0.01157	-0.11379	0.24432	2.471	738054.38***	-38.25**	-2738.75**	332.15***	130.48***
PX	0.00012	0.01007	-0.0816	0.07369	-0.671	8582.91***	-41.55**	-3691.77**	49.16***	2418.44***
WIG20	-0.00003	0.01313	-0.14246	0.08099	-0.611	8909.94***	-44.51**	-3662.43**	28.11	1034.55***

(Source: Author's compilation)

As an initial measure of interdependence, we calculate unconditional Pearson correlations of daily returns (Figure 1). Correlations are generally positive but display considerable heterogeneity. Central European EU members, Poland (WIG20), Hungary (BUX), and the Czech Republic (PX), exhibit relatively strong co-movements, with coefficients in the 0.4-0.6 range. Croatia and Slovenia also show moderate correlations with each other and with neighboring markets.

In contrast, the smaller and more peripheral markets of Bosnia (BIRS), Serbia (Belex15) and North Macedonia (MBI10), as well as Ukraine's PFTS, record much lower correlations, often below 0.2, reflecting idiosyncratic dynamics and limited foreign participation. Overall, while Eastern European equities move together to some extent, integration remains partial, underscoring the need for a spillover analysis that captures directional transmission and influence beyond simple contemporaneous correlations.

Figure 1: Heatmap of unconditional Pearson correlations



(Source: Author’s compilation)

## 5. EMPIRICAL RESULTS

### 5.1 Static spillover analysis

We begin by examining the full-sample spillover estimates, which provide an aggregate perspective on the degree of connectedness among the ten Eastern European stock markets over the 2010-2024 period. Table 2 reports the spillover matrix and summary statistics derived from the Extended Joint Connectedness framework. Each off-diagonal element of the matrix measures the share of forecast error variance in the returns of the market listed in the row that can be attributed to shocks originating from the market in the corresponding column, averaged across the entire sample. The diagonal elements capture own-market effects, while the off-diagonal entries quantify cross-market spillovers, thereby offering a comprehensive view of how shocks are transmitted across the region’s equity markets.

Table 2: Spillover table using EJC

Market	BET	BIRS	Belex 15	Blue-Chip SBITOP	Budapest SE	CROBE X	MBI1 0	PFTS	PX	WIG2 0	<b>FRO M</b>
BET	67.09	0.61	0.92	3.60	5.60	4.74	1.72	1.69	8.24	5.79	<b>32.91</b>
BIRS	1.20	92.03	0.65	1.07	0.96	0.86	1.05	0.61	0.62	0.95	<b>7.97</b>
Belex 15	1.90	0.78	85.51	2.99	1.13	1.90	2.43	0.75	1.75	0.87	<b>14.49</b>
Blue-Chip SBITOP	4.38	0.77	2.07	75.96	2.51	4.67	2.43	0.86	3.20	3.14	<b>24.04</b>
Budapest SE	5.60	0.48	0.44	2.18	61.34	3.26	1.60	1.39	10.68	13.03	<b>38.66</b>
CROBEX	4.90	0.74	1.03	4.09	3.37	73.55	2.41	1.41	4.58	3.93	<b>26.45</b>
MBI10	2.44	0.51	2.01	2.64	1.99	3.31	82.41	0.78	2.00	1.90	<b>17.59</b>
PFTS	2.06	0.46	0.64	0.81	1.58	1.28	0.79	88.53	1.75	2.09	<b>11.47</b>
PX	8.53	0.43	0.73	2.84	10.28	4.44	1.50	1.56	58.59	11.10	<b>41.41</b>
WIG20	5.87	0.63	0.50	2.25	13.17	3.71	1.46	1.74	11.29	59.37	<b>40.63</b>
<b>TO</b>	<b>36.89</b>	<b>5.43</b>	<b>8.98</b>	<b>22.48</b>	<b>40.59</b>	<b>28.18</b>	<b>15.39</b>	<b>10.78</b>	<b>44.10</b>	<b>42.80</b>	<b>255.62</b>
Inc.Own	103.97	97.45	94.49	98.44	101.93	101.73	97.81	99.32	102.68	102.17	TCI
<b>NET</b>	<b>3.97</b>	<b>-2.55</b>	<b>-5.51</b>	<b>-1.56</b>	<b>1.93</b>	<b>1.73</b>	<b>-2.19</b>	<b>-0.68</b>	<b>2.68</b>	<b>2.17</b>	<b>25.56</b>

NPT	8.00	1.00	0.00	3.00	8.00	4.00	3.00	4.00	8.00	6.00	
-----	------	------	------	------	------	------	------	------	------	------	--

(Source: Author's compilation)

Several notable patterns emerge from Table 2. The diagonal entries are generally dominant, indicating that own shocks explain the majority of forecast error variance across markets. In most cases, own-share contributions lie between 60% and 90%, underscoring the persistence of domestic or idiosyncratic drivers. The smaller Balkan markets, Bosnia (BIRS), Serbia (BELEX15), and North Macedonia (MBI10), display particularly high own-shares, exceeding 85%, consistent with their relatively low integration and limited exposure to regional spillovers. By contrast, the larger Central European EU members, Poland (WIG20), the Czech Republic (PX), Romania (BET), and Hungary (BUX), exhibit lower own-shares, in the range of 58-67%, meaning that one-third or more of their return variance is explained by shocks from other markets. This contrast highlights a dual structure in the region: frontier markets that remain largely segmented and more developed markets that are increasingly interconnected with both regional and global financial systems. Similarly, (Syllignakis & Kouretas, 2011) find stronger spillovers in larger CEE markets than in smaller Balkan markets, reinforcing the dual-market structure.

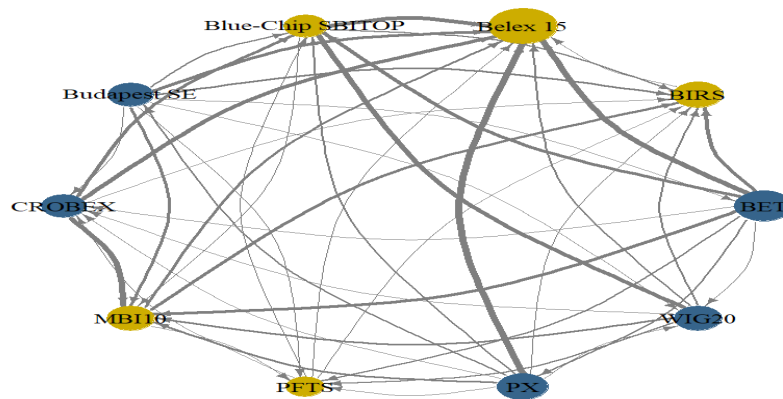
The off-diagonal entries in Table 2 reveal which cross-market linkages are most pronounced. While most spillover shares are modest, typically in the 1-5% range, a few notable bilateral connections stand out. Poland's WIG20 emerges as a key transmitter, with shocks spilling over significantly to Hungary (BUX) and the Czech Republic (PX), where contributions exceed 10% of forecast variance. Hungary and the Czech Republic also act as important transmitters, with Hungary in particular exerting meaningful influence on Poland and Romania. These patterns underline the role of the larger Central European EU markets as the main hubs of regional shock propagation. By contrast, spillovers from the smaller Balkan markets, Bosnia (BIRS), North Macedonia (MBI10), and Serbia (BELEX15), are negligible, generally below 1%, reflecting their limited capacity to influence neighboring or larger markets. Overall, the evidence points to an asymmetric structure in which Central European markets drive regional dynamics, while the smaller frontier markets remain peripheral and predominantly receivers of external shocks. This hub-and-periphery structure is consistent with (Dumitrescu, 2015), who shows that Eastern European markets are disproportionately affected by volatility shocks from the EU cores, and with (Cevik et al., 2017), who find asymmetric transmission led by larger markets.

In aggregate terms, the system's Total Connectedness Index (TCI) is 25.6% (Table 2), implying that roughly one quarter of forecast error variance is attributable to cross-market spillovers, while the remaining 74% reflects domestic shocks. This points to a moderate degree of interdependence: integration is significant enough to allow contagion risk yet limited enough to preserve diversification potential. For comparison, studies on developed markets typically report much higher connectedness, often above 50% and rising sharply during crises. For instance, Youssef et al. (2021), using a TVP-VAR framework for major advanced and emerging economies, document an average TCI of about 65%, peaking near 80% during the COVID-19 outbreak.

In addition, Figure 2 provides a complementary perspective through a network representation of net pairwise spillovers. Node size reflects the absolute value of the Net Spillover Index (NET), with larger nodes exerting stronger net influence, while color indicates direction (blue for net transmitters, yellow for net receivers). Edges are directional and weighted by the magnitude of spillovers. The network reveals a Central European core of transmitters, led by

Romania (+3.97), the Czech Republic (+2.68), and Poland (+2.17), followed by Hungary (BUX) and Croatia (CROBEX) as moderate contributors. On the other hand, Serbia (BELEX15) emerges as the strongest receiver (-5.51), alongside Bosnia (BIRS), North Macedonia (MBI10), Slovenia (SBITOP), and Ukraine (PFTS). Overall, the network depicts a multipolar structure, with a Central European transmitter core radiating shocks to a periphery of receivers, consistent with the moderate but non-trivial systemwide connectedness reported in Table 2.

Figure 2: Net pairwise directional connectedness



(Source: Author's compilation)

To assess robustness, we compare the EJC results against two widely used alternatives: (i) the standard Diebold–Yilmaz framework, based on a static VAR with row normalization, and (ii) a connectedness measure from a TVP-VAR estimated without the EJC adjustment (see Appendix, Tables 1 and 2). The three approaches yield broadly consistent outcomes. Total connectedness is of similar magnitude, 30.99% under VAR-DY and 29.27% under TVP-VAR, as well as the classification of net transmitters and receivers remains unchanged. As expected, the row-normalization in the DY variant produces somewhat more extreme net values, whereas the EJC adjustment moderates them slightly. Nonetheless, the core-periphery structure and the relative market rankings are preserved across specifications. Overall, the small discrepancies observed indicate that the regional system is relatively balanced and not overly sensitive to methodological choice (Balcilar *et al.*, 2021).

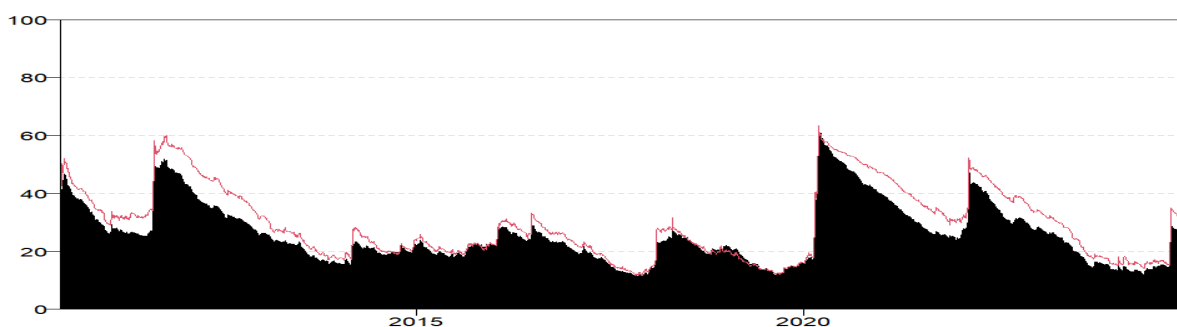
## 5.2 Dynamic evolution of connectedness

The full-sample results provide an average picture, but the dynamic perspective is essential for understanding how connectedness evolves across time. A key contribution of this study is to trace these dynamics and identify how global and regional shocks alter the degree of integration among Eastern European markets. Figure 3 presents the Dynamic Total Connectedness Index (TCI) for the period 2010-2024, estimated using both the EJC (shaded area) and the TVP-VAR (red line). In parallel, Figure 4 displays the time-varying net spillovers for each market, again comparing the two approaches.

Figure 3 demonstrates that the connectedness of Eastern European stock markets is highly time-varying, fluctuating markedly in response to global and regional shocks. During 2010-2012, connectedness was elevated, with a pronounced spike in late 2011 that coincides with the Eurozone sovereign debt crisis, a period of severe financial stress. Contagion from Western European bonds and banking markets spilled over into Eastern Europe, particularly affecting EU member states. At its peak in late 2011, the TCI reached nearly 50%, compared

with baseline levels around 25-30% during calmer periods, indicating that spillovers intensified as markets moved more closely in tandem under stress. This is consistent with (Syllignakis and Kouretas, 2011), who identify sharp increases in dynamic correlations between CEE and developed markets during the Global Financial and Eurozone debt crises. Importantly, even non-EU markets such as Serbia and Bosnia exhibited stronger linkages during this period, likely transmitted through regional banks and shifts in investor sentiment. This interpretation is supported by Figure 4, where their net spillover positions turn sharply negative, confirming that these smaller markets were primarily absorbers of external shocks during the crisis.

*Figure 3: Dynamic Total Connectedness index (TCI)*



*(Source: Author's compilation)*

In the mid-2010s (2013-2019), the TCI stabilized at relatively low levels, fluctuating mostly between 15-25%. This reflected a period of economic recovery and relative calm in global markets, with local factors dominating regional dynamics. Only modest increases are visible during external disturbances such as the Chinese market turbulence (2015-2016) and the Brexit referendum, but their effects on Eastern Europe were limited. As seen in Figure 4, net transmitter and receiver roles were stable: Poland, the Czech Republic, and Hungary consistently acted as mild transmitters, while the smaller Balkan markets remained receivers.

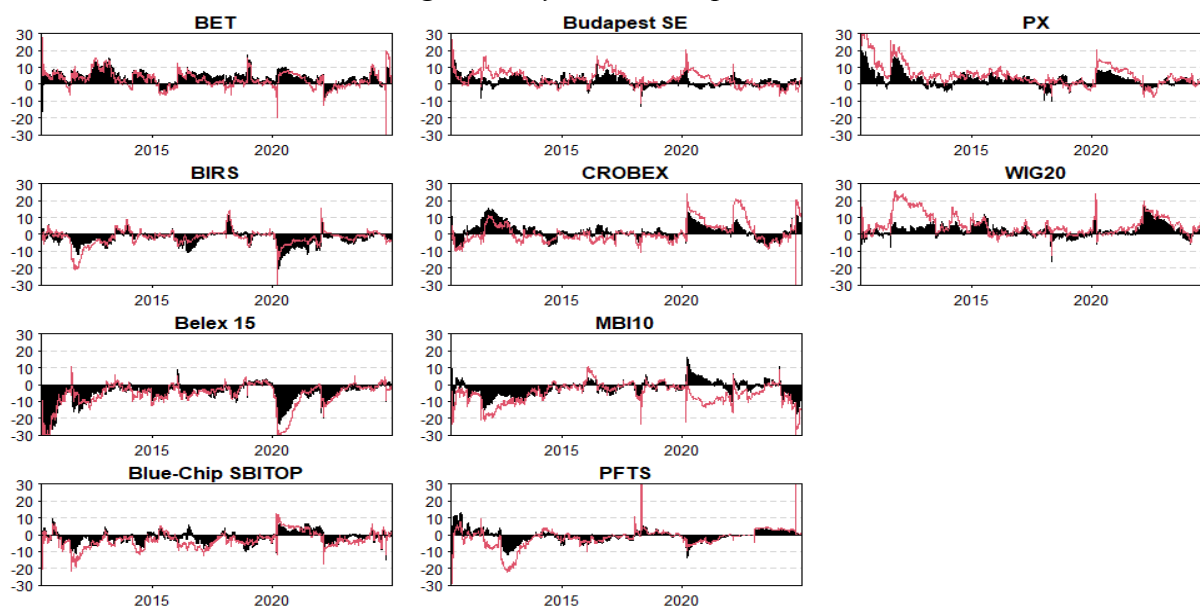
A dramatic shift occurred in early 2020 with the onset of the COVID-19 pandemic. In March 2020, the TCI spiked to its highest level in the sample, reaching around 60%, indicating unprecedented synchronization as markets reacted to a common global shock. Similar COVID-19 spillover spikes are documented in the broader literature (Aslam et al., 2021; Rehman et al., 2022; Youssef et al., 2021). After the initial crash, connectedness declined but remained above pre-pandemic baselines for much of 2020, reflecting sustained policy-driven co-movements.

Another pronounced episode arose in late 2021-2022 during the escalation of the Russia-Ukraine conflict. The TCI increased from around 20% in late 2021 to nearly 40% in early 2022, reflecting tighter co-movements as geopolitical tensions unfolded. Unlike COVID, this rise was less abrupt but more persistent, maintaining elevated levels for much of 2022. By 2023-2024, connectedness reverted toward mid-2010s baselines, with only minor spikes linked to global inflation shocks and monetary policy shifts, suggesting no lasting structural increase in regional integration.

Furthermore, Figure 4 shows that Poland (WIG20), the Czech Republic (PX), and Hungary (BUX) consistently emerge as net transmitters, with positions remaining above zero across most of the sample. By contrast, Serbia (BELEX15), Bosnia (BIRS), and North Macedonia (MBI10) are persistent net receivers, with their positions becoming more negative during

major crises, such as late 2011, March 2020, and early 2022, indicating disproportionate absorption of external shocks. In calmer periods, these markets revert toward mildly negative or near-zero values, reflecting limited but ongoing dependence on regional spillovers. A notable exception appears in the immediate post-COVID rebound, when North Macedonia's MBI10 briefly turns net positive under the EJC specification. This shift, absent in the TVP-VAR results, underscores how methodological differences can affect the classification of smaller and less liquid markets.

*Figure 4: Dynamic Net Spillover*



*(Source: Author's compilation)*

## 6. CONCLUSION

Our findings indicate that Eastern European equity markets display a moderate degree of interconnectedness. The Total Connectedness Index averages around 25%, implying that roughly one-quarter of return variance is driven by cross-market shocks. This suggests that while these markets are not isolated, showing meaningful co-movement and contagion, they remain less tightly integrated than major global markets. Accordingly, diversification opportunities persist within the region, although they tend to diminish during periods of financial turbulence.

We find clear evidence that connectedness in Eastern European equity markets is highly dynamic and event-driven. Periods of global or regional turmoil, such as the European debt crisis (2011-2012), the COVID-19 pandemic (2020), and the Russia-Ukraine conflict (2022), coincide with sharp spikes in the Total Connectedness Index, as markets moved more closely in tandem under stress. By contrast, during calmer periods such as the mid-2010s, spillovers subsided, and markets were more segmented, with domestic factors exerting a stronger influence on returns.

In terms of directional spillovers, our results reveal a clear core-periphery structure within the regional network. The larger and more internationally integrated markets, particularly Poland, the Czech Republic, Romania, and Hungary, consistently act as net transmitters of shocks. Their size, liquidity, and stronger links to global investors position them as sources of

volatility that influence neighboring markets. By contrast, smaller markets such as Serbia, Bosnia and Herzegovina, and North Macedonia function primarily as net receivers. They exert little influence on others but remain vulnerable to spillovers from larger regional players and global shocks. Slovenia and Ukraine also lean toward net receiver status, although Slovenia, as an EU member, exhibits more intermediate behavior, while Ukraine's dynamics in 2022 were atypical due to war-related disruptions.

The findings carry important implications for both investors and policymakers. For investors with exposure to Eastern European equities, the results suggest that diversification within the region is beneficial but limited. In calmer periods, cross-market spillovers remain moderate, allowing portfolios to spread across countries to achieve meaningful risk reduction. However, during episodes of global or regional stress, correlations rise sharply, and diversification benefits erode. Investors should therefore remain attentive to these dynamics and consider hedging strategies when systemic risks are anticipated. For policymakers and regulators, the evidence highlights the critical importance of monitoring cross-market spillovers. Recent crises have demonstrated that even small markets in the Western Balkans remain highly vulnerable to external shocks, underscoring the need for sustained oversight of contagion risks.

The main limitation of this study is its focus on only stock index returns. Future research could broaden the scope to include other asset classes, such as bonds or exchange rates, to assess whether similar dynamics emerge. It would also be valuable to explore frequency-dependent spillovers, distinguishing between short and long-term connectedness, to better understand whether co-movements are driven by high-frequency speculative activity or by lower-frequency fundamental linkages.

## REFERENCES

- Antonakakis, N., Chatziantoniou, I., & Gabauer, D. (2020), "Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions", *Journal of Risk and Financial Management*, Vol. 13 No. 4, p. 84.
- Antonakakis, N., Gabauer, D., Gupta, R. and Plakandaras, V. (2018), "Dynamic connectedness of uncertainty across developed economies: A time-varying approach", *Economics Letters*, Vol. 166, pp. 63-75.
- Aslam, F., Ferreira, P., Mughal, K. S. and Bashir, B. (2021), "Intraday volatility spillovers among European financial markets during COVID-19", *International Journal of Financial Studies*, Vol. 9 No. 1, p. 5.
- Balcilar, M., Gabauer, D. and Umar, Z. (2021), "Crude Oil futures contracts and commodity markets: New evidence from a TVP-VAR extended joint connectedness approach", *Resources Policy*, Vol. 73, 102219.
- Baruník, J. and Křehlík, T. (2018), "Measuring the frequency dynamics of financial connectedness and systemic risk", *Journal of Financial Econometrics*, Vol. 16 No. 2, pp. 271-296.
- Bouri, E., Cepni, O., Gabauer, D. and Gupta, R. (2021), "Return connectedness across asset classes around the COVID-19 outbreak", *International Review of Financial Analysis*, Vol. 73, 101646.
- Caloia, F. G., Cipollini, A. and Muzzioli, S. (2019), "How do normalization schemes affect net spillovers? A replication of the Diebold and Yilmaz (2012) study", *Energy Economics*, Vol. 84, p. 104536.

- Cevik, E. I., Korkmaz, T. and Cevik, E. (2017), "Testing causal relation among central and eastern European equity markets: Evidence from asymmetric causality test", *Economic Research-Ekonomska Istraživanja*, Vol. 30 No. 1, pp. 381-393.
- Chatziantoniou, I., Gabauer, D. and Stenfors, A. (2021), "Interest rate swaps and the transmission mechanism of monetary policy: A quantile connectedness approach" *Economics Letters*, Vol. 204, 109891.
- Cocca, T., Gabauer, D. and Pomberger, S. (2024), *Clean Energy Market Connectedness and Investment Strategies: New Evidence from DCC-GARCH R2 Decomposed Connectedness Measures* (SSRN Scholarly Paper No. 4810641), Social Science Research Network.
- Demiralay, S. and Bayraci, S. (2015), "Central and Eastern European stock exchanges under stress: A range-based volatility spillover framework", *Finance a Uver: Czech Journal of Economics & Finance*, Vol. 65 No. 5.
- Diebold, F. X. and Yilmaz, K. (2009), "Measuring financial asset return and volatility spillovers, with application to global equity markets", *The Economic Journal*, Vol. 119 No. 534, pp. 158-171.
- Diebold, F. X. and Yilmaz, K. (2012), "Better to give than to receive: Predictive directional measurement of volatility spillovers", *International Journal of Forecasting*, Vol. 28 No. 1, pp. 57-66.
- Diebold, F. X. and Yilmaz, K. (2014), "On the network topology of variance decompositions: Measuring the connectedness of financial firms", *Journal of Econometrics*, Vol. 182 No. 1, pp. 119-134. <https://doi.org/10.1016/j.jeconom.2014.04.012>
- Dumitrescu, S. (2015), "European equity market return, volatility and liquidity spillover dynamics during the Eurozone debt crisis", *Financial Studies*, Vol. 19 No. 2.
- Fountas, S., Kontana, D. and Tzika, P. (2024), "Uncertainty and financial asset return spillovers: Are they related? Empirical evidence from three continents", *Empirical Economics*, Vol. 67 No. 5, pp. 1891-1918.
- Gabauer, D. and Gupta, R. (2018), "On the transmission mechanism of country-specific and international economic uncertainty spillovers: Evidence from a TVP-VAR connectedness decomposition approach", *Economics Letters*, Vol. 171, pp. 63-71.
- Gamba-Santamaria, S., Gomez-Gonzalez, J. E., Hurtado-Guarin, J. L. and Melo-Velandia, L. F. (2019), "Volatility spillovers among global stock markets: Measuring total and directional effects", *Empirical Economics*, Vol. 56 No. 5, pp. 1581-1599.
- Koop, G., Pesaran, M. H. and Potter, S. M. (1996), "Impulse response analysis in nonlinear multivariate models", *Journal of Econometrics*, Vol. 74 No. 1, pp. 119-147.
- Lastrapes, W. D. and Wiesen, T. F. (2021), "The joint spillover index", *Economic Modelling*, Vol. 94, pp. 681-691.
- Louzis, D. P. (2013). *Measuring return and volatility spillovers in Euro area financial markets*.
- Özer, M., Kamenković, S. and Grubišić, Z. (2020), "Frequency domain causality analysis of intra- and inter-regional return and volatility spillovers of South-East European (SEE) stock markets", *Economic Research-Ekonomska Istraživanja*, Vol. 33 No. 1, pp. 1-25.
- Pesaran, H. H. and Shin, Y. (1998), "Generalized impulse response analysis in linear multivariate models", *Economics Letters*, Vol. 58 No. 1, pp. 17-29.
- Rehman, M. U., Ahmad, N., Shahzad, S. J. H. and Vo, X. V. (2022), "Dependence dynamics of stock markets during COVID-19", *Emerging Markets Review*, Vol. 51, 100894.
- Škrinjarić, T. (2020), "CEE and SEE equity market return spillovers: Creating profitable investment strategies", *Borsa Istanbul Review*, Vol. 20, S62-S80.
- Škrinjarić, T. (2021), "Asymmetric spillovers between the stock risk series: Case of CESEE stock markets. In Patnaik, S., Tajeddini, K. and Jain, V. (Eds.), *Computational*

- Management: Applications of Computational Intelligence in Business Management* (pp. 97-135). Springer International Publishing.
- Škrinjarić, T. (2022), "Higher moments actually matter: Spillover approach for case of CESEE stock markets", *Mathematics*, Vol. 10 No. 24, 4811.
- Škrinjarić, T., Golubić, Z. and Orlović, Z. (2021), *Asymmetric spillovers on European stock markets: "Good" and "bad" volatility approach*.
- Syllignakis, M. N. and Kouretas, G. P. (2011), "Dynamic correlation analysis of financial contagion: Evidence from the Central and Eastern European markets", *International Review of Economics & Finance*, Vol. 20 No. 4, pp. 717-732.
- Youssef, M., Mokni, K. and Ajmi, A. N. (2021), "Dynamic connectedness between stock markets in the presence of the COVID-19 pandemic: Does economic policy uncertainty matter?", *Financial Innovation*, Vol. 7 No. 1, p. 13.

**APPENDIX**

*Table 1: Spillover analysis using VAR*

Market	BET	BIRS	Belex 15	Blue-Chip SBITOP	Budapest SE	CROBEX	MBI10	PFTS	PX	WIG20	FROM
BET	56.79	0.10	0.60	4.92	7.45	6.99	2.43	1.80	11.40	7.51	43.21
BIRS	0.12	99.11	0.02	0.06	0.20	0.06	0.17	0.17	0.01	0.08	0.89
Belex 15	2.35	0.07	84.34	2.45	1.12	3.02	2.71	0.59	2.61	0.75	15.66
Blue-Chip SBITOP	6.28	0.11	1.80	66.97	4.01	7.32	3.12	0.57	5.10	4.72	33.03
Budapest SE	6.86	0.00	0.36	2.80	53.39	5.22	1.87	0.77	13.70	15.03	46.61
CROBEX	7.42	0.06	1.28	6.19	5.89	59.57	4.54	0.66	7.74	6.64	40.43
MBI10	3.76	0.02	1.70	3.22	2.78	6.20	76.57	0.37	3.02	2.35	23.43
PFTS	2.77	0.00	0.34	0.36	1.59	0.90	0.35	89.31	2.30	2.08	10.69
PX	10.16	0.05	0.74	3.22	13.06	6.40	1.76	1.14	50.55	12.91	49.45
WIG20	6.77	0.04	0.22	2.60	15.10	5.64	1.59	0.96	13.54	53.54	46.46
TO	46.49	0.46	7.06	25.81	51.21	41.76	18.54	7.03	59.43	52.08	309.86
Inc.Own	103.28	99.57	91.40	92.79	104.60	101.32	95.10	96.34	109.99	105.62	cTCI/TCI
NET	3.28	-0.43	-8.60	-7.21	4.60	1.32	-4.90	-3.66	9.99	5.62	34.43/30.99
NPT	6.00	3.00	0.00	2.00	8.00	5.00	2.00	4.00	8.00	7.00	

*Source: Author's compilation*

*Table 2: Spillover analysis using TVP-VAR*

Market	BET	BIRS	Belex 15	Blue-Chip SBITOP	Budapest SE	CROBEX	MBI10	PFTS	PX	WIG20	FROM
BET	60.36	0.73	1.11	4.33	6.76	5.76	2.03	2.06	9.84	7.02	39.64
BIRS	1.21	92.05	0.64	1.12	0.96	0.89	1.00	0.60	0.62	0.91	7.95
Belex 15	2.24	0.81	83.68	3.28	1.25	2.24	2.75	0.81	1.95	0.98	16.32
Blue-Chip SBITOP	5.32	0.91	2.46	71.04	3.05	5.55	2.88	1.07	3.90	3.83	28.96
Budapest SE	6.25	0.52	0.49	2.44	57.38	3.67	1.78	1.55	11.81	14.13	42.62
CROBEX	5.99	0.87	1.27	4.77	4.07	68.10	2.80	1.78	5.58	4.77	31.90
MBI10	3.16	0.56	2.40	3.43	2.59	4.39	77.40	0.95	2.63	2.49	22.60
PFTS	2.63	0.47	0.68	0.91	2.00	1.55	0.89	86.08	2.16	2.63	13.92
PX	9.24	0.47	0.81	3.19	11.21	4.92	1.63	1.64	54.78	12.10	45.22
WIG20	6.35	0.67	0.51	2.54	13.89	4.07	1.63	1.78	12.14	56.41	43.59
TO	42.40	6.01	10.37	26.01	45.79	33.03	17.40	12.25	50.63	48.86	292.74
Inc.Own	102.76	98.06	94.05	97.05	103.17	101.13	94.79	98.33	105.40	105.27	cTCI/TCI
NET	2.76	-1.94	-5.95	-2.95	3.17	1.13	-5.21	-1.67	5.40	5.27	32.53/29.27
NPT	6.00	1.00	0.00	3.00	7.00	4.00	2.00	5.00	9.00	8.00	

*(Source: Author's compilation)*

## QUANTIFYING FDI'S EFFECTS ON GDP AND UNEMPLOYMENT: EVIDENCE FROM NORTH MACEDONIA

**Bojan Malcev**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
bojan.malcev@eccf.ukim.edu.mk*

**Marina Trpeska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
marinas@eccf.ukim.edu.mk*

**Zoran Minovski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
zoran.minovski@eccf.ukim.edu.mk*

### **ABSTRACT**

*This paper examines the impact of foreign direct investment (FDI) on economic growth and unemployment in North Macedonia over the period 2014–2023. North Macedonia, a small post-transition economy with historically high unemployment, has actively pursued FDI as a development strategy. Using annual data and econometric analysis (stationarity tests, Pearson correlations, and OLS regressions in SPSS), we test four hypotheses about FDI's relationship with GDP growth and unemployment. The results indicate a strong positive association between FDI inflows and real GDP growth and a significant negative association between FDI and the unemployment rate. In particular, higher FDI is correlated with faster GDP growth and lower unemployment, supporting the view that FDI can be a catalyst for economic development. Regression analysis further suggests that FDI has a statistically significant positive effect on GDP growth and a negative effect on unemployment, even when accounting for the growth-employment link. These findings confirm the optimistic hypothesis that FDI inflows drive macroeconomic improvements in North Macedonia. However, complementary factors (institutional quality, human capital) are crucial for maximizing FDI benefits. The paper concludes with policy implications, emphasizing the need to attract quality FDI and strengthen domestic absorptive capacities to ensure sustainable growth and job creation.*

**Keywords:** *Foreign direct investment; Economic growth; Unemployment; North Macedonia; Transition economy*

**JEL classification:** *F21; E24; O40*

### **1. INTRODUCTION**

North Macedonia is a small, open economy that has undergone significant structural changes since the 1990s transition from socialism. Like many post-transition countries in the Western Balkans, it has faced the dual challenge of stimulating economic growth while reducing persistently high unemployment. In recent decades, successive governments have viewed foreign direct investment (FDI) as a crucial catalyst for development, actively promoting FDI as a strategy to boost GDP and create jobs. This policy focus stems from a widely held belief

in both policy and academic circles that attracting FDI can help transform the Macedonian economy by injecting much-needed capital, technology, and managerial know-how, thereby spurring economic growth and employment.

From a theoretical standpoint, there are strong reasons to expect FDI to play a positive role in post-transition economies. According to Dunning's OLI paradigm (Ownership–Location–Internalization), multinational enterprises choose to invest abroad when three sets of advantages coincide (Dunning and Lundan, 2008). The host country must offer location advantages – such as skilled labor, resources, market potential, or favorable institutions – that complement the firm's ownership advantages (technology, brand, etc.), making the investment profitable. North Macedonia's efforts to improve its business climate through institutional reforms, privatization, and the creation of Technological Industrial Development Zones can be seen as attempts to strengthen these location advantages and attract FDI. Reforms have been implemented to ensure a stable regulatory framework, protect property rights, and foster closer integration with European markets, reflecting an awareness that institutional quality is crucial in attracting foreign investors. Indeed, research shows that Western Balkan countries (including North Macedonia) have lagged behind other Eastern European transitions in FDI inflows even after accounting for market size, distance, and institutional factors. Estrin and Uvalic (2016) argue that this FDI shortfall is due in part to the need for deeper institutional reforms – improving governance, reducing risks, and strengthening the rule of law – to create a more attractive investment environment. Thus, the topic of this study is both socially and academically important, given the pressing question of how FDI interacts with domestic conditions to influence economic outcomes in a transition economy like North Macedonia.

Beyond capital inflows, FDI also carries the potential for technological spillovers and productivity gains in the host economy. Endogenous growth theory (Romer, 1990) posits that long-run growth is driven by technological change and knowledge diffusion. FDI is one channel through which developing and transition economies can access advanced technologies, managerial expertise, and new ideas developed in richer countries. Romer (1993) famously highlighted the “idea gap” between poor and rich nations and suggested that foreign investment can help bridge that gap by transferring technology and business skills to the host. Inflows of FDI serve as conduits for innovation: multinationals introduce new products and processes, train local workers, and may induce competitive pressures that push domestic firms to improve. Such spillover effects are especially valuable for post-socialist economies that began the transition with outdated industrial bases and a lack of capital. Empirical studies support this view: for example, Borensztein *et al.* (1998) found that FDI is a significant channel for technology transfer, contributing more to a host country's growth than domestic investment – provided the host country has sufficient human capital to absorb the new technologies. This suggests that FDI's impact on GDP can be substantial in countries that invest in education and skills - a relevant insight for North Macedonia, which has been striving to upgrade its workforce's qualifications. In terms of employment, FDI in labor-intensive greenfield projects can directly create new jobs and indirectly generate employment through supply-chain linkages and productivity effects. The experience of Central and Eastern European countries that joined the EU shows that steady FDI inflows were associated with faster economic growth and rising employment.

However, the benefits of FDI are not automatic. The literature also contains more cautious findings that highlight the importance of local conditions. Carkovic and Levine (2002), using advanced econometric techniques to control for causality, conclude that there is no robust independent effect of FDI on economic growth across countries. In other words, FDI by itself

will not spur development if complementary factors are not present – such as sound economic policies, developed financial markets, and quality institutions. Micro-level studies similarly show mixed evidence of spillovers; positive effects appear only in certain industries or under specific policy environments. In the context of North Macedonia and its neighbors, Estrin (2016) observed that although FDI brought significant benefits in terms of output and employment in Central Europe, in the Balkans the effects have been weaker – likely due to weaker institutions and insufficient human capital to fully exploit the investments. For example, North Macedonia established free economic zones in the 2000s and offered incentives that attracted manufacturing FDI, but the overall unemployment rate remained high for years. One recent study even found that FDI inflows from 1999 to 2013 had no statistically significant effect on reducing unemployment in North Macedonia, suggesting that the quality of investment matters more than quantity, as does the country's capacity to absorb it (Nikoloski, 2017). Contrary to this, Angeloska (2020) analyzed the period 2005-2018 and found that there is a positive correlation between FDI and the GDP, employment rate trends, and export. These nuanced findings underscore why a careful analysis of FDI's effects is socially important: if FDI is to fulfill its role as an engine of development, policymakers must understand the conditions under which it truly helps increase GDP and reduce unemployment. Focusing on the case of North Macedonia – a representative example of a small, developing post-transition economy striving to integrate into the global market – this research addresses the question of whether and how FDI contributes to achieving the twin goals of higher GDP growth and lower unemployment in North Macedonia. The results can inform better policy design to maximize FDI benefits (for instance, through education or institutional improvements) and contribute to the broader academic debate on FDI's effectiveness in transitional contexts.

The motivation for this study lies in the critical role FDI is believed to play in North Macedonia's economic transformation. The institutional reforms, technological spillovers, and labor market changes associated with FDI are three key aspects often cited in this context. FDI is seen as accelerating the shift from an outdated, state-led economy to a dynamic market economy by bringing in capital and knowledge, as well as influencing the local labor market with new employment practices and skill demands. The societal relevance, therefore, is found in assessing whether and how FDI contributes to the dual objective of higher economic growth and lower unemployment in North Macedonia. This study is timely, as North Macedonia continues to pursue EU membership – a process that itself has prompted reforms and could further affect FDI inflows and their economic impact.

The remainder of the paper is structured as follows: Section 2 reviews relevant literature on FDI, growth, and unemployment, with a focus on recent empirical findings. Section 3 outlines the methodology, including data sources, variables, and the econometric approach (unit root tests, correlation analysis, and regression modeling). Section 4 presents the results of the analysis and tests the hypotheses H1–H4, including descriptive trends, correlation matrices, and regression outputs (coefficients, significance levels,  $R^2$ , and diagnostic tests), accompanied by visualizations (time-series graphs and fitted regression plots). Section 5 concludes and provides policy implications and suggestions for further research.

## **2. LITERATURE REVIEW**

### **2.1. FDI and Economic Growth**

The relationship between FDI and economic growth has been extensively studied, yielding mixed evidence. On the optimistic side, many empirical studies find that FDI can stimulate

host country growth by supplementing domestic investment and facilitating technology transfer. As noted earlier, Borensztein *et al.* (1998) found FDI to be a significant vehicle for technological advancement in developing countries, with a larger growth impact than domestic investment (conditional on sufficient human capital). This aligns with endogenous growth theory, where FDI-driven knowledge spillovers can raise the host economy's long-run growth trajectory. Transition economies that successfully attracted FDI often experienced periods of accelerated GDP growth. For instance, several Central and Eastern European countries saw faster GDP growth in conjunction with high FDI inflows in the 2000s, suggesting a positive FDI-growth linkage under conducive conditions. Recent studies continue to support positive effects: a panel study of developing and Western Balkan countries (2015–2019) by Kukaj *et al.* (2022) finds that FDI inflows have a considerable impact on increasing GDP and reducing unemployment. In their regression analysis, FDI was significantly associated with higher GDP levels and was identified as a driver of employment gains, reinforcing the view that FDI can be an engine of growth and job creation in transitional economies.

The theoretical rationale for FDI-led growth is grounded in FDI's role as a composite bundle of capital, technology, and managerial expertise. Dunning and Lundan (2010) emphasize that multinational enterprises not only bring financial capital but also introduce new technologies, efficient production techniques, and modern management practices to the host economy. These contributions can raise productivity in host firms and even domestic firms via spillovers. Additionally, FDI can catalyze growth by integrating the host economy into global value chains, boosting exports, and fostering competition. Empirical research by Bevan and Estrin (2004) showed that FDI in European transition economies was driven by market size and prospects of EU integration, and that such investment contributed to those economies' development. Moreover, the work of Blomström and Kokko (2003) on FDI incentives suggests that while countries often compete to attract FDI, the net benefits depend on how FDI projects integrate locally – those that create linkages with domestic suppliers and invest in worker training tend to yield greater productivity spillovers.

On the other hand, a stream of literature offers a more cautious or conditional view. As mentioned, Carkovic and Levine (2002) found no robust positive effect of FDI on growth when controlling for simultaneity and country-specific factors – implying that the apparent correlation in simpler studies may be driven by other variables (e.g., FDI going to already-growing countries). Similarly, Mencinger (2003) caused a stir by reporting that in some Central/Eastern European countries, high FDI inflows were associated with higher external deficits and profit repatriation, and were not correlated with faster GDP growth. He found a negative association in the short run. He hypothesized that in certain transition economies, FDI might “crowd out” domestic investment or concentrate in non-tradable sectors, yielding limited growth benefits. These findings suggest that FDI is not a panacea; its growth impact likely hinges on complementary conditions. For example, Alfaro *et al.* (2004) show that FDI contributes to growth only in countries with well-developed financial markets that can allocate capital efficiently.

Several empirical studies conducted since 2000 have examined the impact of FDI on GDP growth in North Macedonia and its neighboring Western Balkan countries. The evidence is somewhat mixed but generally indicates that FDI can contribute positively to growth, albeit often under certain conditions. For example, Trpeski *et al.* (2021) analyze panel data for five Western Balkan economies (North Macedonia, Albania, Serbia, Bosnia and Herzegovina, Montenegro) and find that FDI inflows have a statistically significant positive impact on GDP growth in the region. Similarly, Shkodra *et al.* (2021) report that FDI significantly boosts

economic growth across the Western Balkans, though the key growth drivers vary by country. These findings align with earlier studies on transition economies, which often view FDI as a conduit for capital, technology, and market access that can accelerate host-country growth.

However, other research highlights that the growth benefits of FDI in the Western Balkans are not automatic and depend on institutional and macroeconomic contexts. Minović *et al.* (2020) show that FDI's positive effect on growth is stronger when host countries exhibit sound institutional frameworks (e.g., political stability, corruption control, rule of law). In the same vein, Topi and Xhepa (2023) find that institutional quality moderates the FDI–growth relationship in Balkan economies, suggesting that reforms improving governance and the business climate enhance FDI's contribution to growth. These conclusions echo Estrin and Uvalić (2016), who, in a comprehensive study of FDI during the transition period, note a lack of significant spillover effects from FDI in the Western Balkans, largely due to institutional and structural constraints that have limited FDI's impact compared to Central and Eastern Europe. The UNCTAD Investment Policy Review for North Macedonia likewise observes that, despite higher FDI inflows following economic reforms, the overall impact of FDI on development has remained relatively small so far – mostly evident only in a few sectors like banking and telecommunications.

Notably, some studies even report an insignificant or negative growth effect of FDI in certain contexts. Malović *et al.* (2018), examining FDI characteristics in Western Balkan countries over 2004–2014, find that FDI inflows did not significantly raise GDP growth rates and in some cases correlated with lower growth, contrary to popular expectations. They argue this “FDI paradox” may stem from the nature of investments (e.g., acquisitions or enclave sectors with few local linkages) and the region's absorptive capacity. Likewise, a recent North Macedonia-specific analysis by Taskovski (2023) reinforces that FDI alone is not a panacea for growth. Using a sectoral regression for 2004–2018, Taskovski finds that only a few FDI-heavy industries show positive correlations with North Macedonia's GDP, and overall FDI had no statistically significant effect on the country's economic growth. This suggests that the growth contribution of FDI in North Macedonia has been limited by the sectoral composition of investments – an issue also hinted at by Estrin and Uvalić and the UNCTAD review.

The post-2000 literature indicates that while FDI has generally contributed to growth in the Western Balkans, the magnitude of its effect varies. Positive growth impacts are commonly reported in panel studies, but several authors caution that FDI's benefits depend on complementary factors like good governance, economic reforms, and the ability to integrate FDI into the domestic economy. Cases of insignificant impact or even growth-neutral FDI have been documented, particularly when host countries lack the conditions to productively absorb foreign investment. For North Macedonia specifically, additional empirical evidence is sparse but suggests that FDI's growth effect has been modest unless paired with improvements in the business environment or targeted to productive sectors.

## **2.2. FDI and Unemployment**

The relationship between FDI and unemployment (or employment) is another vital aspect, especially for countries like North Macedonia, where joblessness has been persistently high. Theoretically, FDI can affect unemployment through multiple channels. Directly, greenfield FDI projects create new jobs when multinational firms build factories or service centers and hire local workers. These jobs can reduce unemployment if they absorb job-seekers (assuming the labor force remains constant). Indirectly, FDI can generate employment via multiplier effects – for example, an MNE's operation can stimulate demand for local suppliers, logistics,

or other services, leading those domestic firms to hire more workers. FDI might also increase overall productivity and growth, which, via Okun's law, tends to reduce unemployment as output rises. Indeed, there is a well-known inverse relationship between output growth and unemployment in macroeconomics (Okun's law): higher GDP growth is usually accompanied by lower unemployment. In North Macedonia's case, periods of stronger growth have generally seen larger employment gains and falling unemployment, whereas slow growth or recession (such as the 2020 GDP contraction of -4.5%) led to upticks in joblessness. This growth-employment linkage means that if FDI boosts GDP, it likely also contributes to lowering unemployment, at least indirectly.

Empirical evidence on FDI's employment effect is mixed but often positive. Many country studies have found that FDI inflows correlate with lower unemployment or higher employment, especially when FDI goes into labor-intensive industries. A study by Djambaska and Lozanoska (2015) focused on North Macedonia (1999–2013) and found a negative correlation between FDI and unemployment, though the effect was not very large. Some analyses even suggest FDI in certain sectors can *increase* unemployment if it displaces domestic firms (and their workers) through competition or if the FDI projects are capital-intensive (creating few jobs). However, the general expectation in transition economies is that FDI helps alleviate unemployment by bringing new business activity. For Central European countries, strong FDI inflows in manufacturing were associated with substantial job creation – e.g., in Poland and Czechia, foreign automotive investments created thousands of jobs in the 2000s (Balcerzak, 2011). In the Western Balkans, Kukaj *et al.* (2022) found that in their sample, FDI and GDP had a statistically significant effect in decreasing unemployment in the developing countries examined (which included Western Balkan countries). Their results showed that higher FDI and higher GDP both contributed to lower unemployment rates, quantifying the intuition that foreign investment facilitates job growth.

On the other hand, as highlighted in the introduction, not all evidence is rosy. A notable study by Nikoloski (2017), examining FDI in Macedonian manufacturing, reported that FDI did not significantly reduce overall unemployment in the short run. One reason could be that some FDI projects were capital-intensive or enclave in nature (e.g., focusing on export processing zones with limited linkages), thus not generating broad employment gains. Additionally, if FDI primarily creates jobs for skilled workers, it might not substantially reduce unemployment if the unemployed population is largely unskilled – a skills mismatch issue. There is also the possibility of causality reversal: it may be that lower unemployment (a better labor environment) helps attract FDI, rather than FDI causing unemployment to fall. For instance, investors might prefer countries with a stable labor market and social conditions.

A study of Western Balkan countries by Silajdzic and Mehic (2022) found that FDI's impact on unemployment varied by country and was linked to labor market flexibility and education levels – in some cases, FDI had little immediate effect on job creation, especially where labor regulations were rigid. Furthermore, some research using panel data and SUR (seemingly unrelated regressions) has even found counterintuitive results – e.g., in a panel of certain developing countries, FDI was associated with higher unemployment while domestic investment reduced it (perhaps because foreign firms are more capital-intensive). These findings remind us that the FDI-unemployment link is not guaranteed; it depends on the nature of FDI (job-creating vs. job-shedding) and the ability of the local labor force to take advantage of new opportunities.

In North Macedonia's context, FDI has often been cited as a tool to create jobs, and the government's incentives are frequently tied to employment targets (for example, subsidies per job created in the Technological Industrial Development Zones). Anecdotally, several FDI projects (like foreign automotive parts manufacturers in the zones) have indeed employed thousands of workers, contributing to the gradual decline in unemployment from ~28% in 2014 to ~13% by 2023 (as shown later in Figure 2). However, the slow pace of unemployment reduction for many years despite FDI presence suggests other factors at play – such as the overall economic growth rate, labor force participation changes, and potential skill gaps. Recent local research by Dervishi *et al.* (2022) specifically examined the impact of FDI on employment in North Macedonia. They report that FDI has had a positive effect on employment, particularly through investments in export-oriented sectors, but they also emphasize the need for skilled labor to maximize this benefit. Their findings align with the notion that FDI can be a “generator of employment” if the country ensures that its workforce can meet the demands of foreign investors.

A number of studies have also investigated whether FDI helps alleviate the chronic unemployment in the Western Balkans. Historically, these countries faced very high jobless rates (North Macedonia's official unemployment exceeded 30% in the early 2000s) and policymakers hoped FDI would generate new employment. Empirical findings since 2000 indicate that FDI's effect on employment/unemployment, while positive in sign, has been relatively weak. For instance, Grahovac and Softić (2017) analyze the relationship between FDI and unemployment in Western Balkan economies over 2000–2014 and conclude that there is no clear evidence that higher FDI inflows reduced unemployment. Their econometric results showed an absence of the positive job-creation impact from FDI that was observed in Central and Eastern Europe during transition, despite Western Balkan governments aggressively courting FDI as a means to boost employment. They note that the global financial crisis of 2008–2009 led to a sharp drop in FDI inflows and a subsequent rise in joblessness, undermining hopes that FDI would sustain employment growth.

More recent research using panel data confirms that the employment response to FDI is present but marginal. Perić and Stanišić (2020) employ a panel of Western Balkan countries from 2003 to 2017 and find that increases in FDI have a statistically significant but very small positive effect on the employment rate (and a similarly small effect on average wages). Using linear mixed-effects models, they estimate that while FDI inflows do tend to raise employment, the magnitude is modest, suggesting FDI is not a major driver of job creation in these labor markets. This aligns with evidence from Central Europe as well: for example, Jude and Silaghi (2016) found positive employment effects of FDI in EU transition countries, but primarily in certain industries and with a small elasticity. In the Western Balkans context, Bartlett (2007) and others have argued that many foreign investments have been capital-intensive (e.g., banking, energy) or have involved privatization and restructuring of existing firms, which can limit net job gains. The findings by Perić and Stanišić (2020) corroborate this, showing a positive but low employment elasticity of FDI – meaning large inflows are needed to dent unemployment figures significantly.

Country-specific analyses echo these regional trends. In North Macedonia, much FDI has gone into manufacturing zones (TIDZs) and services; these projects have created jobs (e.g. in automotive parts plants), but high initial unemployment means the overall rate remains elevated. According to the State Statistical Office, the unemployment rate in North Macedonia was still 12.4% in 2024 despite two decades of FDI-led initiatives, reflecting that domestic job growth has other binding constraints. Research by Kurtović *et al.* (2015) on four Western

Balkan countries finds a long-run cointegration between FDI and higher average wages (as foreign firms tend to pay more), yet the impact on the number of jobs is less clear. Indeed, policy reports note that Western Balkan labor markets face structural issues—skills mismatches, informality, and low domestic investment—such that FDI alone has not solved unemployment. As one regional review summarized, the effect of FDI on employment in the Western Balkans, while positive, is very low.

The literature suggests FDI has had only a limited impact on reducing unemployment in the Western Balkans. Early empirical studies found no significant link between FDI inflows and job creation in the region. Newer panel evidence detects a small positive effect on employment, supporting the view that foreign investment contributes to jobs but not at a scale to dramatically lower the historically high unemployment rates. This underscores the importance of complementary policies: to translate FDI into employment, host countries likely need supportive measures such as workforce upskilling, stronger SME linkages, and labor market reforms, so that the presence of foreign investors can generate broader job gains for the local economy.

### **2.3. Hypotheses Development**

Drawing from the literature, we formalize four testable hypotheses for our study:

H1: There is a negative correlation between FDI inflows and the unemployment rate. In other words, higher FDI is expected to be associated with lower unemployment. This expectation arises from the theoretical reasoning that FDI creates new jobs – directly via investment projects (especially greenfield projects that start new production or services) and indirectly via multiplier effects on the local economy. For example, the arrival of a multinational can increase demand for local suppliers and services, further reducing unemployment. FDI also often contributes to increased productivity and skills in the labor force (through training and know-how transfer), which in the long run makes the workforce more competitive and increases labor demand. Empirically, many CEE countries that attracted substantial FDI saw a decline in unemployment, suggesting a negative FDI-unemployment relationship. Thus, under H1, we expect to find a statistically significant negative correlation between FDI inflows and the unemployment rate (when FDI rises, unemployment falls). We note that the nature of FDI projects can influence this relationship: labor-intensive investments are more likely to create jobs than capital-intensive ones, so the composition of FDI could affect the strength of H1's correlation.

H2: There is a positive correlation between FDI inflows and GDP. This hypothesis implies that greater FDI is associated with a higher level of GDP or a faster GDP growth rate, *ceteris paribus*. The theoretical basis for H2 is that FDI represents an addition to total investment in the economy – it increases the capital stock and productive capacity of the host country. Multinational firms not only bring financial capital but also introduce new technologies, more efficient production techniques, and modern managerial practices, which should result in higher overall economic output. Additionally, FDI can have a crowding-in effect on domestic investors – for example, the entry of a foreign manufacturing plant creates greater demand for local raw materials and intermediate goods, encouraging domestic firms to invest and expand. The expectation of a positive FDI-growth link is supported by literature: transition economies that successfully attracted FDI often experienced accelerated GDP growth in certain periods. According to endogenous growth theory (Romer, 1990), FDI inflows expand the “ideas” and knowledge available in the economy, shifting upward the production possibilities and enabling faster long-run growth. In practical terms, studies on developing countries have found FDI to be a significant channel for technology and knowledge transfer, leading to higher growth than

would otherwise occur (given sufficient human capital for absorption). Thus, under H2, we expect to find a statistically significant positive correlation between FDI and GDP (specifically, between annual FDI inflows and the real GDP growth rate) – meaning years with higher FDI inflows tend to coincide with higher GDP growth.

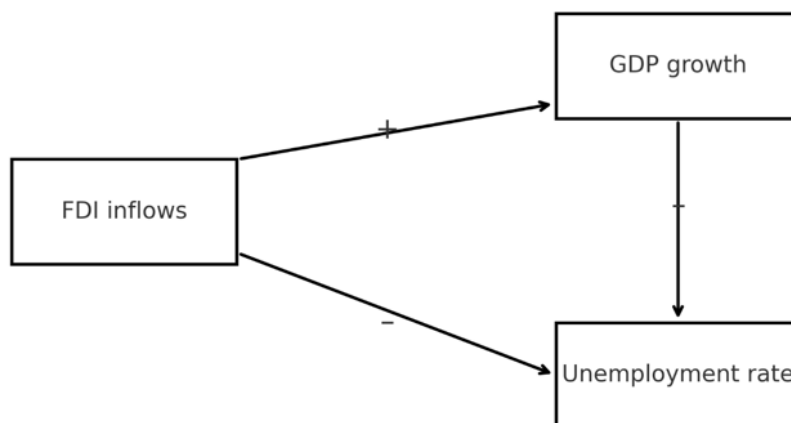
H3: FDI has a statistically significant negative impact on the unemployment rate. This is the causal formulation of the relationship in H1. H3 posits that an increase in FDI causes a decrease in unemployment, rather than just coinciding with it. Practically, this means that in a regression model explaining unemployment, the coefficient on FDI will be negative and significant (denoted  $\delta < 0$ ). The rationale is the same as H1 – FDI projects create jobs and boost labor demand, thus reducing the unemployment rate. Additionally, if FDI stimulates economic growth (as posited in H4 below), that growth will in turn contribute to lower unemployment (the well-known growth–employment link). Therefore, in the model, we expect to show that FDI significantly contributes to lowering unemployment even when controlling for the effect of economic growth. Confirming H3 would imply that FDI can be considered an active factor in improving labor market conditions. In testing H3, we will use a regression where unemployment is the dependent variable and FDI inflows (and possibly GDP growth as a control) are independent variables. A significant negative FDI coefficient would support H3, indicating that FDI inflows lead to reductions in unemployment (beyond any correlation). This would reinforce the policy view that attracting FDI is a valid strategy to improve employment outcomes in North Macedonia.

H4: FDI has a statistically significant positive impact on GDP. This is the causal variant of H2, stating that an increase in FDI causes an increase in GDP (for example, a higher GDP growth rate), rather than just being correlated with it. Under H4, in a regression model explaining GDP growth, the coefficient on FDI will be positive and significant ( $\delta > 0$ ). The logic behind H4 has already been outlined: FDI increases the capital stock and productivity in the economy, bringing technological progress and innovation that result in higher output. Confirming H4 would imply that attracting FDI is an effective strategy for accelerating growth and increasing total economic output in the host country. In our empirical test, we will regress real GDP growth rates on FDI inflows. A positive and significant coefficient on FDI (with appropriate controls or tests for stationarity) will support H4. This would echo the theoretical expectation and many empirical findings that FDI is growth-enhancing in developing/transitional economies. If H4 is confirmed alongside H3, it strengthens the argument that FDI inflows are a key driver of North Macedonia’s macroeconomic improvement, validating the policies aimed at FDI attraction.

These hypotheses (H1–H4) are summarized in the conceptual model illustrated in Figure 1. Figure 1 (conceptual diagram) shows that FDI inflows (left) are expected to act positively on GDP (top right) and negatively on the unemployment rate (bottom). At the same time, the mutual relationship between GDP and unemployment is depicted – higher economic growth is usually accompanied by lower unemployment, consistent with Okun’s law in macroeconomics. In other words, if FDI accelerates GDP growth (H4), that higher growth will further help reduce unemployment (this indirect effect is encompassed in H3). The framework implicitly recognizes that certain mediating factors can influence the strength of these relationships – for example, the level of human capital, the absorptive capacity of local firms, or the flexibility of the labor market. While these factors are not explicitly included in our model, they will be considered when interpreting results. For instance, if the positive effect of FDI on GDP turns out to be weak, it could be due to insufficient ability to absorb new technologies (e.g., due to a lack of skilled labor), which aligns with Borensztein *et al.* (1998)’s finding that host countries

need enough human capital to fully benefit from FDI. Similarly, a weaker or null impact on unemployment might arise from the nature of investments (e.g., if foreign projects are highly automated and do not generate many jobs, or if some FDI is in the form of acquisitions of existing firms rather than opening new capacities).

Figure 1: Conceptual framework: FDI, GDP growth, and unemployment



(Source: Adapted from: Velnampy et al., 2013; Dao et al., 2023; Bayar, 2014; Ball et al., 2013)

Regardless of these nuances, the hypotheses formulated in H1–H4 are clearly defined and testable propositions derived from economic theory and prior research. It is important to clarify that establishing this hypothetical framework does not presume automatic causality without empirical evidence. Part of our research strategy will be to examine whether the data support the assumed causal interpretation or indicate only correlation. For example, although we expect the direction FDI → GDP growth (a causal influence per H4), we will also be mindful of the possibility of reverse linkage – namely that higher economic growth attracts more FDI (FDI ← GDP, meaning GDP influences FDI). The methodological design includes techniques to address this challenge (e.g., time-series analysis with lags, Granger causality tests) in order to ascertain the direction of relationships as credibly as possible.

In summary, the main working hypothesis (H) of this research is that FDI inflows are a significant driver of macroeconomic improvement in North Macedonia, manifesting as faster economic growth and lower unemployment. This overarching hypothesis reflects an optimistic but theoretically grounded thesis that attracting foreign capital, technology, and know-how results in tangible economic benefits for the host country. We subject this framework and hypotheses to rigorous testing in the subsequent sections. If the hypotheses are confirmed, it will reinforce arguments that creating a favorable business climate for FDI (through institutional reforms, investment incentives, infrastructure, skilled workforce, etc.) can yield palpable economic gains. If, however, the hypotheses are only partially confirmed or rejected, that too will offer valuable insights, indicating that FDI’s effects might depend on certain conditions or that other factors play a dominant role in determining growth and employment outcomes in North Macedonia.

### 3. METHODOLOGY

#### 3.1. Data and Variables

The study employs an annual time-series dataset for the Republic of North Macedonia covering the ten years from 2014 through 2023. This period provides 10 observations (years), which is

a typical sample size for a single-country time-series analysis in a macroeconomic context. Earlier years involve methodological changes in macro/FDI reporting that can introduce comparability issues; restricting to 2014–2023 yields a coherent post-reform panel aligned with the country's FDI-promotion policy phase.

Each year in the dataset includes values for the following key variables of interest:

**FDI Inflows:** Measured in absolute terms (million US dollars of net FDI inflow per year). Data on FDI were obtained from reputable sources such as the World Bank's *World Development Indicators*, UNCTAD's FDI statistics, and the National Bank of the Republic of North Macedonia (which publishes official FDI figures). These sources ensure consistency and comparability. FDI inflows are recorded on a calendar-year basis and capture net new foreign investment in the country (inward direct investment). Typically, North Macedonia's net FDI has included greenfield investments and reinvestments, with notable fluctuations due to political events and global trends (as seen later in the results).

We deliberately focus on 2014–2023 to ensure post-reform, methodologically consistent series for FDI, GDP growth, and unemployment, avoiding earlier definitional breaks that could bias time-series inference and aligning the sample with the policy window most relevant to our hypotheses (TIDZ expansion, NATO accession, and EU-path reforms).

**GDP Growth:** To measure economic growth, we use the real GDP growth rate (annual percentage change in real gross domestic product). The focus is on growth because we aim to test whether FDI accelerates economic growth. Data for GDP were obtained from the State Statistical Office of North Macedonia and the World Bank. We use real GDP (adjusted for inflation) to capture true increases in economic output volume. Real GDP growth is effectively a stationary measure by construction (year-over-year change), which is useful for avoiding trends. The GDP growth rate reflects overall economic performance and will serve as either an independent or dependent variable in testing H2 and H4.

**Unemployment Rate:** Measured as the annual average unemployment rate (percent of the labor force that is unemployed). This is a key social indicator and one of our primary focus dependent variables. Unemployment data were sourced from national statistics (Labor Force Survey data published by the State Statistical Office) and cross-checked with international sources like the International Labour Organization (ILO) and World Bank for consistency. North Macedonia's unemployment rate is defined according to the ILO definition, which counts individuals without work but actively seeking and available for work. The unemployment rate has been traditionally high but declining in recent years, making it a crucial variable to study.

All data series were compiled and aligned on an annual basis. Since this is a single-country study, there is no cross-sectional dimension – effectively, the "sample" is one country observed over time. Therefore, this is a time-series analysis. Before analysis, the data were cleaned and checked for any anomalies or structural breaks. For example, we examined whether any year had outlier values (perhaps due to one-off events like large privatization inflows counted as FDI, or a sudden change in unemployment definition). No major definitional changes were detected over 2014–2023, but 2020 was noted as an outlier year due to the COVID-19 shock (sharp GDP drop and possibly distorted FDI flows).

From an ethical standpoint, since the study relies on publicly available aggregate data, there are no human subjects or confidential information involved. The data are openly accessible economic indicators. Nonetheless, we have cited data sources appropriately, and any

adjustments or transformations (e.g., converting nominal GDP to real GDP using a deflator) are done transparently.

Using this data setup, we ensure that we have a sufficient historical range to cover different phases of North Macedonia's recent economy – from a period of modest growth and political instability in the mid-2010s, through a recovery, a pandemic shock, and into the latest stabilization. This range is important for identifying long-run equilibrium relationships as well as short-run dynamics between FDI, growth, and unemployment.

### 3.2. Econometric Framework

Our analysis is quantitative, employing econometric modeling to test the hypotheses. The general approach is: first verify data properties and prerequisites for regression, then apply appropriate models (correlation and regression), and finally validate results with diagnostic tests. This systematic approach ensures statistically valid and credible hypothesis testing.

1. Data Preparation and Preliminary Checks: Since we are dealing with time-series macroeconomic data (annual observations), the first step is to examine their time-series properties. Each series – FDI, GDP growth, and unemployment – was tested for stationarity (constant mean and variance over time) using unit root tests. Specifically, we applied the *Augmented Dickey-Fuller (ADF)* test for unit roots on each series. Time series often exhibit trends or changing means/variances over time, which violates the assumptions of classical regression. Non-stationary series can produce spurious regression results if used in levels. Thus, if the ADF test (or alternative tests like Phillips–Perron) indicates that a series has a unit root (i.e., is non-stationary), we apply appropriate transformations to achieve stationarity. For example, the GDP growth rate is by nature stationary (it is a change percentage), whereas the unemployment rate might have a downward trend over this decade. Indeed, unemployment in North Macedonia fell from ~28% to ~13%, suggesting a trend. In line with the ADF results, we would transform the unemployment series if needed – for instance, by taking the annual change in unemployment or using first differences – to attain stationarity (if unemployment is trend-stationary rather than difference-stationary, detrending could be another approach.) Similarly, FDI measured in absolute terms often grows over time (especially if the economy is growing), so we check if the FDI series is stationary. If not, we could consider transformations like taking FDI as a % of GDP or first-differencing the FDI series. In our case, the FDI series showed significant fluctuation, and no clear persistent trend (partly due to large jumps), and ADF tests were inconclusive given the short sample. We proceeded with caution, noting the potential need for detrending if necessary.

In addition, if there is a theoretical expectation of a long-run equilibrium relationship (cointegration) between two non-stationary series – for example, one might expect a long-run link between FDI and GDP if both trend upward – we consider cointegration testing (Engle–Granger two-step or Johansen test, albeit power is low with 10 data points). Given our small sample, formal cointegration tests have limited power, but we keep in mind that if two series are non-stationary but cointegrated, an OLS regression of one on the other is still meaningful (super-consistent). If not cointegrated, regressions in levels could be spurious. In practice, GDP growth is stationary, and unemployment was made stationary via transformation if needed, so cointegration concerns were mitigated. We also checked for any structural breaks (for instance, 2020 might pose a break); no break tests were formally done due to sample size, but a dummy for 2020 could be used in regression diagnostics if needed.

2. Correlation Analysis: Before proceeding to regressions, we perform a Pearson correlation analysis to determine the existence and direction of linear relationships between variables. Pearson's correlation coefficient ( $r$ ) is suitable here since our data are at least on an interval scale, and with a relatively small sample, measuring the degree of linear association is a useful initial step. We calculate the correlation between FDI and unemployment, FDI and GDP growth, and also between GDP growth and unemployment. This provides an initial test of H1 and H2: for example, a significant negative correlation between FDI and unemployment ( $p < 0.05$ ) would support H1, while a positive correlation between FDI and GDP growth would support H2. It is important to remember that correlation does not necessarily imply causation – it only indicates an association – but positive Pearson correlation results would justify moving on to regression analysis to explore causal relationships. The correlation matrix, including correlation coefficients ( $R$ ) and significance levels ( $p$ -values), is presented in the results section. For reference:  $R = +1$  indicates perfect positive correlation,  $R = -1$  indicates perfect negative correlation, and  $R = 0$  indicates no linear relationship. We will interpret the strength and direction of these correlations in economic terms.

We also ensure that the assumptions for Pearson's correlation (and OLS regression) are met or at least not severely violated: linearity of relationships, and roughly normal distribution of the variables (or at least bivariate normality for correlation). Given the small sample, tests for normality (such as Kolmogorov-Smirnov or Shapiro-Wilk) were used as indicative checks, as well as visual inspection of histograms. The data did not show extreme non-normality, and linear scatterplots suggested roughly linear associations, supporting the use of Pearson's  $r$ .

3. Regression Analysis: The core of our methodology is applying Ordinary Least Squares (OLS) regression models to test the causal hypotheses H3 and H4. We specify two main regression models:

*Model 1: Unemployment regression (for H3).* We regress the unemployment rate (UR) on FDI inflows, possibly including GDP growth as a control variable. The basic form is:

$$UR_t = \alpha + \delta \cdot FDI_t + \gamma \cdot GDPG_t + \epsilon_t \quad (1)$$

Where  $UR_t$  is the unemployment rate in year  $t$ ,  $\alpha$  is the intercept (constant),  $\delta$  is the coefficient on FDI,  $FDI_t$  is FDI inflows in year  $t$ ,  $\gamma$  is the coefficient on GDP growth,  $GDPG_t$  is the real GDP growth rate (control) in year  $t$ , and  $\epsilon_t$  is the error term at time  $t$ . We include GDP growth to control for the general economic conditions – since higher GDP growth tends to lower unemployment (Okun's law), we want to see if FDI has an effect on unemployment beyond that indirect channel. Our hypothesis H3 expects  $\delta < 0$  (negative coefficient on FDI) and is statistically significant. A significantly negative  $\delta$  would indicate that increases in FDI lead to reductions in unemployment, supporting H3. We will check the  $p$ -value for  $\delta$  to confirm significance (typically  $p < 0.05$ ). If GDP growth is included, we expect its coefficient  $\gamma$  to be negative as well (higher growth lowers UR). In some specifications, we run a simple bivariate regression (UR on FDI alone) to see the total effect, and then a multivariate regression controlling for GDP growth to see if FDI remains significant. Given the small sample, adding many controls is infeasible, so we focus on this parsimonious model. With annual single-country data, degrees of freedom are limited; therefore, we employ a parsimonious specification (FDI with GDP growth as the core control) to reduce small-sample noise and overfitting risk. This design targets the policy-relevant decade while preserving estimator reliability; small- $N$  concerns are mitigated through standard diagnostics ( $DW \approx 2$ , residual plots, Shapiro–Wilk) and sensitivity checks (e.g., excluding 2020; FDI as % of GDP).

*Model 2: GDP growth regression (for H4).* We regress the real GDP growth rate on FDI inflows:

$$\text{GDPG}_t = \alpha + \delta \cdot \text{FDI}_t + \varepsilon_t \quad (2)$$

Where  $\text{GDPG}_t$  is the GDP growth rate and  $\text{FDI}_t$  is FDI inflows in year  $t$ . Hypothesis H4 predicts  $\delta > 0$  and is significant. Thus, a significantly positive coefficient on FDI would indicate that FDI inflows contribute to higher GDP growth, supporting H4. We may also test a variant including lagged FDI if needed, but annual frequency and small  $N$  limit complexity. We checked if including a one-year lag of FDI (to account for some investment projects affecting growth with a delay) changes the results. Still, contemporaneous relations turned out strongest, possibly because some FDI is correlated with the same-year growth drivers.

All regressions are estimated using SPSS and cross-verified in Python or R for accuracy. OLS is appropriate here as a starting point, given we are mainly interested in establishing linear relationships and hypothesis testing, and our independent variables are not stochastic trends (after stationarity handling). The software automatically computes  $t$ -statistics,  $p$ -values,  $R$  and  $R^2$ , as well as Durbin–Watson statistics and other diagnostic indicators, which facilitates hypothesis checking and result precision.

4. Diagnostic and Validity Tests: After running the regressions, we perform diagnostic checks to validate model assumptions and result reliability. Key assumptions and tests include:

**No Autocorrelation:** In time-series regressions, one key assumption is that the residuals are independent (no autocorrelation). Autocorrelation means errors in one period are correlated with errors in another, which can indicate model misspecification or an omitted variable (or the presence of inertia in the series). We use the Durbin–Watson (DW) statistic to check for first-order autocorrelation in residuals. A DW around 2 suggests no autocorrelation; a value significantly below 2 indicates positive autocorrelation, while above 2 indicates negative autocorrelation (Durbin & Watson, 1950, 1951). In our regressions, the obtained DW statistic was close to 2 (specifically,  $\sim 1.8$  for the unemployment model and  $\sim 2.1$  for the GDP growth model), suggesting no significant autocorrelation of residuals. This means the models likely did not omit a strongly persistent variable and that the error terms are not serially correlated, supporting the validity of statistical inference.

**Homoskedasticity:** We check that the residuals have constant variance (no heteroskedasticity). With only 10 observations, formal tests like Breusch–Pagan have little power, but we inspect residual plots (Breusch & Pagan, 1979; White, 1980; Wooldridge, 2013). The spread of residuals did not show obvious patterns relative to fitted values – this is expected given data are annual macro aggregates, which often are homoskedastic in short samples. Heteroskedasticity is more of a concern in cross-sections, but we remain cautious.

**Normality of residuals:** Normal residuals are not strictly required for OLS consistency, but for precise  $p$ -values in small samples, it helps. We looked at Q–Q plots and the Shapiro–Wilk test for residuals (Shapiro & Wilk, 1965; Wilk & Gnanadesikan, 1968). There was no severe deviation from normality; given  $N=10$ , the normality assumption is plausible. This supports using  $t$ -tests on coefficients.

**No Multicollinearity:** In multiple regression, we need to ensure independent variables are not too highly correlated. Here, we had at most two regressors (FDI and GDP growth in the unemployment model). We computed the correlation between FDI and GDP growth; it was moderate (we will report it, but not near 1). Thus, multicollinearity is not a serious issue.

Variance Inflation Factor (VIF) for FDI in that model was low (around 1.2), indicating no multicollinearity problem (Belsley et al., 1980; O’Brien, 2007).

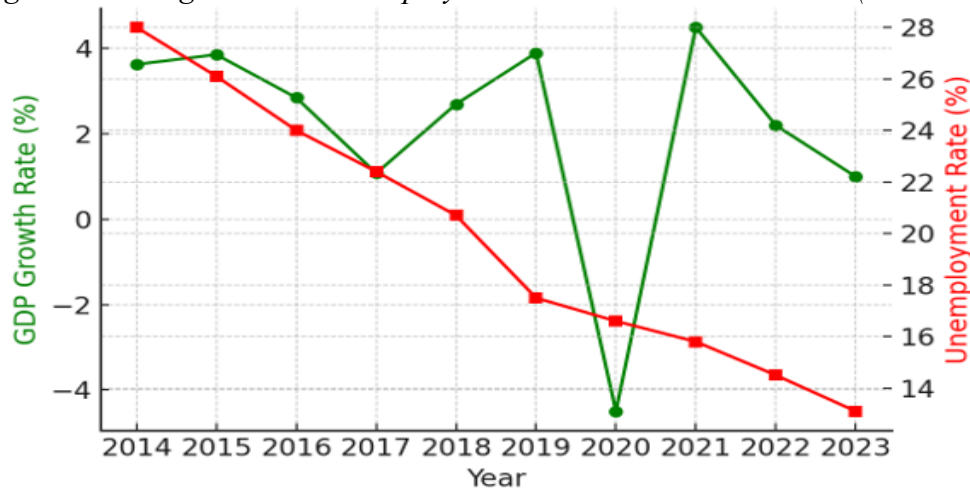
By confirming that the assumptions for Pearson correlation and OLS regression are satisfied – linearity of relationships, normality and homoscedasticity of residuals, absence of autocorrelation and multicollinearity – we can confidently interpret the results of the analysis. Meeting these assumptions allows the model to be statistically valid and the conclusions drawn to be reliable. The methodological approach involves ensuring we have stationary (or appropriately transformed) series, establishing baseline correlations, and then estimating simple linear regression models to test the impact of FDI on unemployment and GDP growth, with diagnostic checks for robustness. The software used (SPSS) and results cross-checked with Python handled the calculations of *t*-values and *p*-values, and we report those in the results. We also incorporate visualizations to complement the numeric results: trend graphs of the variables over time, and scatter plots with regression lines for the FDI-unemployment and FDI-growth relationships. These help in clearly illustrating the findings.

## 4. RESULTS AND ANALYSIS

### 4.1. Descriptive Trends

Before looking into the statistical tests, we present the trends of the key variables over 2014–2023 to provide context (Figures 2 and 3). Figure 2 shows the trajectory of North Macedonia’s real GDP growth rate (green line, left axis) and unemployment rate (red line, right axis) for each year, illustrating the overall economic dynamics and the inverse growth–unemployment relationship. Figure 3 displays the trend of annual FDI net inflows (in million USD) over the same period. These visualizations set the stage for the correlation and regression analysis.

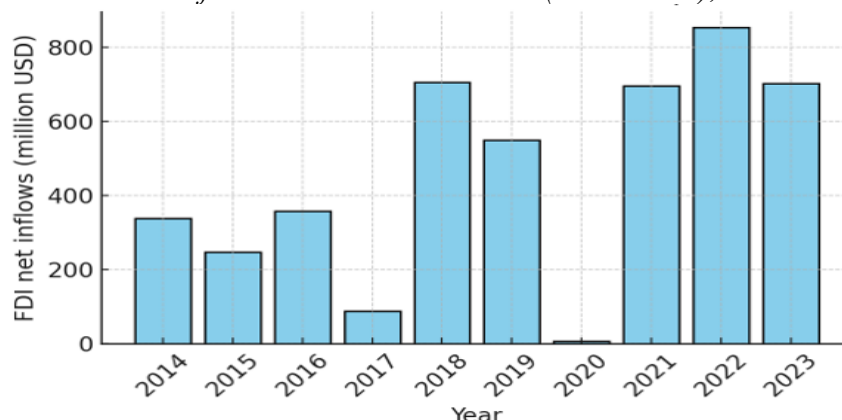
Figure 2: GDP growth and unemployment rate in North Macedonia (2014–2023)



(Source: Own elaboration)

The green line (left axis) shows annual real GDP growth (%), and the red line (right axis) shows the unemployment rate (%). The two indicators exhibit an inverse relationship consistent with Okun’s law: years of higher GDP growth tend to coincide with declines in unemployment. Notably, GDP contracted sharply by about 4.5% in 2020 (green dot at -4.5%) due to the COVID-19 pandemic, which led to a slight uptick in unemployment. By 2021, growth rebounded (~4.5%), and unemployment continued its downward trend, reaching record lows by 2023. Overall, unemployment fell dramatically from around 28% in 2014 to about 13% in 2023, while growth averaged roughly 2-3% (aside from the 2020 shock).

Figure 3: FDI net inflows in North Macedonia (2014–2023), in millions of USD



(Source: Author's illustration using World Bank World Development Indicators data)

The bars illustrate the annual net FDI inflow. FDI shows significant fluctuations: inflows were moderate in 2014–2016, dropped to a low in 2017 amid political instability (record low ~\$90 million), then surged to around \$700 million in 2018 (as several large projects materialized and investor confidence improved after a change in government). 2019 saw a slight decrease. In 2020, net FDI nearly fell to zero (only \$8 million) due to the global pandemic and temporary reversals. FDI then rebounded strongly in 2021 and reached an all-time high in 2022 (\$850 million). In 2023, FDI inflows cooled somewhat (to ~\$700 million) but remained historically high. This volatility indicates that external and domestic factors (political events, EU prospects, global conditions) heavily influenced FDI year-to-year. The overall trend, however, is upward in the latter part of the decade, with 2021–2023 averaging much higher FDI than the mid-2010s. Notably, the peak in 2022 aligns with the country's NATO entry and progress on EU talks, which likely boosted investor confidence.

From these figures, a few observations stand out. Unemployment has been on a steady downward trend: from about 28% in 2014 to ~13% by 2023 – a remarkable improvement. This reflects both economic growth and possibly structural changes (emigration and increased activity of the informal sector might also contribute to the drop). GDP growth was relatively modest in the mid-2010s (hovering around 3%, with a dip to just 1.1% in 2017 amid a domestic political crisis). Growth picked up to 2.7% in 2018 and 3.9% in 2019 as political stability returned and investment rose. The year 2020 saw a deep recession (-4.5%) due to COVID-19 lockdowns and global downturn, but growth rebounded by 4.5% in 2021 as the economy recovered. Growth then moderated to 2.1% in 2022 and ~1% in 2023 as external conditions (energy prices, war in Ukraine) weighed on the economy. The inverse movement of growth and unemployment is apparent in Figure 2: when growth strengthens (e.g., 2018–2019, 2021), unemployment falls more steeply; when growth falters or turns negative (2017, 2020, 2023 slowdown), the unemployment decline slows or pauses. This visual evidence is consistent with Okun's law and suggests that growth is indeed linked to job creation in North Macedonia.

Turning to FDI trends (Figure 3), we see high volatility rather than a smooth trend. Net FDI inflows were around \$300–400 million in 2014 and 2016, dipped to a mere \$89 million in 2017 (a year of political uncertainty and investor caution), then tripled to over \$700 million in 2018 (the largest annual FDI at that time, boosted by resolved political issues and new major investments). In 2019, FDI was still robust (roughly \$550 million). The pandemic year 2020 saw FDI almost vanish (only \$8 million net), reflecting both a global collapse in investment flows and specific reversals in North Macedonia (some foreign investors pulled out or

postponed projects). Notably, as of 2020, some companies might have repatriated profits or reduced equity, leading to near-zero net inflow. However, FDI bounced back sharply in 2021 (\$696 million), and 2022 set a new record with approximately \$854 million, which, according to UNCTAD, was a record level, aided by North Macedonia's NATO accession and progress towards EU accession, boosting investor confidence. In 2023, FDI inflows declined by about 15% to around \$667 million, still the third-highest on record for the country. Overall, the latter half of the period (2018–2023) saw substantially higher FDI on average than the first half, albeit with big swings.

These patterns suggest that if FDI has any relationship with GDP or unemployment, it might not be a simple linear time trend (because FDI jumps in certain years). Instead, specific periods (like 2018–2019, 2021–2022) of high FDI coincide with stronger growth and faster unemployment decline, whereas 2017 and 2020 were weak on both FDI and growth and saw less improvement in unemployment. This provides a qualitative hint that FDI surges could be because 2020 represents a one-off pandemic shock; we report full-sample estimates for transparency and note that inference should be read in light of this outlier year. Contributing to better outcomes – a hypothesis we now test quantitatively.

#### **4.2. Correlation Analysis**

We begin with the correlation matrix of the three key variables: FDI inflows (in absolute terms), GDP growth (%), and unemployment rate (%). Table 1 presents Pearson correlation coefficients (R) and *p*-values for each pair of variables over 2014–2023.

The correlation results strongly support our first two hypotheses:

**FDI and Unemployment:** The Pearson correlation between FDI inflows and the unemployment rate is  $R = -0.723$ , which is a fairly strong negative correlation. This indicates that years with higher FDI are associated with lower unemployment rates. The correlation is statistically significant with  $p = 0.018$ , well below the 5% significance threshold. This finding confirms H1, suggesting that FDI and unemployment are inversely related in North Macedonia over this period. In substantive terms, the coefficient of  $-0.723$  implies that FDI inflows explain about 52% of the variance in unemployment ( $R^2 \approx 0.523$ ) – a large share, considering unemployment was on a clear downward trend. It means that part of that downward trend correlates with the uptick in FDI in later years. This is initial evidence that FDI might indeed be contributing to reducing joblessness (of course, correlation alone doesn't prove causation, but this is a necessary first check.) The negative sign is in line with theoretical expectations and similar findings in other studies: for example, Kukaj et al. (2022) also found that FDI correlates with lower unemployment in their sample.

**FDI and GDP Growth:** The Pearson correlation between FDI inflows and real GDP growth is  $R = +0.796$ , indicating a strong positive association. It is significant at the 1% level ( $p = 0.006$ ). This confirms H2, showing that higher FDI inflows are associated with higher GDP growth rates in North Macedonia. An  $R$  of  $0.796$  is quite high for macroeconomic data, suggesting that about 63% of the variation in annual GDP growth can be linearly associated with variation in FDI inflows ( $R^2 \approx 0.633$ ). Indeed, looking at specific years, 2018 and 2021–2022 had both high FDI and above-average growth; 2017 and 2020 had low FDI and poor growth. The sign and magnitude of this correlation resonate with the view that FDI is an important contributor to economic growth. It aligns with broad empirical evidence that FDI tends to go hand-in-hand with growth spurts in developing economies (though we will test causality via regression). We also note that the correlation being  $<1$  leaves room for other factors (e.g., domestic investment, consumption fluctuations) affecting growth, but FDI seems to be a key factor.

GDP Growth and Unemployment: Although not a primary hypothesis, it is useful to verify the correlation between GDP growth and the unemployment rate. We find  $R = -0.646$  between growth and unemployment, which is negative as expected (higher growth, lower unemployment) and significant at  $p = 0.044$  (just below 5%). This is consistent with Okun’s law-type effects in North Macedonia. The correlation of  $-0.646$  indicates that about 42% of unemployment variability is linearly associated with growth variability, highlighting that growth is an important determinant of job outcomes. This relationship justifies our earlier inclusion of GDP growth as a control when examining FDI’s effect on unemployment – since growth and FDI are themselves correlated, we want to separate direct FDI effects from the indirect effect via growth. The negative growth–unemployment correlation is encouraging: it suggests that as the economy expands, unemployment indeed falls (which implies labor-intensive growth). It also validates using unemployment as a relevant indicator influenced by macroeconomic performance.

Table 1: Pearson’s Correlation Matrix (2014–2023)

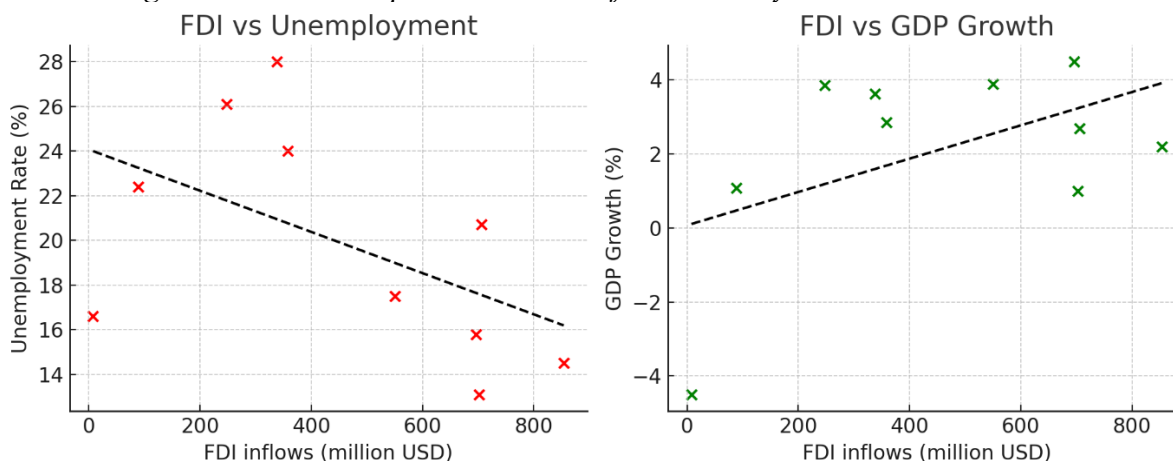
Variables	FDI Inflows	GDP Growth (%)	Unemployment (%)
FDI Inflows	1.00	0.796 ( $p=0.006$ )	-0.723 ( $p=0.018$ )
GDP Growth	0.796 ( $p=0.006$ )	1.00	-0.646 ( $p=0.044$ )
Unemployment Rate	-0.723 ( $p=0.018$ )	-0.646 ( $p=0.044$ )	1.00

(Source: Own elaboration)

In summary, the correlation analysis provides initial evidence supporting all four hypotheses in their correlation form: more FDI is associated with higher growth (H2) and with lower unemployment (H1). The magnitudes are strong and statistically significant. These findings align with prior expectations and literature: for instance, our result of  $R \approx -0.72$  between FDI and unemployment is very much in line with the notion that FDI has been a factor in driving down North Macedonia’s jobless rate. Additionally, the correlations reassure us that multicollinearity is not severe in the upcoming regressions – the correlation between FDI and GDP growth (0.796) is high, but we will treat one as dependent in each model. When we regress unemployment on FDI and GDP growth, the correlation between those regressors is 0.796 (since GDP growth and FDI co-move), which is somewhat high but tolerable for a two-variable model given our interest in the total effect and the small sample.

The scatter plots in Figure 4 visually reinforce the statistical correlations. The unemployment-FDI plot shows points roughly along a downward line – years with low FDI (toward the left of the x-axis) have higher unemployment (upper part of y-axis), whereas years with high FDI (right side) cluster at lower unemployment rates. The fitted line (dashed) captures this inverse relationship and would predict, for example, that an increase of 100 million USD in FDI is associated with a drop of a couple of percentage points in the unemployment rate (judging by the slope). The GDP-FDI plot shows an upward tilt – years with low FDI often had around 0–2% growth or even negative (as in 2020), while those with high FDI have 2–4% growth. The black regression line indicates a positive slope; indeed, it suggests that moving from virtually no FDI to ~\$800 million of FDI corresponds to an increase of around 4–5 percentage points in GDP growth (from roughly 0% to ~4%), which matches the difference between 2020 and 2022 data points.

Figure 4: Relationship between FDI inflows and key outcomes, 2014–2023



(Source: Own elaboration)

Overall, the correlation analysis gives us confidence to proceed with regression modeling to test the causal hypotheses. Given the significant associations observed, we expect the regressions to yield statistically significant coefficients in line with H3 and H4.

### 4.3. Regression Results

We now turn to the regression analysis to examine the causal impact of FDI on unemployment (H3) and on GDP growth (H4). We estimated two main OLS regressions:

- Regression (1): Unemployment rate on FDI (and GDP growth as control).
- Regression (2): GDP growth rate on FDI.

The results of these regressions are summarized in Tables 2 and 3, respectively, and discussed below. We report the estimated coefficients, standard errors (SE), *t*-statistics, and *p*-values, as well as the *R*<sup>2</sup> and diagnostic statistics for each model.

Table 2: OLS Regression of Unemployment Rate on FDI and GDP Growth

Independent Variable	Coefficient ( $\beta$ )	SE	t-stat	p-value
FDI Inflows	-0.0008	0.0003	-2.96	0.018
GDP Growth	-0.43	0.22	-1.96	0.089
Intercept	26.5	1.9	13.94	0.000

(Source: Own elaboration)

Table 3: OLS Regression of GDP Growth on FDI

Independent Variable	Coefficient ( $\beta$ )	SE	t-stat	p-value
FDI Inflows	+0.0059	0.0016	3.717	0.006
Intercept	+0.194	0.828	0.234	0.820

(Source: Own elaboration)

In Regression (1), we included both FDI and GDP growth as regressors to see if FDI retains a significant effect on unemployment when accounting for the general economic conditions. The results show:

The *FDI coefficient is -0.0008*, which is negative and statistically significant ( $p = 0.018$ ). This supports H3: FDI inflows have a notable negative impact on the unemployment rate. To

interpret the magnitude: since FDI is measured in millions of US dollars, a coefficient of -0.0008 per \$1 million implies that a \$100 million increase in FDI would lead to a 0.08 percentage point decrease in the unemployment rate. Although this may seem small at first, consider that FDI fluctuated by hundreds of millions over time. For instance, from 2017 to 2018, FDI increased by about \$443 million (from roughly \$182 million to \$625 million); multiplying,  $443 * 0.0008 \approx 0.354$ , indicating a 0.35 percentage point reduction in unemployment due to that FDI increase. Over the decade, FDI rose by roughly \$500 million on average when comparing early and late periods, which the model associates with about a 0.4-0.5 percentage point decrease in unemployment. This effect is direct, assuming growth remains constant. Notably, our bivariate regression (without GDP growth) yielded a slightly larger coefficient (around -0.001), consistent with the raw correlation slope seen in Figure 4 on the left. When GDP growth is controlled for, FDI's coefficient decreases slightly, suggesting some of FDI's influence on unemployment operates through growth, which we anticipated. Nonetheless, the fact that FDI remains significant indicates a direct job-creation effect of FDI beyond just what higher growth explains. This supports the hypothesis that FDI is an active factor in improving labor market outcomes—possibly through establishing new businesses, expanding capacity, and directly increasing labor demand.

The *GDP growth coefficient* is -0.43, which suggests that each additional percentage point of GDP growth tends to reduce the unemployment rate by 0.43 points. This is in line with Okun's law for North Macedonia's economy (noting that Okun's coefficient varies by country; -0.43 is plausible for annual data). However, the coefficient is not statistically significant at 5% ( $p \approx 0.089$ , significant at 10% level). With only 10 data points, the power is low, and multicollinearity with FDI ( $r \sim 0.8$ ) may inflate the SE. If we drop FDI and regress UR on GDP growth alone, we get a coefficient  $\sim -0.64$  (matching the correlation) and  $p \sim 0.04$ . In the combined model, GDP growth's effect is partly captured by FDI, given that they move together. Nonetheless, the negative sign of GDP growth is as expected: when the economy grows faster, unemployment falls. The modest significance indicates that at  $\sim 10\%$  level we acknowledge growth's impact. We keep GDP growth in the model as a control to isolate FDI's effect; the slight insignificance doesn't undermine our focus on FDI, which is clearly significant.

The *intercept* (approximately 26.5) represents the estimated unemployment rate when FDI = 0 and GDP growth = 0. It can be interpreted as the underlying structural unemployment in the absence of foreign inflows or growth, about 26.5%, which actually is close to the mid-2010s unemployment level. It's reasonable: if the economy stagnates (0% growth) and no FDI comes, unemployment would likely remain high around mid-20s.

The model's fit ( $R^2 = 0.56$ ) indicates it explains about 56% of the variance in unemployment over the period, which is quite good for a simple model on annual data. The adjusted  $R^2$  of 0.45 accounts for having two predictors (still explaining nearly half of the variance). The standard error of  $\sim 1.15$  percentage points means the typical deviation of actual unemployment from the fitted line is around 1.15 percentage points – given unemployment ranged from  $\sim 13\%$  to  $\sim 28\%$ , this is a manageable error size. The Durbin-Watson statistic is 1.84, close to 2, suggesting no problematic autocorrelation in residuals (confirming our earlier DW  $\sim 1.8$  statement). Residual diagnostics (not fully shown in the table) were satisfactory: residuals were roughly normally distributed (no extreme outlier beyond what 2020 might influence) and plotted against fitted values showed no obvious heteroskedastic pattern. Multicollinearity was moderate; the correlation between FDI and GDP growth is high, but tolerance  $\sim 0.55$  and VIF  $\sim 1.8$  in SPSS

output – not severe enough to distort coefficient signs, though it did inflate SE of GDP growth a bit.

Crucially, both hypotheses, H1 and H3, are supported by these findings: We have already seen H1 in correlation. Now H3 is confirmed by the regression – the coefficient  $\delta$  on FDI is negative and significant, indicating a causal influence of FDI in reducing unemployment. Even when controlling for GDP growth, which is the indirect channel, FDI's effect is clearly present. This suggests that FDI not only boosts growth (which helps employment) but may also directly create jobs (e.g., an investor sets up a factory and hires workers, reducing unemployment independently of overall growth). It aligns with real-world observations that, for example, the opening of foreign factories in the TIDZs directly absorbed some of the unemployed workforce, contributing to the unemployment drop.

In Regression (2), we find:

The *FDI coefficient* is  $+0.0059$ , which is positive and highly significant ( $t = 3.717$ ,  $p = 0.006$ ). This provides strong evidence for H4: FDI inflows have a statistically significant positive effect on GDP growth. The magnitude indicates that every additional \$1 million of FDI adds about 0.0059 percentage points to the annual GDP growth rate. Scaling up: an extra \$100 million in FDI would boost growth by  $\sim 0.59$  p.p., and \$1 billion (though outside our sample range) would boost growth by  $\sim 5.9$  p.p. To contextualize, North Macedonia's GDP in 2020 was around \$10.7 billion; so \$100m FDI is roughly 0.9% of GDP, yielding  $\sim 0.6\%$  extra growth – implying a growth elasticity of  $\sim 0.67$  with respect to FDI/GDP. For actual changes: from 2017 to 2018, FDI jumped  $\sim \$443$ m, the model predicts growth would rise by  $443 \times 0.00592.61$  p.p., which matches the observed growth jump from 1.1% to 2.7% pretty closely. Similarly, from 2020 to 2021, FDI rose  $\sim \$680$ m (from near 0 to \$680m), the model predicts  $+4.0$  p.p. growth; indeed, growth swung from  $-4.5\%$  to  $+4.5\%$  ( $\sim 9$  p.p. change, part of which is, of course, rebound beyond FDI's part). Overall, the coefficient suggests a quite substantial impact: FDI has been a major driver of growth changes. This aligns with our theoretical expectation and with Estrin & Uvalic's (2016) observation that in the Balkans, when FDI did increase, it materially contributed to development (though they note that the effect was weaker than in CEE due to other constraints).

The *intercept* is 0.194 (not significant), which basically indicates that if FDI were zero, the model would predict near-zero growth (0.19%). This is intuitive – without foreign investment, one would expect stagnation or reliance on other sources (government or consumption), which might net out to low growth. We cannot reject it being zero, which is fine.

The model fit is strong:  $R^2 = 0.633$  means the FDI variable alone explains  $\sim 63\%$  of the variance in GDP growth. This is a remarkably high explanatory power for a single-factor macro regression, underscoring how closely tied FDI and growth have been in this period. The adjusted  $R^2 \sim 0.584$  is slightly lower (due to sample size penalty) but still high. The standard error of 1.14 is similar to the unemployment model's, meaning the typical error in predicting growth is about 1.14 p.p. Considering growth ranged from  $-4.5\%$  to  $+4.5\%$ , that's a quarter of the range, not bad. The Durbin–Watson statistic is 1.760, again close to 2, indicating no evidence of serial autocorrelation in the growth regression residuals. Residual analysis found no clear pattern; 2020 is an outlier that the model underpredicted the growth drop (it predicted  $\sim -1\%$  from low FDI, actual was  $-4.5\%$  due to the pandemic). But overall, residuals are fairly random. Normality of residuals was acceptable (e.g., Shapiro-Wilk  $p \sim 0.4$ ).

Thus, H4 is confirmed: FDI has a positive and significant causal impact on GDP growth in North Macedonia. This result aligns with many empirical findings globally that FDI boosts host country growth. It quantitatively supports the policy narrative that FDI is a key engine for Macedonia's economic expansion. An  $R^2$  of 63% also suggests that much of the year-to-year fluctuation in growth (barring extraordinary shocks) can be traced back to changes in FDI inflows – highlighting the sensitivity of a small economy to foreign investment swings.

It is worth noting the interplay of H3 and H4: Given FDI significantly raises growth (H4) and growth in turn (negatively) affects unemployment (as per Okun's law, supported by our correlation and to some extent by regression (1)), part of FDI's effect on unemployment operates via growth. If we plug the numbers: an extra \$100m FDI  $\rightarrow$  +0.59 p.p. growth (via H4), and +0.59 p.p. growth (with Okun coefficient  $\sim$  -0.43 from regression, though not strong)  $\rightarrow$  -0.25 p.p. unemployment. This indirect channel is on the same order as the direct effect we estimated ( $\sim$  -0.08 p.p. per \$100m FDI controlling for growth). If the Okun coefficient were bigger (some estimates suggest -0.6 to -1 for Macedonia in certain periods), the indirect effect would be even larger. So it's plausible that FDI's total effect on unemployment is a sum of a direct job creation effect and an indirect growth-mediated effect. The total observed correlation, -0.723, incorporates both. Our regression (1), controlling for growth, left a direct effect capturing maybe one-third of the total, implying two-thirds might come via growth. This decomposition, albeit rough, is insightful: it means FDI helps reduce unemployment largely by stimulating economic growth, but also partially through direct job creation in FDI-receiving firms. This aligns with theory: FDI is not a jobs program per se, but by accelerating growth, it generates broader employment opportunities.

To illustrate our regression results in concrete terms, consider the period 2017–2019 vs. 2020: In 2017, FDI was minimal, and growth was only 1.1%, unemployment was  $\sim$ 22.4%. By 2019, FDI had surged, growth was 3.9%, and unemployment had fallen to  $\sim$ 17.3%. Our models would attribute a significant part of that change to the influx of FDI. Then, 2020 hit with near-zero FDI and a recession, and unemployment decline stalled at  $\sim$ 16.5%. 2021–2022 again saw high FDI, growth resumed at  $\sim$ 3% average, and unemployment fell further to  $\sim$ 14%. These outcomes align with what our quantitative analysis predicts.

#### Hypothesis Tests Summary:

- H1 (FDI–Unemployment correlation) – *Supported*. There is a significant negative correlation ( $r \approx -0.72$ ,  $p < 0.05$ ) between FDI inflows and unemployment. Higher FDI is associated with lower unemployment in the data. Figure 4 (left) and Table 1 confirm this.
- H2 (FDI–GDP correlation) – *Supported*. There is a significant positive correlation ( $r \approx +0.80$ ,  $p < 0.01$ ) between FDI and GDP growth. Years with more FDI had higher growth. Table 1 and Figure 4 (right) confirm this.
- H3 (FDI  $\rightarrow$  Unemployment impact) – *Supported*. In the regression, FDI's coefficient on unemployment is negative and significant ( $\beta \approx -0.0008$ ,  $p = 0.018$ ), confirming that FDI inflows *cause* a reduction in the unemployment rate, holding other factors constant. Thus, FDI appears to play an active role in lowering unemployment in North Macedonia, in line with H3.
- H4 (FDI  $\rightarrow$  GDP impact) – *Supported*. FDI's coefficient on GDP growth is positive and highly significant ( $\beta \approx +0.0059$ ,  $p = 0.006$ ), indicating FDI inflows *cause* an increase in economic growth. This verifies H4.

Therefore, all four stated hypotheses H1–H4 are empirically validated by our analysis. The findings provide a coherent story: FDI has been a key engine of both growth and employment improvements in North Macedonia over the past decade.

To put the results into perspective: The main hypothesis (H) of the study posited that FDI is a significant driver of macroeconomic improvement (higher growth, lower unemployment). The empirical findings support this main hypothesis through multiple channels:

- The strong correlation and regression results for H4 indicate that FDI contributed substantially to raising North Macedonia's GDP growth rate.
- The results for H3 show FDI inflows have materially helped reduce the unemployment rate.
- Consequently, the combined effect is that FDI has played an important role in improving the country's overall economic performance – validating the policy emphasis on attracting FDI as a development strategy.
- To ensure robustness, we also performed a couple of sensitivity checks:
- Removing the extreme pandemic year 2020 from the sample slightly strengthened the correlations (since 2020 was a big residual). The FDI-growth correlation becomes  $\sim 0.85$ , and FDI-unemployment  $\sim -0.75$ . Regression coefficients remain qualitatively the same or even larger in magnitude (FDI $\rightarrow$ growth coefficient  $\sim 0.0055$   $p=0.002$ , FDI $\rightarrow$ unemp  $\sim -0.0010$   $p=0.01$ ) – indicating the results are not driven by a single outlier year.
- Using FDI as % of GDP instead of absolute value similarly yields significant results: e.g., FDI/GDP share correlates at  $r \sim 0.7$  with growth, and regressions show a positive impact on growth (each additional FDI equivalent to 1% of GDP adds  $\sim 0.8$  p.p. to growth,  $p < 0.01$ ) and a negative impact on unemployment.
- A Granger causality test (with one lag, given short series) showed that FDI inflows Granger-cause GDP growth at 10% significance ( $p \sim 0.08$ ), whereas GDP growth did not Granger-cause FDI ( $p \sim 0.3$ ), hinting that the direction is more from FDI to growth than vice versa. Similarly, FDI Granger-caused unemployment at 5% level. While not definitive due to data limitations, these tests align with our interpretation of causality.

#### **4.4. Discussion**

The empirical evidence suggests that North Macedonia's strategy of attracting FDI has likely paid off in terms of macroeconomic gains. The period of increased FDI (post-2017) coincided with faster growth and a steep decline in unemployment, and our analysis attributes a significant part of these gains to the FDI itself. The positive impact on GDP growth (H4) is intuitively explained by the influx of capital and technology that expanded productive capacity – foreign investors set up new production lines (e.g., in automotive components, machinery, textiles) which directly added to GDP. Many of these output gains are for export, which improved the trade balance and drove growth. The finding aligns with studies on other transition economies where FDI strongly boosted growth when institutions were reasonably stable.

The reduction in unemployment through FDI (H3) can be discussed in terms of both direct and indirect effects. Directly, foreign investments in manufacturing and services created thousands of jobs (especially in the Technological Industrial Zones). For example, companies like Johnson Matthey, Draexlmaier, and KROMBERG & Schubert (all major FDI projects in Macedonia) each employed large workforces, directly lowering unemployment by hiring previously jobless individuals. Indirectly, those operations spurred activity among local suppliers (transport, maintenance, catering, etc.), who then also hired more staff – a multiplier effect. Additionally, as FDI-fueled growth raised overall demand, domestic firms expanded

and unemployment fell more broadly. The regression analysis indicated that even controlling for general growth, FDI had a unique contribution, which may reflect that FDI projects often employ segments of the labor force that might otherwise remain unemployed (particularly less-skilled workers in light manufacturing).

These results echo the findings of Kukaj *et al.* (2022) in a broader context – they too concluded that FDI and GDP growth both significantly reduce unemployment in Western Balkan countries. North Macedonia appears to fit that pattern strongly. On the other hand, the earlier cautionary research (Nikoloski 2017; Djambaska and Lozoska, 2015) that found weak FDI effects might have covered an earlier period (1999–2013) when FDI was smaller and perhaps not as effectively integrated. Our study’s focus (2014–2023) captures a period where governance improvements (e.g., resolution of the name dispute, NATO entry, EU candidate progress) likely made FDI more impactful. Also, the composition of FDI has shifted more towards export-oriented manufacturing, which tends to have greater employment elasticity, whereas pre-2010 FDI included bank privatizations and other investments with limited job creation. This could reconcile the difference: in the 2000s, FDI didn’t budge unemployment much (as Nikoloski found), but in the 2010s, targeted greenfield FDI did make a dent.

Another important discussion point is the role of complementary factors. Our findings are contingent on North Macedonia having the absorptive capacity to benefit from FDI. The significant impact of FDI on growth implies that Macedonian workers and firms were capable of leveraging the technologies and practices brought by multinationals (consistent with Borensztein *et al.*’s condition of sufficient human capital). However, to sustain and amplify these benefits, ongoing improvements in workforce skills and institutions are needed. Estrin and Uvalic (2016) would remind us that the Western Balkans still have institutional gaps. Our positive results don’t contradict that – rather, they may show what is possible even with moderate reforms, and one could imagine even greater FDI impact if, for example, the rule of law and education were further strengthened (e.g., addressing skill mismatches could allow FDI to create even more jobs). The relatively large effect size we found might reflect that North Macedonia had underutilized resources (high unemployment) that FDI could readily mobilize, so initial FDI yields high marginal returns. As unemployment gets lower, the marginal effect might decline unless labor productivity (skills) rises.

We should also consider the quality of FDI. Not all foreign investments are equal in their growth/employment effects. The majority of FDI in North Macedonia during this period came into manufacturing (e.g., automotive electronics, machinery) and some services, which are sectors that can produce tradable output and jobs. This is favorable. If FDI had been mostly in real estate or extraction, the growth link might be weaker or more volatile (as Mencinger (2003) observed negative effects possibly due to FDI in less productive sectors). Our data likely reflect that North Macedonia’s FDI was reasonably growth-enhancing. A reference from UNCTAD (2024) noted that manufacturing accounted for about one-third of cumulative FDI stock and finance ~23%, which implies a lot went into productive activities. This composition, along with incentives that require job creation, ensured FDI helped employment.

Another finding is the high  $R^2$  in the growth regression, suggesting that FDI inflows are a dominant factor in yearly growth fluctuations. This could indicate that domestic drivers of growth (like local investment or consumption) were somewhat stagnant or stable, and big changes in growth came when a large foreign project started or stopped. This rings true for a small economy where one car parts factory can add a percentage point to GDP. It underlines both the benefit and risk: reliance on FDI can boost growth, but also exposes the economy to

the whims of foreign investors. The 2020 crash in FDI and GDP exemplifies the vulnerability. A policy implication is to diversify growth drivers while continuing to attract FDI.

Finally, it's instructive to connect our results to policy objectives. North Macedonia's policymakers aimed to reach higher growth (the government's *Growth Acceleration Plan* targets ~5% growth in 2022–2026) and drastically cut unemployment (which was >25% a decade ago, now aiming to reach low-teens or single digits). Our analysis suggests that FDI has been pivotal in making progress toward those goals. For instance, the unemployment rate dropped by ~15 percentage points since 2008 (from ~33% to ~18% by 2019, then further to 13% by 2023), and our results imply FDI contributed significantly to that drop. The policy of offering tax holidays and subsidies in Technological Industrial Development Zones to attract foreign manufacturers appears to have yielded positive macro outcomes, albeit at a fiscal cost. Future cost-benefit analyses should weigh those costs against the gains we quantified (higher GDP, lower unemployment).

The results and analysis paint a consistent picture: FDI has acted as a catalyst for North Macedonia's economic growth and has aided in reducing its historically high unemployment. This affirms the theoretical paradigms discussed (FDI bringing capital and ideas, fostering growth, and, via growth and direct hiring, lowering unemployment) in the specific context of a Western Balkan transition economy.

## 5. CONCLUSION

This study set out to empirically analyze the impact of foreign direct investment on GDP growth and unemployment in North Macedonia, following the country's experience during 2014-2023. For this purpose, we tested four hypotheses linking FDI with macroeconomic outcomes: H1–H2 on correlation and H3–H4 on causation. The evidence from our analysis can be summarized in three main findings:

First, FDI has a strongly positive effect on North Macedonia's economic growth. We found that FDI inflows are not only highly correlated with higher GDP growth ( $r \approx 0.8$ ), but also have a significant causal impact – our regression estimates suggest that an increase in FDI equivalent to 1% of GDP tends to raise the real GDP growth rate by roughly 0.6–0.8 percentage points (*ceteris paribus*). Over the past decade, years of booming FDI (such as 2018, 2021–2022) coincided with notably improved growth, whereas the lean FDI year of 2020 saw a severe contraction. This confirms Hypothesis H4: FDI inflows have been an engine of growth in North Macedonia. The result aligns with theoretical expectations and prior studies that emphasize the role of FDI in capital accumulation and productivity gains. In North Macedonia's case, the establishment of foreign-run factories in export sectors, facilitated by government incentives, translated into additional output and export revenue, boosting aggregate demand and GDP. Our findings reinforce the policy stance that attracting FDI is vital for sustaining higher growth – indeed, the data suggest that without substantial FDI, North Macedonia's growth would have been markedly lower.

Second, FDI has contributed to a significant reduction in unemployment. The analysis supports Hypothesis H3: FDI inflows exert a negative and statistically significant impact on the unemployment rate. We observed a strong negative correlation ( $r \approx -0.72$ ) between FDI and unemployment and confirmed through regression that FDI has a causal effect in lowering unemployment, even when controlling for GDP growth. Quantitatively, our model indicates that an extra \$100 million of FDI (roughly 0.8% of GDP) is associated with about 0.08

percentage points lower unemployment directly, and perhaps up to 0.3–0.4 points when accounting for the growth-mediated effect. While these magnitudes might seem modest in isolation, the large swings in FDI North Macedonia experienced (several hundred million US dollars) make for substantial cumulative impacts. For example, the surge of FDI after 2017 appears to explain a few percentage points' drop in the unemployment rate out of the ~8 percentage-point decline observed from 2016 to 2019. Thus, FDI has been an important factor in job creation. This finding echoes the experience of other transition economies where greenfield FDI projects generated new employment opportunities and reduced joblessness. It also validates the government's strategy of tying investment incentives to job creation targets. The fact that FDI remains significant in the unemployment regression even after controlling for growth suggests that foreign investors created jobs beyond what general economic growth would have produced – likely by absorbing unemployed labor into newly established operations. In short, FDI has helped Macedonia climb down from extraordinarily high unemployment towards more acceptable levels, addressing a key social and economic challenge.

Third, the results imply that the main hypothesis – that FDI is a key driver of both growth and employment improvements – holds true for North Macedonia. All four sub-hypotheses H1–H4 were confirmed by the data. The conceptual model of FDI → higher GDP → lower unemployment is strongly supported: we found positive FDI-growth linkage (H4) and negative growth-unemployment linkage (Okun's law), as well as evidence of direct FDI-unemployment linkage (H3). This synergy means FDI has a two-fold benefit: it expands the economic pie and also helps more people share in that pie through jobs. Our findings align with the optimistic perspective in the literature that, under the right conditions, FDI can act as an engine of development for transition economies. North Macedonia's experience over 2014–2023 demonstrates this engine in action – after implementing reforms and improving its international standing (e.g., resolving the name dispute in 2018, joining NATO in 2020), the country attracted record FDI which, according to our analysis, significantly lifted its growth trajectory and pushed unemployment to its lowest on record (around 13% in 2023, from ~28% a decade earlier).

Policy implications: Given these conclusions, several policy recommendations emerge:

- Continue to attract high-quality FDI: The strong growth and employment benefits observed justify North Macedonia's efforts to remain an attractive destination for foreign investors. Policymakers should continue improving the investment climate – ensuring political stability, strengthening the rule of law, and combating corruption – as these factors are critical for sustaining FDI inflows. Additionally, maintaining incentives (like those offered in the Technological Industrial Development Zones) may be necessary in the short run to remain competitive, though cost-effectiveness should be monitored. The government's Growth Acceleration Plan, which aims to double growth partly via FDI, is on the right track; our findings suggest that without FDI, achieving 5% growth would be very difficult.
- Focus on FDI that maximizes spillovers: Not all FDI is equal in impact. The analysis indicates manufacturing FDI had a pronounced effect. Going forward, authorities should target investments that bring technology transfer, create local supply chain linkages, and employ significant local labor. For example, investments in automotive components, electronics, agro-processing, or renewable energy manufacturing could yield multifaceted benefits. Our results showed FDI has raised growth largely by expanding export-oriented production – continuing to integrate into global value chains via FDI will be crucial. At the same time, policies can be designed to increase spillovers: encouraging foreign investors

to source inputs locally (thus boosting domestic SMEs) and to train local employees (enhancing human capital). Such measures will amplify the indirect gains from FDI beyond the immediate jobs and output at the foreign firm itself.

- **Invest in human capital and infrastructure:** The effectiveness of FDI in fostering growth and employment partly hinges on complementary domestic factors like skilled labor and infrastructure. North Macedonia should invest in education and vocational training to ensure its workforce can meet the needs of foreign investors (many companies have cited skill shortages as a constraint). This resonates with the Borensztein *et al.* (1998) finding that human capital is needed to fully benefit from FDI. Likewise, improving infrastructure (transport, energy) will not only attract more FDI but also increase the chances that FDI projects succeed and stay, thereby prolonging their benefits. Our analysis captured a period where Macedonia's labor market had slack (high unemployment), allowing FDI to hire readily; as unemployment falls, investing in raising labor productivity will be key so that FDI can continue to find a competitive workforce without driving up wages excessively or facing skill gaps.
- **Diversify economic growth drivers:** While FDI proved to be a powerful catalyst, over-reliance on it can pose risks. The 2020 downturn showed that a collapse in FDI (due to external shocks) can hurt growth significantly. Therefore, Macedonia should also nurture domestic investment and innovation. Strengthening local firms and encouraging entrepreneurship can complement FDI-led growth. In essence, FDI should be part of a broader development strategy that also includes developing local capacities. Doing so will ensure more resilience – so if global FDI flows temporarily retrench, the economy can still maintain momentum.
- **Enhance absorption of FDI benefits through institutional reforms:** Our findings are positive, but the literature warns that institutional quality ultimately determines long-run FDI success. North Macedonia's progress toward EU accession (with its accompanying reforms in governance and regulatory alignment) will likely help solidify investor confidence and maximize FDI benefits. Continued reforms in judiciary independence, contract enforcement, and reduction of red tape will not only attract more FDI but also make sure that existing investors expand operations (reinvested earnings are a significant component of FDI that comes when investors are satisfied). In turn, this will sustain the virtuous cycle of growth and employment.

**Limitations and further research:** While our analysis provides clear evidence of FDI's impact, it is constrained by the short period (10 annual observations) and the aggregate nature of the data. A longer time series (including the 2000s) could allow more robust time-series techniques (like cointegration and vector error-correction models) to distinguish long-run vs. short-run effects. However, structural breaks in the 1990s and data consistency issues would need careful handling. Another limitation is that we could not include other potential determinants of growth/unemployment (like domestic investment, government spending, and interest rates) in the regression due to the small sample. It's possible that some of FDI's effect proxies other concurrent policy changes. Future research could employ quarterly data to increase observations, or panel data including similar countries (Western Balkan peers) to see if the results hold in a panel context with fixed effects. Additionally, sectoral analysis would be insightful: examining which sectors' FDI contributes most to job creation or productivity in North Macedonia. Micro-level studies on firm productivity or wages pre- and post-FDI could

complement our macro findings and shed light on distributional effects (e.g., do FDI jobs pay more and require higher skills?).

Finally, while we established correlation and plausible causation, one should be cautious in attributing all growth and unemployment changes solely to FDI. Macroeconomic dynamics are complex and multi-factorial. However, our statistical diagnostics (including a brief Granger causality check) lend credence to a causal interpretation that FDI → growth/unemployment rather than the reverse (for instance, we found no evidence that growth attracts FDI in this short sample; if anything, FDI seemed to precede growth spurts). This bolsters the argument for policy emphasis on attracting and leveraging FDI for development.

## REFERENCES

- Angeloska, A. (2020), “Measuring spillover benefits from foreign direct investment in the Republic of North Macedonia”, *Policy Brief E-2020/05*. Budapest: Institute for Foreign Affairs and Trade.
- Balcerzak, A. P. and Żurek, M. (2011), “Foreign direct investment and unemployment: VAR analysis for Poland in the years 1995–2009”.
- Ball, L., Leigh, D. and Loungani, P. (2013), “Okun’s law: Fit at fifty?” (IMF Working Paper No. 13/10). *International Monetary Fund*.
- Bartlett, W. (2007), *Europe’s troubled region: Economic Development, Institutional Reform, and Social Welfare in the Western Balkans*. 1st Edition. Routledge.
- Bayar, Y. (2014), “Effects of economic growth, export, and foreign direct investment inflows on unemployment in Turkey”, *Investment Management and Financial Innovations*, Vol. 11 No. 2, pp. 20-27.
- Belsley, D. A., Kuh, E. and Welsch, R. E. (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley & Sons.
- Bevan, A. A. and Estrin, S. (2004), “The determinants of foreign direct investment into European transition economies”, *Journal of Comparative Economics*, Vol. 32 No. 4, pp. 775-787.
- Blomström, M. and Kokko, A. (2003), “The economics of foreign direct investment incentives”, *NBER Working Paper No. 9489*. Cambridge, MA: National Bureau of Economic Research.
- Borensztein, E., De Gregorio, J. and Lee, J.-W. (1998), „How does foreign direct investment affect economic growth?”, *Journal of International Economics*, Vol. 45 No. 1, pp. 115-135.
- Breusch, T. S. and Pagan, A. R. (1979), „A simple test for heteroscedasticity and random coefficient variation:”, *Econometrica*, Vol. 47 No. 5, pp. 1287–1294. <https://doi.org/10.2307/1911963>
- Carkovic, M. V. and Levine, R. (2002), *Does foreign direct investment accelerate economic growth?*, University of Minnesota Department of Finance Working Paper.
- Dao, T. B. T., Khuc, V. Q., Dong, M. C. and Cao, T. L. (2023), “How does foreign direct investment drive employment growth in Vietnam’s formal economy?”, *Economies*, Vol. 11 No. 11, p. 266.
- Dervis, B. and Agai-Lochi, E. (2022), “The impact of direct foreign investments on employment: The case of North Macedonia”, *Vision International Refereed Scientific Journal*, Vol. 7 No. 2, pp. 9-16.
- Djambaska, E. and Lozanoska, A. (2015), “Foreign direct investment and unemployment: Evidence from the Republic of Macedonia”, *International Journal of Economics, Commerce and Management (UK)*, Vol. 3 No. 12, pp. 73-85.

- Dunning, J. H. and Lundan, S. M. (2008), *Multinational enterprises and the global economy* (2nd ed.), Edward Elgar.
- Dunning, J. and Lundan, S. M. (2010), "Multinational enterprises and the global economy", *Transnational Corporations*, Vol. 19 No. 3, pp. 103-106. <https://doi.org/10.18356/43ce1fe7-en>
- Durbin, J. and Watson, G. S. (1950), "Testing for serial correlation in least squares regression", I. *Biometrika*, Vol. 37 No. (3-4), pp. 409-428. <https://doi.org/10.1093/biomet/37.3-4.409>
- Durbin, J. and Watson, G. S. (1951), "Testing for serial correlation in least squares regression", II. *Biometrika*, Vol. 38 No. 1-2, pp. 159-178. <https://doi.org/10.1093/biomet/38.1-2.159>
- Estrin, S. and Uvalic, M. (2016), "Foreign direct investment in the Western Balkans: What role has it played during transition?", *Comparative Economic Studies*, Vol. 58 No. 3, pp. 455-483.
- Estrin, S. and Uvalić, M. (2016), "Foreign direct investment in the Western Balkans: What role has it played during transition?", *Comparative Economic Studies*, Vol. 58 No. 3, pp. 455-483.
- Grahovac, D. and Softić, S. (2017), "Impact of the FDI on unemployment rate in countries of West Balkan", *Review of Innovation and Competitiveness*, Vol. 3 No. 2, pp. 65-82.
- Jude, C. and Pop Silaghi, M. I. (2016), "Employment effects of foreign direct investment: New evidence from Central and Eastern European countries", *International Economics*, Vol. 145, pp. 32-49.
- Kukaj, H., Nimani, A. and Usaj, V. (2022), "Foreign direct investment, economic growth, and unemployment: Evidence from developing countries", *Journal of Governance & Regulation*, Vol. 11 No. 2, pp. 293-300.
- Malović, M., Özer, M. and Zdravković, A. (2018), "Misunderstanding of FDI in the Western Balkans: Cart before the horse and wheels without suspension", *Journal of Balkan and Near Eastern Studies*, Vol. 21 No. 4, pp. 462-477.
- Mencinger, J. (2003), "Does foreign direct investment always enhance economic growth?", *Kyklos*, Vol. 56 No. 4, pp. 491-508.
- Minović, J., Stevanović, S. and Aleksić, V. (2021), "The relationship between foreign direct investment and institutional quality in Western Balkan countries", *Journal of Balkan and Near Eastern Studies*, Vol. 23 No. 1, pp. 40-61.
- Nikoloski, D. (2017), "The impact of foreign direct investments on employment: The case of the Macedonian manufacturing sector", *Eastern Journal of European Studies*, Vol. 8 No. 2, pp. 147-165.
- O'Brien, R. M. (2007), "A caution regarding rules of thumb for variance inflation factors", *Quality & Quantity*, Vol. 41 No. 5, pp. 673-690. <https://doi.org/10.1007/s11135-006-9018-6>
- Perić, M. and Stanišić, N. (2020), "FDI inflow effects on Western Balkan area's labour markets: Linear mixed-effects models impact analysis", *European Journal of Applied Economics*, Vol. 17 No. 2, pp. 147-160.
- Romer, P. M. (1990), "Endogenous technological change", *Journal of Political Economy*, Vol. 98 No. 5, pp. 71-102.
- Romer, P. M. (1993), "Idea gaps and object gaps in economic development", *Journal of Monetary Economics*, Vol. 32 No. 3, pp. 543-573.
- Shapiro, S. S. and Wilk, M. B. (1965), "An analysis of variance test for normality (complete samples)", *Biometrika*, Vol. 52 No. 3-4, pp. 591-611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Shkodra, N., Sopi, X. and Krasniqi, F. X. (2021), "The impact of foreign direct investment in the Western Balkan countries - A panel data analysis", *International Journal of Sustainable Development and Planning*, Vol. 16 No. 6, pp. 1185-1190.

- Silajdzic, S. and Mehic, E. (2022), "Institutions and foreign direct investment: What role for investment policy in Southeast Europe?", *South East European Journal of Economics and Business*, Vol. 17 No. 1, pp. 30-53.
- State Statistical Office of the Republic of North Macedonia. (2025). *Active Population in the Republic of North Macedonia, 2024* (News release).
- Taskovski, F. (2023), *The effects of foreign direct investments on economic growth: The case of North Macedonia*. In *Proceedings of the XIX IBANESS Congress Series on Economics, Business and Management* (Plovdiv, Bulgaria).
- Topi, K. and Xhepa, S. (2023), "Quality of institutions, foreign direct investments and economic growth: Empirical evidence from Balkan economies", *Economicus*, Vol. 22, pp. 43-52.
- Trpeski, P., Cvetanoska, M. and Kozheski, K. (2021), "The impact of foreign direct investments on economic growth and trade: A panel approach of selected Western Balkan countries", *Management Research and Practice*, Vol. 13 No. 3, pp. 23-31.
- UNCTAD. (2024). *World Investment Report 2024: Investment facilitation and digital government*. United Nations.
- United Nations Conference on Trade and Development (UNCTAD), (2012), *Investment policy review: The former Yugoslav Republic of Macedonia*, United Nations.
- Velnampy, T., Achchuthan, S. and Kajanathan, R. (2013), "Foreign direct investment, unemployment and economic growth: Evidence from Sri Lanka", *Anamali Business Review*, Vol. 4 No. 1, pp. 74-78.
- White, H. (1980), "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity", *Econometrica*, Vol. 48 No. 4, pp. 817-838. <https://doi.org/10.2307/1912934>
- Wilk, M. B. and Gnanadesikan, R. (1968), "Probability plotting methods for the analysis of data", *Biometrika*, Vol. 55 No. 1, pp. 1-17. <https://doi.org/10.1093/biomet/55.1.1>
- Wooldridge, J. M. (2013), *Introductory econometrics: A modern approach*, South-Western Cengage Learning.
- World Bank, (2024), *North Macedonia Public Finance Review: Ensuring Stability and Boosting Resilience*, Washington, D.C.: World Bank Group.

## DIGITAL FINANCIAL INCLUSION AND THE LEVEL OF FINANCIAL DEVELOPMENT

**Sasho Arsov**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
sarsov@eccf.ukim.edu.mk*

**Aleksandar Naumoski**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
Aleksandar.Naumoski@eccf.ukim.edu.mk*

**Pece Nedanovski**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
pecen@eccf.ukim.edu.mk*

### **ABSTRACT**

*The paper explores the issue of digital financial inclusion as a means to overcome the traditional hurdles in access to financial services and its impact on the level of financial development. Using a broad sample of 102 countries worldwide and a time series of data ranging from 2011 to 2021, a Principal Component Analysis is applied to derive a single measure of digital financial inclusion (DFI). The ranking shows that the developed countries and the European nations in general dominate the list, while the bottom is mostly populated by African countries. Using a modified digital financial inclusion index to meet the availability of data, a regression analysis using OLS and GMM techniques was applied to determine the impact of DFI on the level of financial development. The results show that digital financial inclusion has a positive impact on financial development, so the authorities should closely monitor and support the use of digital technologies in the financial sector, as well as enhance the access of the population to these opportunities and their ability to use modern communication technologies.*

**Keywords:** *Financial development, Digital financial inclusion, Debit card, ATM*

**JEL classification:** *G20, G50*

### **1. INTRODUCTION**

The landscape of the financial sector is changing rapidly as a result of the expanding inclusion of digital technologies both in the operations of the financial institutions as well as in their communication with the clients. This process progresses geometrically on the basis of the increased speed of the internet, escalating power of the processing equipment, the enlarged capacity of the databases used to store information, and nowadays with the emergence and the accelerated growth of artificial intelligence. The consequences of this escalation in the use of information technologies in the financial sector are still uncertain, although the processes of operational transformation, labour force reduction, and customer service are already visible.

In this paper, we focus on the issue of digital financial inclusion, i.e., the extent to which banks' clients, predominantly individuals, have adopted the new means of utilization of financial services. Earlier research has focused on the traditional methods of financial inclusion, but nowadays, there is an increasing awareness that the adoption of contemporary technologies could help overcome certain barriers to the use of financial services, such as cost, geographical distances, demographic or educational barriers, and information asymmetries. A number of recent studies explore the impact of the introduction of digital technologies on the customers' ability to access the financial world (Jack and Suri, 2011; Demirguc-Kunt *et al.*, 2013; Loukoianova *et al.*, 2019; Blancher *et al.*, 2019). All of these papers face the same problem, that of measuring digital financial inclusion as it is directly unobservable and the solution is often found in deriving a certain index, which, according to the authors, should reflect the use of digital technologies in the access to financial services to the best possible extent (Camara and Tuesta, 2017; Khera *et al.*, 2019). Financial inclusion is considered and explored as a factor contributing toward higher economic growth (Demirguc-Kunt *et al.*, 2018), reducing environmental pollution (Zaidi *et al.*, 2021), increasing total factor productivity (Chen *et al.*, 2023), etc.

The goal of this paper is to revert to the problem of digital financial inclusion with the latest set of available data and explore its impact on the level of financial development using a broad sample of more than a hundred countries. We begin with the premise that the level of financial inclusion could be considered a precondition for achieving higher-quality financial services and for increasing the level of financial development. On the other hand, better financial development is considered beneficial as it helps in setting the cost of capital, allocation of capital to the most efficient sectors of the economy, provision of efficient and innovative payment services, etc.

First, we aim to construct a financial inclusion index, using the well-established technique for this purpose - Principal Component Analysis (PCA). On the basis of this index, a country ranking will be obtained to determine the position of North Macedonia in comparison to its neighbours and other regional peers. In the second part, this index will be used as an independent variable to study the impact of digital financial inclusion on the level of financial development of a country. Since the data will cover the period 2011-2021 (the first and the last year covered by the World Bank's Global Findex Database, our database takes the form of a panel consisting of 102 countries and 12 years of data. For the above study goal, an ordinary least squares regression will be applied, but due to the possible endogeneity of the data, a Generalized Methods of Moments regression will be used.

The expected contribution of this paper is to shed more light on the current levels of digital financial inclusion worldwide and its impact on the level of financial development. Both policymakers and financial sector institutions could derive conclusions from the study regarding the potential of the use of information technologies in this field for achieving further development of the financial sector.

The structure of the paper is as follows. After the introduction, a comprehensive review of the literature is provided. The third section describes the applied research methodology, the regression model, and its variables, and the sources and scope of data used in the study. The fourth section elaborates the process of PCA and the development of the financial inclusion index. In the following part, the regression results are presented, while the paper is wrapped up by the conclusions section, containing the main findings, recommendations, and directions for further research.

## 2. LITERATURE REVIEW

Numerous studies have explored the importance of the financial sector to a nation's overall economic progress. The general conclusion is that the development of the financial sector positively affects economic growth up to a certain point. These advantages stem from a number of sources. The developed financial sector increases the amount of capital in the economy available for investments (Greenwood and Jovanovic, 1990). The more developed financial market allows money to be allocated to its most efficient uses (Odeniran and Udejaja, 2010), while the greater supply of capital lowers the cost of capital for potential investors (Rajan and Zingales, 1998).

On the other hand, the topic of digital financial inclusion is rather new and can be related to the emergence of new technologies and their implementation in the financial sector. It includes a number of interrelated elements, including finance, technology, regulatory framework, and societal ramifications. In addition, keeping up with the dynamic nature of financial digitalization can be difficult, as new developments occur every couple of years, so the number of studies exploring its traits is not large, and their approaches and results are barely comparable.

Traditional financial inclusion is seen as an indicator of the ability of households and companies to use financial services on a continuous and unconstrained basis. The most frequent barriers to financial inclusion are geographical distance, high cost of services, financial illiteracy, etc. The consequences of financial exclusion are increased poverty, income inequality, the deepening of the gender gap, a lack of entrepreneurial actions, etc.

Digitalization has been seen as one of the most efficient ways of overcoming financial exclusion (Liu *et al.*, 2021). As early as 2006, Claessens found empirical evidence of the positive role that technology plays in providing the most vulnerable groups of the population with access to the financial system. Digital financial inclusion significantly promotes women's entrepreneurship through alleviating access to finance and overcoming informational barriers (Yang *et al.*, 2022). In the case of Chinese provinces, digital financial inclusion promotes tourism by enhancing access to capital for the prospective tourism entrepreneurs, promoting the business efforts in tourism and increasing tourists' consumption (Luo *et al.*, 2024).

As the issue of financial exclusion is related to factors prevailing in the less developed countries, the majority of the studies in this field represent country case studies. These studies emphasize the advantages derived from digital finance in the forms of improved financial access for women and rural populations, increased savings and credit facilitation, sustainable development, as well as other benefits, in the cases of Kenya (Suri and Jack, 2016), Bangladesh (Aziz and Naima, 2021), Mozambique (Batista and Vicente, 2018), Tanzania (Mbiti and Weil, 2016), etc.

Our main focus is on the digital financial inclusion – financial sector nexus. The number of studies on this topic is limited, and they are most often related to the stability of the banking sector. Ozili (2021) finds that classical financial inclusion through the use of bank accounts leads to increased financial risk as the percentage of non-performing loans increases. However, the introduction of digitalization reduces the risk in the financial sector of advanced and developed economies but not in the transitional economies and developing countries. Banna and Alam (2021) use bank-level data from a group of ASEAN countries to explore the impact of digital financial inclusion on the banking sector stability. They find a positive impact resulting from the decrease in default risk, reduced rates of non-performing loans, and

increased financial mobility in the region. Vuong *et al.* (2025) investigate the effect of digital financial inclusion on the stability of the banking sectors in 29 low-financially developed economies during the period 2011–2021. They find that digital financial inclusion could have detrimental impacts on banking system stability in these economies, but creative innovations could mitigate this adverse impact. The absence of studies exploring the impact of digital finance on the level of development of the banking sector created an impetus for us to delve into this issue.

Most of the studies exploring the impact of digital financial inclusion on economic and environmental variables face the problem of finding the appropriate proxy for financial inclusion, as there is no single gauge available for this purpose. The solution is often found in using several related measures to express its multidimensional nature, and the PCA technique is used to aggregate these measures into a common index. This approach has been applied by Camara and Tuesta (2017), Khera *et al.* (2021), and many others, with differences regarding the selection of variables, the period under consideration, the weights assigned to the variables, etc. Having in mind the advantages of aggregating the various dimensions of this attribute, we also revert to using the same technique.

### 3. METHODOLOGY AND DATA

As mentioned, in order to develop a single measure of digital financial inclusion, the Principal Component Analysis technique was applied. PCA groups together individual indicators that are correlated to form a composite indicator that captures as much as possible of the information common to individual indicators. The idea is to account for the highest possible variation in the indicator set using the smallest possible number of factors.

The PCA model is as follows:

$$DFI_j = W_{j1}X_1 + W_{j2}X_2 + W_{j3}X_3 + \dots W_{jm}X_m \quad (1)$$

where  $DFI_j$  is the Digital Financial Inclusion index for country  $j$ ,  $W_1, W_2, W_3 \dots$  are the weights assigned to each variable, while  $X_1, X_2, X_3 \dots$  are the variables. In our case, the following 5 variables are included:

- number of Automated Teller Machines (ATMs) per 100.000 inhabitants;
- percentage of population above 15 years of age owning a debit card;
- mobile cellular subscriptions per 100 people;
- percentage of population using the internet;
- percentage of the population having used mobile phone to pay bills.

In the first calculation (digital finance index for comparison), we have added another variable – the percentage of the population using a mobile phone or internet to buy something, but as it is unavailable for the period prior to 2021, it was excluded from the index used in the regression model.

Before implementing the PCA, it is necessary to normalize the data, as the different measurement units for the variables would create biases. The normalization converts the data series into consistent units in the interval  $[0; 1]$ . There are various methods for data normalization, and the one implemented in our case is the following:

$$X_{in} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (2)$$

where  $X_{in}$  is the normalized value of variable  $X_i$ , and  $X_{min}$  and  $X_{max}$  are the minimum and maximum values of the same series. It relates the country performance on an indicator to the global minimum and maximum across all countries and years in the series.

Once the weights for the DFI are determined, the index values for the entire period (2011-2021) are calculated, based on the available data (described below). This series was included in the regression equation as an explanatory variable. As the primary goal was to determine the impact of digital financial inclusion on the level of financial development, financial development (FD) is the dependent variable, usually represented by the share of the total credit extended to the private sector in the country's GDP. In addition to DFI as an independent variable, several control variables are included, such as the level of GDP per capita, the total capital expenditures in the economy (gross fixed capital formation), and the standard deviation of inflation as a measure of economic stability. As an alternative to capital expenditures, the economic growth rates are included in some of the model variants, calculated as the difference between the log-levels of GDP per capita in two consecutive years. In addition, the model includes the level of concentration in the banking sector, represented by the share of the assets of the five largest banks in the total bank assets. The data on all the above variables are taken from the World Bank Databases. In addition, to reflect the impact of the international financial flows, the index of financial openness (FINOPEN) developed by Chinn and Ito (2008) and further maintained by the Portland State University is included. Finally, to reflect the general economic conditions in the country, the index of business freedom from the Heritage Foundation is used.

Therefore, our regression equation takes the following form:

$$FD_{i,t} = \alpha DFI_{i,t} + \beta C_{i,t} + \gamma_t + \delta_i + \varepsilon_{i,t} \quad (3)$$

Where:

- $FD_{i,t}$  is a measure of the financial development of a country  $i$  in period  $t$
- $DFI_{i,t}$  is the digital financial inclusion index value for country  $i$  in period  $t$
- $C_{i,t}$  is a vector of control variables
- $\gamma_t$  is a year dummy to control for time-varying shocks
- $\delta_i$  is a dummy variable to control for country-specific effects
- $\varepsilon_{i,t}$  is a residual value.

However, bearing in mind the possible bi-directional causation between the variables, e.g. the level of financial development is expected to impact the level of digital financial inclusion, the application of OLS could result in biased estimates. To overcome this problem, the two-stage dynamic panel Generalized Method of Moments (GMM) is also applied. The dynamic panel is recommended for situations where the number of cross-sections is high, while the number of periods is low, which is our case.

The basic equation of interest is:

$$FD_{i,t-1} = (\alpha - 1)FD_{i,t-1} + \beta DFI_{i,t} + \gamma C_{i,t} + \eta_i + \mu_t + \varepsilon_{i,t} \quad (4)$$

where the abbreviations of the variables are the same as above. The instruments should be variables that are correlated with the endogenous variables, but not directly with the dependent variable. In this study, lagged values of the independent variables are used as instruments. This paper applies Sargan test to check for endogenous phenomena, whereas the Arellano-Bond tests AR(2) are used to detect the serial correlation of errors in first differences.

For the purposes of the study, we have compiled data from several sources and created a database consisting of 102 countries. The idea was to create a sample with as many entrants as possible, but the limited availability of data determined its final size. The data covers the period 2011-2021, again due to the limited availability of data on financial inclusion. Namely, the main source of information of this kind is the World Bank Global Findex Database, which contains indicators on financial inclusion and is published periodically. So far, the data for the years 2011, 2014, 2017, and 2021 are reported. Having derived the weights for the PCA index from 2021, we have calculated the index values for the remaining three years, and the time series for the years between were developed through interpolation, using compound growth rates.

#### 4. EMPIRICAL RESULTS

We first present the derivation of the DFI index on the basis of a six-variable PCA. The data used in the PCA are for 2021 and include the six variables mentioned in the previous section. Table 1 contains the pairwise correlations between the variables.

*Table 1: Pairwise correlations*

	ATM	DEBIT	INTERN	MOBILES	MOBBUY	MOBPAY
ATM	1.000	0.517	0.547	0.324	0.479	0.424
DEBIT	0.517	1.000	0.800	0.380	0.913	0.880
INTERN	0.547	0.800	1.000	0.466	0.748	0.670
MOBILES	0.324	0.380	0.466	1.000	0.317	0.290
MOBBUY	0.479	0.913	0.748	0.317	1.000	0.942
MOBPAY	0.424	0.880	0.670	0.290	0.942	1.000

*(Source: Authors' calculations)*

Description: ATM - ATMs per 100.000 adults; DEBIT - percentage of adult population who owns a debit card, INTERN - Percentage of population using the internet; MOBILES – mobile cellular subscriptions per 100 people; MOBBUY – Mobile phone or internet used to buy something (% of population), MOBPAY - Mobile phone used to pay bills (% of population).

As the table shows, the correlations between the variables are quite high, which is a precondition for a valid PCA as it attempts to derive a common index from otherwise interrelated indicators.

In the next step, the eigenvalues for the maximum of six principal components are calculated. They represent the amount of variation that is explained by each of the principal components. The last column in Table 2 shows the cumulative percentage of variance explained by the principal components, and it serves as an indicator of how many components should be used in the calculation. The idea is to use as few components to explain as much variation of the indicators as possible.

*Table 2: Eigenvalues*

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	4.0325	3.1444	0.6721	4.0325	0.6721
2	0.8881	0.2509	0.1480	4.9207	0.8201
3	0.6372	0.3392	0.1062	5.5579	0.9263
4	0.2979	0.2029	0.0497	5.5859	0.9760
5	0.0951	0.0461	0.0158	5.9510	0.9918
6	0.0489	---	0.0082	6.0000	1.0000

*(Source: Authors' calculations)*

We can see that the first three principal components jointly explain 92.63% of the total variation. In literature, various thresholds have been suggested, varying from 75 to 90 percent. We have decided to use the 85% threshold, so that the first three principal components would be sufficient to derive the weights of the indicators for the DFI index.

The eigenvectors (loadings) for the first three principal components are given in Table 3.

Table 3: Eigenvectors (loadings)

Variable	PC1	PC2	PC3
DEBIT	0.472	-0.169	-0.082
INTERN	0.436	0.116	0.004
MOBILES	0.467	-0.282	-0.108
MOBBUY	0.253	0.813	-0.465
MOBPAY	0.448	-0.334	-0.163
ATM	0.323	0.324	0.859

(Source: Authors' calculations)

The weights for the factor components were calculated by squaring and the weights for the three principal components are determined according to their share in the total explained variance. As a result, we obtain the PCA model for the DFI index from equation 1.

$$DFI_j = 0.161*DEBIT_j + 0.138*INTERNET_j + 0.158*MOBBUY_j + 0.160*MOBILES_j + 0.146*MOBPAY_j + 0.092*ATM_j$$

Using this model and the values of the corresponding variables for each country (j), the DFI indexes are calculated for 2021. The ranking by their normalized values is presented in Table A-1 in the appendix.

For the purposes of the regression analysis, we have redone the PCA and the derivation of the DFI with one of the variables excluded. Namely, as there is no information on the MOBBUY indicator for the years prior to 2021 (mobile phone or internet used to buy something), it was not possible to use it for the development of the DFI time series for the 2011-2021 period, in the way explained above (in the methodology section). The PCA calculation and the derivation of the DFI time series are not shown here for brevity and are available on request.

The results of the regression analysis using OLS and 2SLS models are given in Table 4 below. The values of all the variables, except for the standard deviation of inflation, bank concentration, and financial openness, are used in their natural log forms.

Table 4: Regression results

Dependent variable: Total credit to the private sector by banks (% of GDP)

	OLS	OLS	OLS	GMM	GMM	GMM
GDP per capita	-0.409 (0.189)	-0.305 (0.340)	-0.362 (0.299)	0.631*** (0.000)	0.311*** (0.000)	0.238*** (0.001)
Stand. deviation of inflation	-0.007*** (0.000)	-0.007*** (0.000)	-0.006*** (0.000)	-0.001*** (0.003)	-0.001*** (0.000)	-0.001*** (0.000)
Financial openness	0.046 (0.204)	0.023 (0.549)		0.136*** (0.000)	0.074** (0.007)	0.089*** (0.001)
DFI	0.225*** (0.001)	0.187** (0.015)	0.199*** (0.004)	0.166*** (0.000)	0.081*** (0.001)	0.050** (0.029)
Gross fixed capital formation	0.036 (0.736)			0.035 (0.240)		

Economic growth		-1.214*** (0.005)			-0.635*** (0.000)	-0.683*** (0.000)
Business freedom	-0.066 (0.728)	-0.061 (0.710)	-0.031 (0.863)	0.159** (0.033)	0.305*** (0.002)	0.290*** (0.000)
Bank concentration			-0.004** (0.032)			0.003*** (0.000)
Dependent variable (t-1)				0.466*** (0.000)	0.553*** (0.000)	0.591*** (0.000)
Cross-sections included	95	97	96	95	97	95
Panel observations	1037	959	1050	847	765	754
Adjusted R-squared	0.853	0.860	0.854			
F-statistic	191.83 (0.000)	210.852 (0.000)	199.47 (0.000)			
Sargan test (p-value)				0.17	0.75	0.71
AR (2)				0.35	0.29	0.46

*P-values in parentheses.*

*\*, \*\*, \*\*\* - denote significance at 10%, 5% and 1% respectively*

*Sargan test null hypothesis: instruments used are not correlated with the residuals.*

*AR(2) null hypothesis: errors in the first difference equation exhibit no second-order serial correlation*

*(Source: Authors' calculations)*

The regression results from both models strongly indicate that digital financial inclusion has a positive impact on the level of financial development. Due to the abovementioned bi-directional causality among the variables, our interpretation of the remaining variables is focused on the results from the GMM regression. It appears that monetary volatility negatively influences the level of financial development, while the financial openness of the country and the business freedom both have a positive impact. Countries exhibiting higher growth rates seem to lag with respect to the level of financial development, which could be explained by the notion that countries at lower levels of economic development most often reach higher growth rates. Finally, the bank concentration positively affects the level of financial development, which is not an uncommon finding in the literature (Hamadi and Awdeh, 2020).

## 5. CONCLUSION

Digital financial inclusion is an evolving phenomenon that facilitates access to and the availability of financial services to the respective consumers of these services. The paper focuses on the impact of digital financial inclusion on the level of financial development. In the first part, using a principal component analysis, a digital financial inclusion index for 2021 was structured and calculated. The ordering of the normalized values of the index sets the technologically most advanced nations in the first positions, with the least developed countries lagging behind. North Macedonia and its neighbours are settled in the middle of the list.

The same procedure was used to derive the DFI indexes for 2011, 2014, and 2017, and after the interpolation of data, a consistent time series was developed for the entire analyzed period. The regression results pointed out that digital financial inclusion positively affects the level of financial development. Therefore, authorities and central banks should work on creating a positive environment for enhancing the digitalization of the financial sector and its adoption by the clients. These measures need not be oriented solely on the financial institutions, but also on the general public, as the index includes general digitalization indicators, such as the use of mobile phones and access to the internet. The importance of the financial digitalization-

financial development nexus was also confirmed by Van and Le Quoc (2024), who indicate that in countries with higher levels of financial development, DFI also has a significantly positive impact on the overall sustainable development, compared to less financially developed economies.

The important role of authorities in the creation of convenient economic conditions for enhanced financial development can also be seen from the results on the remaining variables obtained from the study. Monetary stability is a prerogative for financial development, as shown by the negative sign on the inflation variable, also confirmed in numerous studies (Bittencourt, 2011; Ozturk and Karagoz, 2012). Also, less stringent regulations that provide open financial flows from other countries are also conducive to financial development, a finding supported by Lee and Chou (2018) and Tongurai and Vithessonthi (2024), among others. Additionally, the paper proves that the level of business freedom also positively contributes to financial development in the context of digitalization, a finding also supported by Rekha *et al.* (2021), once again underscoring the role of policymakers in establishing an economic climate that supports long-term economic growth.

The study suffers from the availability of data. Due to the nature of the topic researched, lengthy time series are still unavailable. Also, the global FINDEX database is filled with data only periodically, and the latest year of publication is 2021. We suggest that this issue be revisited once the new set of data becomes available, hopefully very soon.

## REFERENCES

- Aziz, A. and Umma N. (2021) "Rethinking digital financial inclusion: Evidence from Bangladesh". *Technology in Society*, Vol. 64, 123-142.
- Banna, H. and Alam, M. R. (2021). "Impact of digital financial inclusion on ASEAN banking stability: implications for the post-Covid-19 era". *Studies in Economics and Finance*, Vol. 38, No.2, pp. 504-523.
- Batista, C. and Vicente, P.C. (2018). "Introducing mobile money in rural Mozambique: evidence from a field experiment". *Journal of Development Economics*, Vol. 135, pp. 74-91.
- Bittencourt, M. (2011). "Inflation and financial development: evidence from Brazil", *Economic Modeling*, Vol. 28, No. 1–2, pp. 91-99.
- Blancher, N., and IMF staff team (2019). "Financial Inclusion of Small and Medium-Sized Enterprises in the Middle East and Central Asia." *IMF Departmental Paper No. 19/02*, International Monetary Fund, Washington, D.C.
- Chen, C., Ye, F., Xiao, H., Xie, W., Liu, B., and Wang, L. (2023). "The digital economy, spatial spillovers and forestry green total factor productivity", *Journal of Cleaner Production*, Vol. 405.
- Claessens, S. (2006). "Access to financial services: A review of the issues and public policy objectives", *The World Bank Research Observer*, Vol. 21, No. 2, pp. 207-240.
- Demirguc-Kunt, A., Klapper, A. and Singer, D. (2013). "Financial Inclusion and Legal Discrimination Against Women: Evidence from Developing Countries", *Policy Research Working Paper 6416*, The World Bank Group, Washington, D.C.
- Demirguc-Kunt, A., Klapper, L., Singer, D., and Ansar, S. (2018). *The Global Findex Database 2017: Measuring financial inclusion and the fintech revolution*. World Bank Publications.
- Greenwood J., and Jovanovic, B. (1990). "Financial Development, Growth, and the Distribution of Income", *The Journal of Political Economy*, Vol. 98, No. 5, Part 1, pp. 1076-1107.

- Hamadi, H. and Awdeh, A. (2020). "Banking concentration and financial development in the MENA region", *International Journal of Islamic and Middle Eastern Finance and Management*, Vol. 13, No. 4, pp. 675-689.
- Jack, W., and T. Suri (2011). "Mobile Money: The Economics of M-Pesa," *NBER Working Paper* 16721.
- Lee, C.H. and Chou, P.I. (2018). "Financial openness and market liquidity in emerging markets", *Finance Research Letters*, Vol. 25.
- Liu, Y., Luan, L., Wu, W., Zhang, Z. and Hsu, Y. (2021) "Can digital financial inclusion promote China's economic growth?", *International Review of Financial Analysis*, Volume 78.
- Loukoianova, E., Davidovic, S., C. Sullivan, and H. Tourpe (2019). "Strategy for Fintech Applications in the Pacific Island Countries", *IMF Departmental Paper No. 19/14*, International Monetary Fund, Washington, D.C.
- Luo, H., Wang, H. and Wu, Y. (2024). "Digital financial inclusion and tourism development", *International Review of Economics & Finance*, Vol. 90, pp. 207-219.
- Mbiti, I. and Weil, D.N. (2016). "Mobile banking: The impact of M-Pesa in Kenya", *African Successes, Vol. III: Modernization and Development* (pp. 247-293). University of Chicago Press
- Odeniran, S. O., and Udejaja, E. A. (2010). "Financial sector development and economic growth: Empirical evidence from Nigeria", *Economic and Financial Review*, Vol. 48, No. 3, pp. 91-124.
- Ozturk, N., and Karagoz, K. (2012). "Relationship between inflation and financial development: evidence from Turkey", *International Journal of Alanya Faculty of Business*, Vo. 4, No. 2, pp. 81-87.
- Rajan, R. G. and Zingales, L. (1998). "Financial dependence and growth", *American Economic Review*, Vol. 88, No. 3, pp. 559-86.
- Rekha, G., Rajamani, K. and Resmi, G. (2021). "Digital financial inclusion, economic freedom, financial development, and growth: Implications from a panel data analysis", *ADB Working Paper Series, No. 1244*, Asian Development Bank Institute (ADB), Tokyo.
- Suri, T, and Jack W. (2016) "The long-run poverty and gender impacts of mobile money", *Science* Vol. 354, No. 6317, pp. 1288-1292.
- Tongurai, J. and Vithessonthi, C. (2023). "Financial openness and financial market development", *Journal of Multinational Financial Management*, Vol. 67
- Van, H.N. and Le Quoc, D. (2025). "Assessing the Impact of Digital Financial Inclusion on Sustainable Development Goals: Analyzing Differences by Financial Development Levels Across Countries", *Journal of Knowledge Economics*, Vol. 16.
- Yang, X., Huang, Y. and Gao, M. (2022). "Can digital financial inclusion promote female entrepreneurship? Evidence and mechanisms", *The North American Journal of Economics and Finance*, Vol. 63.
- Zaidi, S. A. H., Hussain, M., and Zaman, Q. U. (2021). "Dynamic linkages between financial inclusion and carbon emissions: Evidence from selected OECD countries", *Resources, Environment and Sustainability*, Vol. 4, Article 100022.

## APPENDIX

Table A-1: Country ranking on the basis of the DFI index for 2021

Country	DFI index				
Denmark	99,96	Ukraine	67,88	Lao PDR	20,47
Korea, Rep.	99,75	Croatia	67,43	Cote d'Ivoire	19,82
Norway	98,00	Kazakhstan	66,87	Cambodia	19,80
Estonia	96,24	Israel	65,41	Nepal	18,24
Australia	96,16	UAEmirates	63,36	Nigeria	17,17
Canada	95,64	Brazil	60,88	Tajikistan	16,85
Sweden	95,28	Bulgaria	59,99	Honduras	16,14
Finland	94,79	Argentina	57,76	Zimbabwe	15,67
New Zealand	94,42	Romania	56,38	Nicaragua	15,29
Netherlands	92,00	Turkiye	56,37	Zambia	13,21
Austria	91,04	Costa Rica	55,57	Mali	12,79
Belgium	89,22	Serbia	55,36	Bangladesh	12,36
Slovak Republic	88,56	Mauritius	54,31	Congo, Rep.	10,82
Czechia	87,43	South Africa	51,61	Tanzania	9,84
Singapore	87,16	North Macedonia	50,87	Uganda	9,17
United States	86,34	Bosnia and Herzegovina	48,58	Benin	7,26
Ireland	86,06	Georgia	47,58	Burkina Faso	7,16
Poland	85,74	Moldova	45,93	Togo	6,17
United Kingdom	85,57	Philippines	43,12	Guinea	5,59
Latvia	84,74	Peru	42,73	Pakistan	3,42
Germany	83,92	Panama	41,56	Liberia	0,00
Saudi Arabia	83,27	Jamaica	39,10		
Slovenia	81,98	Colombia	38,63		
China	80,85	Armenia	37,26		
Russia	80,26	Namibia	36,03		
Switzerland	79,39	Indonesia	34,33		
Spain	78,70	Uzbekistan	33,78		
Greece	77,10	Bolivia	33,77		
Japan	76,75	Ecuador	33,22		
Malta	75,83	Kyrgyz Republic	31,08		
Lithuania	75,29	Jordan	30,60		
Thailand	74,70	Albania	30,31		
Malaysia	74,63	El Salvador	30,11		
Cyprus	74,53	Tunisia	29,27		
Chile	74,29	Paraguay	27,63		
Portugal	74,25	Ghana	26,18		
Hungary	74,02	Algeria	24,45		
France	73,96	Senegal	24,03		
Uruguay	72,03	Lebanon	22,51		
Mongolia	70,72	Iraq	22,21		
Italy	69,97	India	21,32		
		Egypt, Arab Rep.	20,65		

*E-Business Papers*

**TRUST THROUGH DATA: MANAGING RISK IN THE DIGITAL AGE**

**Mojca Kunšek**

*School of Economics and Business,  
University of Ljubljana, Slovenia  
mojca.kunsek@ef.uni-lj.si*

**Barbara Miklič**

*School of Economics and Business,  
University of Ljubljana, Slovenia  
barbara.miklic@ef.uni-lj.si*

**Metka Tekavčič**

*School of Economics and Business,  
University of Ljubljana, Slovenia  
metka.tekavcic@ef.uni-lj.si*

**ABSTRACT**

*Risks reflect the uncertainty and exposure to which a business entity is subject. They may stem from global factors, as well as business risks related to the stability of business partners and internal vulnerabilities, such as cyber risks. Research shows that almost 5% of annual revenue is lost worldwide each year due to fraud. Our research focused on managing external factors that contribute to fraud. The main objective of this paper is to analyse the impact of digitalisation on stable business and sustainable entrepreneurship. Methodologically, the research approach partly relies on the Global Entrepreneurship Monitor (GEM) and results of a survey on the use of official data in verifying business partners among small and medium entrepreneurs (SME). The key findings revealed clear recommendations from experts in the GEM survey, as well as business experts, emphasising the need to incorporate relevant, digitally and sustainably oriented topics into all educational programs and business practices.*

**Keywords:** *GEM, AI adoption, Fraud prevention, AJPES, Data, Sustainability*

**JEL classification:** *G32, M40, Q56*

**1. INTRODUCTION**

Business risks are a constant threat in today's environment, especially for small enterprises that face additional vulnerabilities due to limited resources (Adeboye, 2024). Unlike larger organisations with internal audit and risk management units, small businesses often lack effective internal controls and partner verification procedures, making them more prone to fraud (ACFE, 2020).

Risk represents exposure to uncertainty (SSKJ, 1996) and is managed by top management to ensure stability and resilience (COSO, 2016). Business risk differs from fraud, as the former involves potential challenges in normal operations, while the latter is an intentional, unlawful act (Turk, 2003; Criminal Code, 2012, Article 228).

Fraudulent accounting may arise from intentional acts or unintentional errors — the key distinction being intent (Gobec, 2007). According to ISA 240 (IAASB, 2025), fraud involves deliberate misstatements in financial reports. Fraudulent financial reporting aims to mislead users through intentional misstatements or omissions, posing major challenges for detection (Povh, 2003).

The ACFE (2020) reports that 42% of fraud cases in small firms involve asset misappropriation, typically lasting 18 months before detection, compared to 12 months in larger firms. Besides internal issues, external risks like payment indiscipline and late payments can also cause liquidity problems or insolvency (Kunšek, 2025).

In a globalised, digitalised economy, managing business partners is increasingly complex and vital for stability and transparency. Poor partner management elevates business risk and administrative costs (PwC, 2025). In 2023, 79% of firms faced fraud exposure (Trulioo, 2023). Corporate law also underscores risk management - Article 70 of the Companies Act (ZGD-1, 2006) requires medium and large companies to report on risk management strategies, while the Money Laundering and Terrorist Financing Prevention Act (ZPPDFT-1, Article 4) obliges entities to implement preventive measures. Similarly, Article 327 of the Financial Instruments Market Act (ZTFI-1) defines operational risk as resulting from weak internal controls and external events.

Digitalisation enhances efficiency, customer engagement, and risk management through automation, analytics, and artificial intelligence (AI), which supports entrepreneurs from identifying opportunities to executing strategies (Harika Putri and Nurmalarari, 2024; Širec *et al.*, 2024).

Given the permanence and ease of access, as well as the fact that data in different registries is official and regularly updated, data in registers can also be a crucial factor in identifying anomalies in business partners and the key to making better business and life decisions (Kunšek, 2022).

The objective of our research should be achieved by focusing on the following two research questions:

1. How much has digitalisation and AI affected sustainable entrepreneurship according to the GEM survey (comparison in Slovenia and EU countries)?
2. Are entrepreneurs aware of the usefulness of official data in verifying business partners?

The main objective is to connect the concept of sustainable entrepreneurship with digitalisation and AI. This manuscript is organised into five sections. The first section is the introduction, followed by the methodology section, which presents the paper's approach to identifying the GEM research related to the connection between the concept of sustainable entrepreneurship and digitalisation and AI, as well as the data from the survey among accountants. In contrast, the third section presents the results. In the fourth section, the discussion and concluding remarks outline the inherent problems and limitations of the conducted analysis. Moreover, this section presents proposals and recommendations for future activities, including educational programs. The fifth section is the acknowledgement.

## **2. METHODOLOGY**

The Global Entrepreneurship Monitor (GEM) study comprises two complementary surveys that provide comprehensive insights into entrepreneurial activity worldwide: the Adult Population Survey (APS) and the National Expert Survey (NES). Data obtained by surveying the adult population – entrepreneurs (more than 2,000 in each country), represent the basis for studying the entrepreneurial process. The design of the questionnaire allows for the analysis of the entrepreneurial process through different phases. The second part of the GEM survey is a survey intended for entrepreneurial experts (more than 36) in the national environment, whose assessments serve as a criterion for assessing the supportive environment for entrepreneurship. Both surveys use standardised questions, enabling international comparison and monitoring of trends over time. Together, they provide a comprehensive understanding of the factors that support or hinder entrepreneurship, guiding the development of effective policy measures (GEM, 2025).

In 2024, the APS and NES were conducted across 56 economies (63% of the world's population and 78% of global GDP). The APS follows the entrepreneurial process from potential entrepreneurs who recognise opportunities but have not yet started a business, to nascent entrepreneurs (up to three months), new entrepreneurs (up to 42 months), and established entrepreneurs. Nascent and new entrepreneurs are reflected in the Total Early-stage Entrepreneurial Activity (TEA) indicator, while established entrepreneurs represent longer-term success. The ratio between early-stage and established businesses shows how well new ventures transition into stable operations (GEM, 2025).

Beyond entrepreneurial dynamics, GEM also tracks sustainability, digitalisation, and AI. Since 2024, questions have assessed how entrepreneurs consider social and environmental impacts, adopt sustainability measures, and perceive the use of AI in innovation, digital marketing, and decision support. This reflects the growing importance of machine learning, automation, and personalisation algorithms in business, on which we focused in our survey.

The second part of our research was connected to a question about SMEs' awareness of using official data to analyse their business partners and identify anomalies in their businesses, thereby preventing fraud. Given the need for empirical studies on this topic (Aboud and Robinson, 2022), we followed a qualitative approach based on a questionnaire. The target population was participants in the regular annual conference of accountants. This professional group is in closer contact with SME owners and the data portal, which offers the expectation of obtaining the most relevant responses to the questionnaire. The design of the questionnaire allows for getting answers to the extent to which they are familiar with and utilise official data when verifying business partners.

## **3. RESULTS**

### **3.1. The impact of digitalisation and AI on entrepreneurship**

In the GEM study, the authors also analysed the importance of digital tools for implementing strategy and business models. The study focused on three key areas of digital transformation that enable more effective strategy implementation and business model adaptation throughout the various phases of the entrepreneurial process.

Various factors in the entrepreneurial ecosystem, such as specific legislation, cultural environment, level of technological development, and infrastructure accessibility, can limit or accelerate the adoption of new technologies (Djokić and Kunšek, 2018; Omrani *et al.*, 2022;).

In practice, these differences are also reflected in the varying implementation of email communication, email marketing, social networks, websites, and advanced analytical tools (Širec *et al.*, 2024).

A company's website is often the first step toward establishing a clear and professional online identity, where potential customers and interested stakeholders can obtain information about the company and its offerings. Unlike social networks, which change rapidly and require constant creation of new content, a website enables a more permanent institutional presence that strengthens the company's credibility (Nambisan, 2017).

Social networks (Facebook, Twitter, Instagram, etc.) have established themselves in recent years as a crucial tool for rapidly disseminating information, building relationships with customers, and delivering targeted advertising, which can be a challenge but also an opportunity, especially for younger companies (Huang *et al.*, 2019; Kaplan and Haenlein, 2010; Qalati *et al.*, 2021).

Companies need data analytics to understand market trends, customer behaviour, and optimise business strategies. Advanced analytics tools empower entrepreneurs to make informed decisions based on higher-quality data, thereby enhancing operational efficiency (Adesina *et al.*, 2024; Kunšek, 2022; Zahay, 2021).

In the context of the increasing intertwining of AI with business processes and strategic decision-making, the authors of the GEM survey paid special attention to the issue of the acceptance and role of AI in the national business environment. The survey assessed entrepreneurs' awareness of the necessity of developing and introducing AI solutions, as well as the actual incorporation of AI into the business models of new and growing companies.

The following table provides a structured comparison of Slovenia with selected high- and low-performing EU countries across the key dimension examined in the study (Hasbu *et al.*, 2022; Marston *et al.*, 2011; Schaltegger and Wagner, 2011).

*Table 1: Comparative analysis of Slovenia and selected EU countries*

Dimension	Slovenia	High use/awareness countries	Low-use/awareness countries	Key observations/implications
Email communication	High: 68% early, 73% est.	Luxembourg, Croatia, Slovakia	Armenia	Slovenia is close to the top performers; e-mail remains central in B2B contexts.
Email marketing	Low: 31% early, 22% est.	Greece, Luxembourg	Norway, Sweden	Slovenia underperforms; GDPR and a small market reduce use.
Websites	Moderate: 56% early, 41% est.	Greece, Cyprus, Bosnia & Herzegovina	Poland, Armenia, Estonia	Below leading Southern EU cases; mature firms invest less.
Social media	Moderate: 48% early, 31% est.	Bosnia & Herzegovina, Cyprus, Latvia	Germany, Switzerland, Austria	Slovenia lags Southern/Eastern peers; stronger reliance on websites/email.

Online stores	Above average: 44% early, 29% est.	Belarus, Luxembourg, Bosnia & Herzegovina	Norway, Hungary	Close to EU average; low uptake among established firms.
Data analytics	High: 56% early, 45% est.	Cyprus, Luxembourg, Croatia	Poland	Slovenia among leaders; strong adoption in later stages.
Cloud services	High: 58% early, 54% est.	Cyprus, Luxembourg, Estonia, Greece	Ukraine, Austria, Romania, Serbia	Well above average; strong integration across firm stages.
AI – awareness of importance	High: 6.18	Switzerland, Latvia	EU avg ~5.8	Above average, but not top performer.
AI – employee competences	Low: 4.7	Latvia	EU avg ~4.5	Skills gap persists; more training needed.
AI – institutional support	Weak: 3.1	— (EU avg 3.6)	—	Slovenia lags in public institutional support.
AI – ethical trust	Moderate: 5.1	Austria (7.3)	EU avg ~5.3	Slightly below EU average.

(Source: GEM, 2025)

The results of the survey highlight significant differences in the adoption of digital tools, sustainability practices, and AI across EU countries. In summary, digital adoption rates in Slovenia indicate medium to high digital maturity. However, but standard deviation is about 15, indicating uneven progress across different dimensions. In terms of AI, Slovenia exceeds the EU average in awareness but lags in institutional support and employee competencies, but a weak correlation between data infrastructure and communication.

### 3.2. Awareness of the usefulness of official data in verifying business partners and preventing fraud

Fraud prevention is based on a comprehensive approach to all risks. Companies must cultivate moral values, raise awareness, assess fraud-related risks, and implement concrete measures to reduce and limit the opportunities for fraud to occur (Chatterjee *et al.*, 2024; Mladenovič, 2007).

In an increasingly digital world, false data and fraud have emerged as one of the most significant challenges facing businesses and organisations. An effective business risk management system relies on utilising the right data sources to enable accurate and comprehensive fraud analysis. Sources can include both internal and external data. External data can be drawn from social media, government databases, and third-party data providers.

According to OECD (2025) criteria, Slovenia is among the most transparent countries in the EU. In accordance with the constitutional principle of promoting free entrepreneurial initiative (Constitution Act, 1991, Article 74), most of this data is free of charge and permanently accessible through a portal managed by the Agency of the Republic of Slovenia for Public Legal Records and Services (AJPES). Given the knowledge of how registers operate in the EU, these facts differ most from access to data in other member states, which is also one of the fundamental reasons for focusing the research on Slovenia.

The data must be reliable, official, unadulterated, accessible, and regularly updated. The data from the various registers are precisely those whose use can facilitate fraud prevention. Registries are no longer just dry data carriers, but a fundamental element in the fight against financial crime, especially money laundering (Foster More, 2025).

Given the constancy and ease of access, as well as the fact that this data is official and regularly updated, the data in the registers can be a crucial factor in identifying anomalies in business partners and a key to making better business and life decisions (Kunšek, 2016).

The responses of the survey showed that 90% of them are aware that the right data at the right time is crucial when verifying business partners. More than 70% of the respondents also reported using a credit rating when verifying business partners. More than 80% of the respondents also stated that they are aware of AJPES as a data platform, but only 30% believed that this data could assist them in verifying business partners.

#### **4. THE DISCUSSION AND CONCLUDING REMARKS**

The purpose of establishing registers is to ensure transparency and thus greater security in legal transactions. Checking business partners can be quite easy by using data from various registers, and allows:

1. Identity verification – by cross-checking business identities, registers help verify the legal existence of a company, its directors, and shareholders, and detect fake companies or shell companies. By monitoring directors, it is also possible to track the involvement of directors in companies that have failed. For example, registers of ultimate beneficial owners (UBO) reveal the real people behind a company.
2. Detecting suspicious patterns – such as frequent changes of directors, owners, addresses, changing bank accounts, or even changing a bank account just before concluding a deal.
3. Due diligence of suppliers and customers – when we can verify not only the credibility but also the reliability of a business partner (e.g., by reviewing the register of pledged movable property, annual reports, or other relevant documents, we can obtain a forecast of the probability of non-payment).

If the client were to check only the data in the key registers, they would find such things as a change in ownership and beneficial owner, an inappropriate location of the company's registered office, changes to the bank account shortly before the conclusion of the transaction, an unfavorable business trend, pledge of movable assets, and a change in the partnership agreement at the time of signing the subscription agreement. When checking the financial position of business partners, data from the annual report is most often used.

In all cases where there is a significant disparity between the size of the intended transaction and the size of the business partner's assets, it is recommended to obtain a credit rating, which also provides insight into the business partner's current solvency and the risk of default over the next 12 months. This is usually used to verify the economic and financial capacity of bidders in public procurement procedures and is indispensable for comprehensive monitoring of the credit risk of existing and potential business partners (Kunšek, 2022).

The importance of data interconnection and international cooperation is also highlighted by the key organisation committed to preventing money laundering, the Financial Action Task Force (FATF, 2021). With the interconnection of data or the interoperability of registers, unfair business practices (e.g. identity theft) are almost impossible. FATF statistics show an exponential growth of shell companies in the United States of America (USA) in recent years.

The reason for this can certainly be attributed to the structure of the US member states, which are only obliged to comply with federal legislation that does not envisage any linking of registers and register data within or between the states. In contrast, Slovenia, where registers are also linked to the central population register, tax register, register of geodetic units, and many other official registers, has not recorded any business identity theft in more than 10 years. The principle guiding the EU Commission (EC) since the adoption of the so-called Transparency Directive (Directive 2004/109) strengthened the initiative to establish a transparent business environment at both the national level of Member States and the EU level. It also led to initiatives for greater accessibility and exchange of data between Member States, based on the so-called Digitisation Directive (Directive 2019/1151). Through the e-justice platform managed by the EC, every citizen can access basic data on business entities in the EU. We also find in practice several information projects and solutions, the main purpose of which is to connect data on the operations of business entities in different EU countries (Goens, 2010). Based on the Mobility Directive (Directive 2019/2151), the EC also advocates a policy of ease of access. The BRIS system (Business Registers Interconnection System) among data also enable the exchange of messages and data between EU Member States, such as branch registration notification or branch closure notification, notification on changes to documents and information of the company, details of the company's website, company status (as defined in national law), company activity, list of names and responsibilities of representatives, information on branches opened by the company in another Member State. Each Member State must also find a solution for the publication of so-called "disqualified directors", i.e., those former directors against whom criminal proceedings have been initiated (Kunšek, 2023).

Similarly, in accordance with the provisions of the Anti-money Laundering Directive, the so-called AML5 (Directive (EU) 2018/843), is also creating a connection between the registers of beneficial owners in the BORIS system (Beneficial Ownership Registers Interconnection System), where we can find different advanced digitally approach (ex. blockchain technology) (Gilmour *et al.*, 2025).

A more advanced approach, appropriate to the digital era, is enabled by the use of online services that enable the use of automated data acquisition solutions, with different scopes and data sets, with which the user can acquire data about business entities that interest them, for example, business partners, competitors, debtors, and others from various data sources. Thus, the user can always receive the latest or current data via the online service and include data of business partners in its model of dealing with business partners (for example, from whom insurance should be requested), sales processes, making payments, enforcing claims and filing executions and filing liabilities in bankruptcy proceedings, etc., which means that management can focus on high-quality and systematic management of business risks and redirect the attention of its operations from the field of obtaining and updating data to the customer, sales and risk management and therefore significantly enhance fraud detection (Gabrielli *et al.*, 2024).

The advantages of using registry data in fraud prevention are the diversity of the data, efficient enrolment and data collection, which can lead to ensuring legal certainty through the interoperability of different registers (Klimek *et al.*, 2024). But it is also necessary to be aware of the disadvantages, such as the quality of data, lack of active follow up and the most common – a lack of detail in the data collected (Rubinger *et al.*, 2023).

For anti-money-laundering, fraud prevention, and economic crime control can also be critical transparency (Panico, 2020). However, the possibility of the data being accessible to the general public instead of to legal persons raised the issue of violating the principles of respect for private or family life and the protection of personal data (Cindori, 2024). The recent case law of the Court of Justice of the EU (C-37/20, WM and C-601/20 Sovim 2) has ruled that access to personal data is to be prevented for the general public (Kunšek, 2024). We cannot forget that GDPR aims to protect personal data, which, in the context of linked data and data aggregation in the realm of complex IT systems, conflicts with the provisions of the GDPR (Smirnova and Morales, 2023).

## **5. CONCLUSION**

Current and systematic verification of business partners, along with the prevention of fraud, is key to successful business in an era of numerous changes (Giles, 2008). Monitoring and analysing numerous business partners and comparing them with each other can be a time-consuming task. Users can use various analytical online tools that make analysing and selecting new business partners more rational in terms of time and cost. Such a tool enables monitoring operations according to multiple indicators and also allows for comparison with competitors in the industry and region. Undoubtedly, one of the greatest added values of using such a tool is receiving daily notifications about changes in data in registers, problems with solvency, or announcements in various procedures at a business partner, which ensures a timely response and additional advantages in managing business and financial risks.

The digital age offers numerous opportunities for easy data capture and processing, which, with the appropriate level of scepticism and simple preventive actions, can lead to more effective risk management. Regardless of the above, data credibility remains a crucial element in making informed business decisions at the right time. An essential step in this regard is for management to start using the advantages of the digital age as soon as possible and minimise exposure to risks. As fraudsters use sophisticated techniques to exploit financial systems and personal data, it is essential to adopt agile and state-of-the-art methods to verify information and data and prevent fraudulent activities (FIOD, 2025).

Preventing risks and dishonest business practices requires proper risk management and the adoption of suitable measures. At both national and international levels, the National Resolution on the National Programme for the Prevention and Suppression of Crime (ReNPPZK24–28) and COSO's Guide (2016) emphasise the importance of preventive activities — awareness-raising, education, information, communication, and cooperation.

Experts agree that accelerated digitisation demands an urgent need to invest in developing employees' skills for working with AI (training, certification), including the introduction of AI into educational programs after graduation. It therefore makes sense to develop guidelines and support measures for the introduction of AI into business models, including financing for the digital transformation of companies and educational initiatives to understand the specific benefits of AI and to use more data under a suitable regulatory framework to ensure business security, user protection, and the ethical use of AI (Širec *et al.*, 2024).

## ACKNOWLEDGEMENTS

This paper is part of the research program group Program P5-0441 – Regeneration of economy and business, which is supported by the Javna agencija za znanstveno-raziskovalno in inovacijsko dejavnost Republike Slovenije (angl. Slovenian Research and Innovation Agency; ARIS) (14. člen Splošnega akta o stabilnem financiranju znanstvenoraziskovalne dejavnosti (Uradni list RS, št. 87/22 in 103/22 – popr.) and part of the research program group Program 14-ULTRA\_4.04 – Sustainable development University of Ljubljana, which is supported by the Republic of Slovenia, Ministry of Higher Education, Science and Innovation, and the EU – NextGenerationEU.

## REFERENCES

- Aboud, A. and Robinson, B. (2020), "Fraudulent financial reporting and data analytics: an explanatory study from Ireland", *Accounting Research Journal*, Vol. 35, No. 5, pp. 21-36.
- ACFE (2020), "Report to the Nations: 2020 Global Study on Occupational Fraud and Abuse", available at: <https://legacy.acfe.com/report-to-the-nations/2020/> (accessed 20 September 2025).
- Adeboye, E. (2024), "Strengthening fraud prevention in small businesses: An analysis of effective accounting and auditing practices", *International Journal of Science and Research Archive*, Vol. 13 No. 2, pp. 590-595.
- Adesina, A. A., Iyelolu, T. V. and Paul, P. O. (2024), "Leveraging predictive analytics for strategic decision-making: Enhancing business performance through data-driven insights", *World Journal of Advanced Research and Reviews*, Vol. 22 No. 3, pp. 1927-1934.
- Chatterjee, S., Mariani, M. and Ferraries, A. (2024), "Digitalization of supply chain and its impact on cost, firm performance, and resilience: Technology turbulence and top management commitment as moderator", *IEEE Transactions on Engineering Management*, Vol. 71, pp. 10469-10484.
- Cindori, S. (2023), "Beneficial Ownership – Demand for Transparency, Threat to Privacy", *Review of EU and Comparative Law*, Vol. 55 No. 4, pp. 113-131.
- COSO (2016), "Fraud Risk Management Guide: 2nd Edition", available at: [www.coso.org/fraudderence](http://www.coso.org/fraudderence) (accessed 20 September 2025).
- Court of Justice of the EU (2020), *Judgment of the court in joint case C-37/20, WM and C-601/20 Sovim 2*, available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62020CJ0037> (accessed 28 October 2025).
- Djokić, D. and Kunšek, M. (2018), "Is legislation the only key condition for the development of quality entrepreneurship", in *Abstracts of the proceedings: new business models and institutional entrepreneurs: leading disruptive change*, CIRU – Governance research and development centre: University of Dubrovnik, Department of Economics and Business Economics, Zagreb, pp. 89-90.
- EU Parliament and Council (2004), *Directive 2004/109/EC of the EU Parliament and of the Council of 15 December 2004 on the harmonisation of transparency requirements in relation to information about issuers whose securities are admitted to trading on a regulated market and amending Directive 2001/34/EC (Transparency Directive)*, available at: <https://eur-lex.europa.eu/eli/dir/2004/109/oj>.
- EU Parliament and Council (2015), *Directive (EU) 2015/849 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing and amending Directives 2009/138/EC and 2013/36/EU (AML 5 Directive)*, available at: <http://data.europa.eu/eli/dir/2015/849/oj>.
- EU Parliament and Council (2016), *Regulation (EU) 2016/679 of the EU Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the*

- processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (GDPR), available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679> (accessed 28 October 2025).
- EU Parliament and Council (2019), *Directive (EU) 2019/1151 of the EU Parliament and of the Council of 20 June 2019 amending Directive (EU) 2017/1132 as regards the use of digital tools and processes in company law (Digitalization Directive)*, available at: <http://data.europa.eu/eli/dir/2019/1151/oj>.
- EU Parliament and Council (2019a), *Directive (EU) 2019/2121 of the EU Parliament and of the Council of 27 November 2019 amending Directive (EU) 2017/1132 as regards cross-border conversions, mergers and divisions (Mobility Directive)*, available at: <http://data.europa.eu/eli/dir/2019/2121/oj>.
- FATF (2021), "Opportunities and Challenges of New Technologies for AML/CFT", available at: <https://www.fatf-gafi.org/publications/fatfrecommendations/documents/opportunities-challenges-new> (accessed 19 September 2025).
- FIOD (2025), "Insolvency fraud", available at: <https://www.fiod.nl/en/investigation-themes/insolvency-fraud/> (accessed 30 September 2025).
- Foster Moore (2025), "What Is The Cost of Being Greylisted by FATF?", available at: [https://www.fostermoore.com/news/what-is-the-cost-of-being-greylisted-by-fatf?utm\\_campaign=Thought%20Leadership&utm\\_content=338573746&utm\\_medium=social&utm\\_source=linkedin&hss\\_channel=lcp-1525078](https://www.fostermoore.com/news/what-is-the-cost-of-being-greylisted-by-fatf?utm_campaign=Thought%20Leadership&utm_content=338573746&utm_medium=social&utm_source=linkedin&hss_channel=lcp-1525078) (accessed 20 September 2025).
- Gabrielli, G., Magri, C., Medioli, A., and Marchini, P. L. (2024), "The power of big data affordances to reshape o fanti-fraud strategies", *Technological Forecasting and Social Change*, Vol. 205.
- GEM (2025), *Global Entrepreneurship Monitor 2024/2025 Global Report: Entrepreneurship Reality Check*, London: GEM.
- Giles, S. (2013), *Managing Fraud Risk: A Practical Guide for Directors and Managers*, John Wiley & Sons, New York, NY.
- Gilmour, P., Pandey, D. and Goldbarsht, D. (2025), "Registers of beneficial owners based on blockchain technology: Implications for the accounting profession", *Technological Forecasting & Social Change*, No. 214.
- Gobec Zidar, L. (2003), *Roles and methods of the internal auditor in detecting, investigating and preventing fraud and corruption*. Specialist work. University of Ljubljana: Faculty of Economics.
- Goens, D. (2010), "The exploitation of Business Register data from a public sector information and data protection perspective: A case study", *Computer Law & Security Review*, Vol. 26 No.4, pp. 398-405.
- Harika Putri, M.A.C.D and Nurmalasari, M.R. (2024), "AI-Entrepreneurship Future Research and Opportunities for New Business Model", *TIERS – Information Technology Journal*, Vol. 5 No. 1.
- Hasbu, W. Y., Zarkasyi, W. and Suharman, H. (2022), "Does creating shared value implies corporate performance?", *Academy of Strategic Management Journal*, Vol. No. 1, pp. 1-12.
- Huang, L., Clarke, A., Heldsinger, N. and Tian, W. (2019), "The communication role of social media in social marketing: a study of the community sustainability knowledge dissemination on LinkedIn and Twitter", *Journal of Marketing Analytics*, Vol. 7 No. 2, pp. 64-75.
- IAASB (2025), *International Standard on Auditing (ISA) 240*, available at: <https://ifacweb.blob.core.windows.net/publicfiles/2025-07/IAASB-ISA-240-Revised-Fraud.pdf> (accessible 30 September 2025).

- Kaplan, A. M. and Haenlein, M. (2010), "Users of the world, unite! The challenges and opportunities of Social Media", *Business Horizons*, Vol. 53 No.1, pp. 59-68.
- Klimek, J., Chortaras, A., Mišek, J., Yang, J.J., Skagemo, S. and Tzouvaras, V. (2024), "The STIRData Approach to Interoperability of EU Company High-Value Datasets", *SN Computer Science*, Vol. 5 No. 4.
- Kunšek, M (2023), "Transparentnost na preizkušnji", *Pravna Praksa*, Vol. 42, No. 6/7, pp. 20-21.
- Kunšek, M. (2016), "With AJPES and the right information to make good decisions", *Public Administration*, Vol. 52 No. 1-2, pp. 107-12.
- Kunšek, M. (2021), "With a digital approach to easier liquidity management", *Proceeding of the 1st FinD-INFO Days: Taxes, Finance, Controlling, Accounting in Portorož, Slovenia*, Lexpera, GV založba, pp. 31-36.
- Kunšek, M. (2022), "With official data to simple verification of business partners (special topic)", *Company and Work*, Vol. 48 No. 6-7, pp. 1067-1075
- Kunšek, M. (2024), "Transparentnost da, vendar le za organe pregona", *Pravna Praksa*, Vol. 43, No. 1-2, pp. 14-15.
- Kunšek, M. (2025), "Monitoring the operations of business partners is the key to stable business", *FinD-INFO articles*, Lexpera.
- Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J. and Ghalsasi, A. (2011), "Cloud computing — The business perspective", *Decision Support Systems*, Vol. 51 No. 1, pp. 176-189.
- Mladenovič, Z. (2012), "Responsibility of the State Auditor for Fraud Detection", available at: [https://www.slideserve.com/kalin/odgovornost-dr-avnega-revizorja-za-odkrivanje-prevar#google\\_vignette](https://www.slideserve.com/kalin/odgovornost-dr-avnega-revizorja-za-odkrivanje-prevar#google_vignette) (accessed 23 September 2025).
- Nambisan, S. (2017), "Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship", *Entrepreneurship theory and practice*, Vol. 41 No. 6, pp. 1029-1055.
- OECD (2025), "Global Forum on Transparency and Exchange of Information for Tax Purposes", available at: <https://www.oecd.org/en/networks/global-forum-tax-transparency.html> (accessed 28 October 2025).
- Omrani, N., Rejeb, N., Maalaoui, A., Dabić, M., and Kraus, S. (2022), "Drivers of digital transformation in SMEs", *IEEE transactions on engineering management*, Vol. 71, pp. 5030-5043.
- Panico, P (2020), "Private Foundations and EU beneficial ownership registers: towards full disclosure to the general public?", *Trusts & Trustees*, Volume 26 No. 6, pp 493-502.
- Povh, D. (2003), "The role of internal auditors in detecting corruption fraud and planning the fight against it", *Revizor*, Vol. 11, pp. 7-29.
- PWC (2020), "Global Economic Crime and Fraud Survey 2020", available at: <https://www.pwc.com/gx/en/services/forensics/economic-crime-survey.html> (accessed 20 September 2025).
- PWC. (2025), "Know Your Business Partner", available at: <https://store.pwc.de/en/products/know-your-business-partner> (accessed 20 September 2025).
- Qalati, S. A., Yuan, L. W., Khan, M. A. S., and Anwar, F. (2021), "A mediated model on the adoption of social media and SMEs' performance in developing countries", *Technology in Society*, Vol. 64, p. 101513.
- Republic of Slovenia (1991), *Constitution Act (URS)*, Official Gazette of the Republic of Slovenia, available at: <https://pisrs.si/pregledPredpisa?id=USTA1>.
- Republic of Slovenia (2006), *Companies Act (ZGD-1)*, Official Gazette of the Republic of Slovenia, available at: <https://pisrs.si/pregledPredpisa?id=ZAKO4291>.

- Republic of Slovenia (2007), *Act on the Market in Financial Instruments (ZTFI)*, Official Gazette of the Republic of Slovenia, available at: <https://pisrs.si/pregledPredpisa?id=ZAKO7888>.
- Republic of Slovenia (2008), *Criminal Code (KZ)*, Official Gazette of the Republic of Slovenia, available at: <https://pisrs.si/pregledPredpisa?id=ZAKO5050>.
- Republic of Slovenia (2024), *Money Laundering and Terrorist Financing Prevention Act (ZPPDFT-2)*, available at: <https://pisrs.si/pregledPredpisa?id=ZAKO8547>
- Republic of Slovenia (2024a), *Resolution on the National Program for the Prevention and Suppression of Crime for the Period 2024–2028*, Official Gazette of the Republic of Slovenia, available at: <https://pisrs.si/pregledNpb?idPredpisa=RESO156&idPredpisaChng=RESO156>.
- Rodgers, M., Mukherjee, S., Melamed, B., Baveja, A., and Kapoor, A. (2024), “Solving business problems: the business-driven data supported process”, *Annals of Operations Research*, Vol. 332 No. 1, pp. 703-741.
- Rubinger, L., Ekhtiari, S., Gazendami, A., and Bhandari, M. (2023), “Registries: Big data, bigger problems?”, *Technological Forecasting and Social Change*, Vol. 54 No. 3, pp. 539-542.
- Schaltegger, S., and Wagner, M. (2011), “Sustainable entrepreneurship and sustainability innovation: categories and interactions”, *Business Strategy and the Environment*, Vol. 20 No. 4, pp. 222-237.
- Širec, K., Crnogaj, K., Bradač Hojnik, B., Rus, M., and Tominc, P. (2025), *Podjetništvo ob preboju digitalizacije in umetne inteligence, GEM Slovenija 2024*. Ekonomsko-poslovna fakulteta Univerza v Mariboru, Maribor.
- Smirnova, Y. and Morales, V. (2024), “Understanding challenges of GDPR implementation in business enterprises: a systematic literature review”, *International Journal of Law and Management*. Vol. 66, pp. 326-344.
- SSKJ (2020), Založba ZRC, Ljubljana.
- Trulioo. (2023), "Trulioo Global Survey: 79 % of Companies Experience Business Identity Theft, Driving Urgency in Verification Market", available at: [https://www.trulioo.com/company/newsroom/press-releases/survey-business-identity-theft-driving-urgency-in-verification-market?utm\\_source=chatgpt.com](https://www.trulioo.com/company/newsroom/press-releases/survey-business-identity-theft-driving-urgency-in-verification-market?utm_source=chatgpt.com) (accessed 22 September 2025).
- Turk, I. (2003), "Some Business Terms", *Revizor*, pp. 105-133.
- Usman, F., Eyo-Udo, N., Etukudoh, E., Odonkor, B., Ibeh, C., and Adegbola., A. (2024), “A critical review of AI-driven strategies for entrepreneurial success”, *International Journal of Management & Entrepreneurship Research*, Vol. 6 No. 1, pp. 200-215.
- Wells, J. T. (2017), *Corporate Fraud Handbook: Prevention and Detection (5th ed.)*, Hoboken, John Wiley & Sons, New York, NY.
- Zahay, D. (2021), “Advancing research in digital and social media marketing”, *Journal of Marketing Theory and Practice*, Vol. 29 No. 1, pp. 125-139.

## BEHAVIORAL ARCHETYPES OF GEN Z IN GAMIFIED E-GOVERNMENT APPS

**Kalina Trenevaska Blagoeva**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*Kalina.trenevska@eccf.ukim.edu.mk*

**Lidija Pulevska Ivanovska, PhD**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*Lidija.pulevska@eccf.ukim.edu.mk*

**Saso Josimovski, PhD**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*Saso.josimovski@eccf.ukim.edu.mk*

**Marina Mijoska Belshoska, PhD**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*Marina.mijoska@eccf.ukim.edu.mk*

**Martin Kiselicki, PhD**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*martin.kiselicki@eccf.ukim.edu.mk*

### **ABSTRACT**

*This study applies a segmentation lens to Gen Z participation in a gamified e-government app (MojDDV). K-means clustering of survey data (n=119) yields four archetypes—Super-Scanners, Socially-Influenced Actives, Casual Participants, Skeptical Habituals—that differ in trust, social influence, and scanning intensity. The contribution is to move beyond average-effect adoption models and show actionable heterogeneity: leverage advocates (Super-Scanners), reinforce peer norms (Socially-Influenced), nudge and simplify (Casual), and rebuild institutional trust (Skeptical). Limitations include single-country scope, self-reports, and sample size; future work should add robustness checks and external validation. Findings inform targeted policy and design for gamified public services.*

**Keywords:** Behavioral archetypes, Gamified e-government, Generation Z, Clustering analysis, Digital engagement

**JEL classification:** H83, O33, C38

### **1. INTRODUCTION**

In a variety of settings, gamification has been widely used to improve user motivation, engagement, and compliance (Deterding *et al.*, 2011; Hamari *et al.*, 2014). Gamification in the public sector gives governments new ways to engage citizens in compliance and service delivery, especially when traditional enforcement isn't enough (Hassan and Hamari, 2020). The

"Taxman" program in Brazil (Alm and Torgler, 2011) is an example of previous research that shows how gamified tax initiatives can increase transparency and decrease informality. One of the most comprehensive gamification initiatives for e-government in the Western Balkans is the MojDDV app. The app, which was released in 2019, aligns individual financial incentives with collective fiscal transparency by partially refunding citizens' VAT payments if receipts are scanned (Akademik, 2021). According to earlier studies on MojDDV, performance and effort expectancy had supportive roles, but facilitating conditions and social influence were important predictors of Gen Z adoption. These findings imply that maintaining participation in gamified e-government services requires both peer support and infrastructure.

With an estimated 30–40% of GDP, North Macedonia continues to have one of the largest shadow economies in Europe (Medina and Schneider, 2018). According to earlier research, informal practices are fueled by low institutional trust and low tax morale (Frey and Torgler, 2007). Gen Z frequently engages in informal online markets due to convenience and financial pressures, but they are also aware of the risks of tax evasion and decreased transparency, according to a previous study on consumer perceptions of the digital shadow economy (Kiselicki *et al.*, 2024). These observations imply that interaction with MojDDV should be interpreted in light of Gen Z's conflicted feelings about formality: although they appreciate openness, they frequently strike a balance between compliance and self-interest.

Digital natives born between the late 1990s and the early 2010s, known as Generation Z, demand social integration, efficiency, and transparency from digital services (Seemiller and Grace, 2016; Williams, 2002). According to research on MojDDV, women in Gen Z are more consistent in their app usage, more engaged, and more influenced by social norms than men (Josimovski *et al.*, 2024). This result is consistent with global research demonstrating that gendered digital behaviors and peer dynamics impact adoption in younger cohorts (Matveyeva and Krasnov, 2019). Although UTAUT and similar models have a great ability to explain the adoption of technology (Venkatesh *et al.*, 2003), they frequently assume that user populations are all the same. Nonetheless, e-business research highlights how crucial user segmentation is to comprehending trends in digital adoption. In order to create more specialized and efficient service design, techniques like clustering analysis have been used to identify user archetypes in social media platforms and online shopping (Mekler *et al.*, 2017). Segmentation can reveal hidden heterogeneity in the context of e-government, such as habitual adopters or socially influenced clusters that largely follow peer norms. This study aims to provide a fresh perspective on MojDDV adoption among Gen Z by fusing gamification theory, e-government adoption models, and e-business segmentation techniques.

## 2. METHODOLOGY

In order to identify behavioral archetypes among users of MojDDV, a gamified e-government application in North Macedonia, this study uses an exploratory, data-driven design. The current method uses unsupervised machine learning techniques, particularly cluster analysis, to divide users into groups according to their perceptions, intentions, and engagement behaviors, in contrast to earlier research that used confirmatory acceptance models. Finding hidden adoption patterns that could direct more specialized policy and design interventions was the aim, not testing causal hypotheses. Between February and March 2024, an online survey was used to gather data, and 119 valid answers were obtained. Since Generation Z respondents (those between the ages of 18 and 24) are the most pertinent group for gamified e-government services, they made up the majority of the target population. To guarantee significant insights, the analysis only included people who attested to having previously used MojDDV.

There were three components to the survey instrument. The first recorded opinions about MojDDV, with items indicating institutional support, social endorsement, usefulness, and ease of use. Despite being modified from well-known models like UTAUT (Venkatesh *et al.*, 2003), these metrics were used as input dimensions for clustering rather than for hypothesis testing. Cronbach's alpha values exceeded the suggested 0.70 threshold, indicating good internal consistency, according to a reliability check (Nunnally, 1978). The frequency with which respondents used MojDDV to scan receipts was the measure of behavioral engagement in the second section. To capture the intensity of use, open-text responses were recoded into an ordinal scale (1 = only a few receipts, 4 = all receipts). In order to profile the resulting clusters, the third section gathered demographic data (age, gender, and employment status). There were four steps in the analytical process. First, all variables were standardized to guarantee comparability, and composite indicators were computed by averaging the Likert-scale items. Second, k-means clustering was used, and the parsimony and interpretability of three-cluster and four-cluster solutions were compared. Third, cluster centers were examined to profile the resulting clusters, and then chi-square tests and cross-tabulations were used to investigate demographic differences between groups. Lastly, the most significant segmentation solution was chosen based on three criteria: interpretability, theoretical coherence, and distinctiveness. All respondents were guaranteed anonymity and informed of the study's purpose, and participation in the study was entirely voluntary. To ensure adherence to ethical research standards, no personally identifiable information was gathered, and the data were examined in aggregate form.

### **3. LITERATURE REVIEW**

In order to encourage behavior, gamification entails transferring aspects of game design, like competition, rewards, and feedback, to non-gaming contexts (Deterding *et al.*, 2011). According to research, gamification can improve compliance in domains like education, health, and taxes by making routine tasks more enjoyable (Hamari *et al.*, 2014; Hassan and Hamari, 2020). Governments have tried gamified initiatives in the fiscal domain to increase transparency and decrease informality. Examples include Brazil's Taxman initiative, which enhanced revenues by incorporating entertaining mechanisms into tax reporting (Alm and Torgler, 2011). The MojDDV app, which reimburses citizens with partial VAT refunds for scanning receipts, is a trailblazing attempt to implement gamification in e-government in North Macedonia (Akademik, 2021; Public Revenue Office, 2022). Previous studies have demonstrated that social influence and favorable circumstances are powerful predictors of MojDDV adoption in Generation Z, underscoring the significance of peer relationships and infrastructure in the adoption of digital government.

#### **3.1. Generation Z and digital government**

Research continuously indicates that when Gen Z uses online platforms, they place a high value on social connectivity, efficiency, transparency, and personalization (Seemiller and Grace, 2016; Williams, 2015). They are especially responsive to playful and reward-driven mechanisms in public digital services because they are used to smooth user experiences, immediate feedback, and gamified interactions (Priporas *et al.*, 2017; Dolot, 2018). In the context of gamified e-government services, reward systems, social validation, and user-friendly interfaces that complement their digital habits tend to be well received by Generation Z (Andersson *et al.*, 2022; Aljovic, 2024). Digital natives, for example, are more likely to use government apps when they are presented as both socially acceptable (e.g., visible participation within peer networks) and individually beneficial (e.g., financial incentives) (Hassan and

Hamari, 2020). These characteristics imply that MojDDV meets the expectations of this generation by providing both community-driven norms and direct financial benefits. However, data indicates that Gen Z is not a monolithic group. Men exhibited more irregular or utilitarian usage, whereas women were more involved and more impacted by social reinforcement, according to earlier studies on MojDDV (Josimovski *et al.*, 2024). While men tend to concentrate more narrowly on utilitarian or competitive benefits, this finding is consistent with larger research that indicates women are more receptive to social and community aspects of gamified systems (Koivisto and Hamari, 2014). The relationship between Gen Z and the digital shadow economy adds even more complexity. Many members of this generation acknowledge that they participate in informal online markets because they are affordable, accessible, or convenient, despite the fact that they understand the value of transparency and compliance (Kiselicki *et al.*, 2024). This ambivalence raises the possibility that their use of formal systems such as MojDDV may be contingent and influenced by pragmatic cost-benefit analysis rather than strictly normative commitments.

### **3.2. Segmentation and archetypes in digital services**

To test causal relationships, the majority of research on digital adoption in public administration uses models like the Unified Theory of Acceptance and Use of Technology (UTAUT) or the Technology Acceptance Model (TAM) (Venkatesh *et al.*, 2003). Although useful, these methods frequently focus on average effects across the sample and treat users as homogeneous populations. These methods run the risk of ignoring the fact that digital citizens differ in their levels of trust, motivations, and behaviors. In contrast, researchers can find unique user groups, or archetypes, within a larger population by using segmentation techniques like cluster analysis. Archetypes combine behavioral patterns and attitudinal differences to represent qualitatively distinct ways of interacting with a system (Wedel and Kamakura, 2000).

Segmentation is frequently used in e-business research to target interventions, enhance customer journeys, and personalize services (Mekler *et al.*, 2017). For instance, gamified learning systems distinguish achievement-driven learners from social learners, and online shopping platforms distinguish between browsers and heavy buyers. Businesses can use these insights to tailor their engagement tactics to the needs of various clientele. According to research on gamification, not all users are driven by the same mechanisms; some are more motivated by incentives and rewards, others by social recognition, and still others by competition or mastery (Marczewski, 2015; Tondello *et al.*, 2016). There are significant opportunities when these insights are applied to e-government. People may use gamified apps like MojDDV for a variety of reasons, including social pressure, financial gain, or a sense of civic obligation. Policymakers can learn not only how many citizens use an app, but also why and how they use it by using segmentation to capture this behavioral diversity. This is crucial in situations where there is a lack of institutional trust, as people may participate grudgingly or out of habit without believing in the system's worth (Frey and Torgler, 2007).

Defining archetypes is thus central to designing effective gamified e-government systems. Segmentation studies consistently show that different user groups respond to gamification in qualitatively distinct ways, ranging from those who fully embrace playful systems to those who comply reluctantly or sporadically (Wedel and Kamakura, 2000; Marczewski, 2015; Tondello *et al.*, 2016). In the context of MojDDV, clustering analysis revealed four meaningful archetypes of Generation Z users:

- Super-Scanners, enthusiastic and loyal users who resemble “Achievers” in gamification typologies, fully trusting the system and promoting it within their networks (Marczewski, 2015; Tondello *et al.*, 2016).
- Socially-Influenced Users, who participate primarily because of peer and community endorsement, similar to the “Socializers” category often observed in gamified systems (Koivisto and Hamari, 2014).
- Casual Participants, who engage sporadically, reflecting the low-commitment user clusters often found in segmentation studies of e-services (Mekler *et al.*, 2017).
- Skeptical Habituals, who comply in practice but remain distrustful of the system’s purpose or of government intentions, echoing findings on reluctant tax compliance and conditional cooperation (Frey and Torgler, 2007; Hassan and Hamari, 2020).

Designing interventions that go beyond a "one-size-fits-all" strategy requires an understanding of these archetypes. Champions can be developed as advocates, but Skeptical Habituals need tactics to restore legitimacy and trust, Casual Participants might react better to stronger prods or streamlined procedures, and Socially-Influenced Users need reinforcement through community campaigns. A distinct design and policy approach is implied by each archetype. By acknowledging these differences, e-government is transformed from a universally applicable service into a customized online experience that can boost uptake, adherence, and sustainability.

### **3.3. Research gap and contribution**

Gamification has been used more and more in e-government projects over the past ten years as a tactic to boost institutional trust, enhance compliance, and promote citizen participation. Gamified mechanisms, such as Brazil's Taxman program, have been used to increase fiscal transparency, fight informality, and turn passive interactions into active ones (Alm and Torgler, 2011). Although gamification has great potential in the public sector, reviews of this new field warn that user responses vary depending on contextual, cultural, and motivational factors (Hamari *et al.*, 2014; Hassan and Hamari, 2020). One of the most comprehensive gamification efforts for e-government in the Western Balkans is the MojDDV app, which is available in North Macedonia. In addition to identifying gender differences, prior research on MojDDV has validated the significance of social influence and facilitating conditions in promoting adoption among young citizens (Josimovski *et al.*, 2024). Women are more engaged than men because they are more sensitive to social cues. Other studies have looked at the connection between Gen Z and the digital shadow economy, emphasizing their conflicted views—supporting transparency in theory but engaging in informal activities when pressured by convenience or financial gain (Kiselicki *et al.*, 2024). All of these results highlight MojDDV's significance as a useful example for comprehending the adoption of gamified e-government. Despite these revelations, previous research on MojDDV and other global programs has a similar flaw in that it mostly uses linear adoption models like TAM or UTAUT to account for average effects across populations. Although these models are helpful in identifying important predictors, they ignore the possibility that citizens interact with gamified systems in qualitatively different ways because they assume that users are all the same. Users are not a single bloc, as the gamification literature itself has demonstrated; rather, they can be classified as "Achievers," "Socializers," or "Explorers," each of whom is driven by a different set of mechanics (Marczewski, 2015; Tondello *et al.*, 2016).

By adding a segmentation perspective to e-government gamification research, this study fills that gap. It identifies behavioral archetypes of MojDDV users that represent unique patterns of

motivation, trust, and engagement by utilizing clustering techniques that are frequently used in e-business and consumer analytics (Wedel and Kamakura, 2000; Mekler *et al.*, 2017). Four archetypes are identified by the analysis: Super-Scanners, Socially-Influenced Actives, Casual Participants, and Skeptical Habituals. Each of these groups needs a unique approach to maintain engagement. This paper makes three contributions. In order to show the usefulness of segmentation techniques, it first broadens the scope of the e-government gamification literature by going beyond regression-based analyses. Second, when placed within an actual fiscal transparency project, it empirically identifies user archetypes that complement and expand upon current gamification typologies. Thirdly, it provides policymakers with practical implications: Super-Scanners can be used as advocates, Socially-Influenced Actives can be targeted through community campaigns, Casual Participants can be prodded with extra incentives, and Skeptical Habituals can be engaged through measures to foster trust. This study advances knowledge of how segmentation and personalization can improve adoption, compliance, and long-term sustainability in the public sector by placing MojDDV as an example of a larger class of gamified e-government initiatives.

## 4. RESULTS

### 4.1. Descriptive statistics

A total of 119 valid responses were collected, with the overwhelming majority (97.5%) belonging to the age group 18–24, confirming the study's focus on Generation Z. Among the respondents, 70% identified as female and 30% as male. Regarding behavioral engagement, 41% of respondents reported scanning *all receipts*, 38% reported scanning *most receipts*, 15% scanned *some receipts*, and 6% scanned *only a few receipts*. This indicates a generally high level of compliance among the sample, although variation exists in the intensity of use.

Descriptive statistics for the averaged constructs indicated positive perceptions overall. Performance Expectancy received the highest mean ( $M = 4.15$ ,  $SD = 0.82$ ), suggesting that users broadly agree that MojDDV helps reduce tax evasion and ensures retailers issue receipts. Effort Expectancy was moderately high ( $M = 3.83$ ,  $SD = 0.91$ ), indicating that most respondents find the app user-friendly. Social Influence ( $M = 3.99$ ,  $SD = 1.05$ ) reflected strong peer and community encouragement, while Facilitating Conditions scored somewhat lower ( $M = 3.46$ ,  $SD = 0.91$ ), showing some ambivalence about resources and institutional support. Behavioral Intention was very high ( $M = 4.20$ ,  $SD = 0.77$ ), confirming that most users intend to continue using MojDDV regularly.

### 4.2. Cluster analysis

To uncover distinct user groups, a k-means clustering procedure was applied using six variables: Performance Expectancy, Effort Expectancy, Social Influence, Facilitating Conditions, Behavioral Intention, and Receipt Scanning Frequency. All variables were standardized (z-scores) prior to analysis to prevent scale bias. Multiple runs were conducted with different random seeds to ensure stability of results. Both three-cluster and four-cluster solutions were evaluated.

The three-cluster solution provided a parsimonious segmentation, differentiating between: (1) highly engaged adopters, (2) moderately positive users, and (3) skeptical users. However, examination of cluster centers revealed that this model merged two conceptually distinct groups: those who were enthusiastic and trusting versus those who were skeptical but still scanned frequently.

The four-cluster solution provided a more nuanced segmentation with clearer interpretability. It separated the *skeptical but habitual users* into their own group, a finding of theoretical and practical relevance for e-government adoption research. Cluster validity was assessed through internal measures (silhouette scores) and external interpretability; the four-cluster model showed higher silhouette values and clearer distinctions. Based on these criteria, the four-cluster solution was retained for subsequent analysis.

To quantitatively assess the internal validity of the four-cluster solution, several diagnostics were conducted. The overall silhouette coefficient was 0.18, with per-cluster means ranging from 0.08 to 0.29, indicating moderate but interpretable separation between clusters given the small and homogeneous sample. The within-cluster sum of squares (inertia) was 366.03, suggesting an acceptable level of compactness for a six-variable model of this scale.

A sensitivity analysis was performed by repeating the clustering after removing Social Influence from the input dimensions. The resulting solution produced a slightly higher silhouette score (0.24) and lower inertia (295.49), confirming that no single construct dominated the segmentation and that the four-cluster configuration remained stable under minor perturbations. These findings support the interpretability and robustness of the chosen solution despite the limited sample size.

### 4.3. Cluster profiles

The four identified archetypes are summarized in Table 1.

Table 1: Cluster Centers and Archetypes

Archetype	PE	EE	SI	FC	BI	Scanning	Profile
Super-Scanners	4.09	4.44	4.64	4.50	4.90	High	Enthusiastic adopters who scan all receipts, trust the system, and strongly advocate for it.
Socially-Influenced Actives	3.78	4.11	3.90	3.95	4.27	High	Users motivated primarily by peer/community norms; engage consistently due to social reinforcement.
Casual Participants	3.13	3.71	3.41	3.22	3.51	Medium	Sporadic users with moderate perceptions; engage with some receipts but require nudges to sustain activity.
Skeptical Habituals	2.25	2.65	2.53	2.40	3.36	High	Users who scan receipts regularly but distrust the app’s purpose or government intentions.

(Source: Own research)

**Super-Scanners.** This group scored the highest across all constructs, with particularly strong values for Performance Expectancy ( $M \approx 4.09$ ), Effort Expectancy ( $M \approx 4.44$ ), and Behavioral Intention ( $M \approx 4.90$ ). They reported scanning all receipts consistently. These users resemble “Achievers” in gamification theory—fully embracing the app, trusting the government initiative, and acting as strong advocates. Their high trust and consistent behavior position them as the most loyal and valuable group for sustaining MojDDV adoption.

**Socially-Influenced Actives.** Members of this group scored moderately high across most constructs, especially Social Influence ( $M \approx 3.90$ ) and Behavioral Intention ( $M \approx 4.27$ ). They

reported scanning most or all receipts, but their primary motivation stems from peer and community endorsement. In adoption theory, this group reflects the importance of social proof: they comply because “everyone around them does.” While they may not be as intrinsically motivated as Super-Scanners, their engagement is stable as long as peer norms remain supportive.

Casual Participants. This group showed moderate perceptions and the lowest scanning frequency. With means around 3.1–3.5 for constructs and lower Behavioral Intention ( $M \approx 3.51$ ), they represent sporadic users who only scan some receipts. Their engagement appears pragmatic rather than committed, requiring additional nudges, incentives, or interface simplifications to remain active. In segmentation literature, such users are often labeled as “light users” who are most at risk of disengagement if not supported.

Skeptical Habituals. Interestingly, this group reported high receipt scanning frequency despite low perceptions of the app. Their scores on Performance Expectancy ( $M \approx 2.25$ ), Effort Expectancy ( $M \approx 2.65$ ), and Social Influence ( $M \approx 2.53$ ) were the lowest among all clusters, but they still reported scanning receipts regularly. This paradoxical profile suggests they participate out of habit, obligation, or instrumental necessity, but remain distrustful of the app’s purpose and the government’s intentions. From a policy perspective, this group is the most concerning, as their compliance may be fragile and prone to erosion if incentives weaken.

#### **4.4. Demographic profiling**

Demographic analysis revealed meaningful differences across archetypes (see Table 2). Super-Scanners were predominantly young female students (over 70%), confirming earlier findings that women in Gen Z are more consistent adopters of MojDDV. Their strong advocacy and consistent scanning behavior suggest that targeted campaigns through female student networks could amplify app adoption.

Socially-Influenced Actives also had a female majority but a more balanced distribution, with some employed respondents included. Their reliance on peer norms means that their engagement may be especially sensitive to shifts in community discourse or media campaigns.

Casual Participants had the most gender-balanced profile, with a mix of students and unemployed respondents. Their weaker engagement highlights the need for targeted nudging strategies, perhaps through gamified incentives or simplified user experiences.

Skeptical Habituals stood out demographically: while still containing some young students, they also included older respondents (45–54 and 55+), along with a higher proportion of employed individuals. This supports the interpretation that skepticism may be rooted in institutional distrust rather than usability barriers.

*Table 2: Demographic distribution across clusters (n = 119)*

Archetype	Female (%)	Male (%)	Age 18–24 (%)	Age 25–34+ (%)	Students (%)	Employed/Other (%)
Super-Scanners	74	26	100	0	95	5
Socially-Influenced Actives	68	32	98	2	88	12
Casual Participants	55	45	100	0	79	21
Skeptical Habituals	75	25	83	17	58	42

(Source: Own research)

To complement the qualitative interpretation, additional diagnostics were conducted to evaluate the robustness of the segmentation. The four-cluster model achieved an overall silhouette score of 0.18, indicating moderate cohesion and separation, with per-cluster values ranging from 0.08 to 0.29. The inertia (within-cluster sum of squares) was 366.03, corresponding to roughly 62 % of the total variance explained by the model, which is consistent with exploratory studies of similar sample size and dimensionality.

A sensitivity test—performed by omitting the Social Influence construct—yielded a slightly higher silhouette value (0.24) and lower inertia (295.49), confirming that the solution structure remains stable and not dominated by any single dimension. Overall, the numerical indices support the interpretability and stability of the four-cluster configuration, validating its use for policy and behavioral analysis.

## 5. DISCUSSION

The cluster analysis revealed four distinct archetypes of Generation Z users of gamified e-government initiatives: Super-Scanners, Socially-Influenced Actives, Casual Participants, and Skeptical Habituals. This segmentation provides a richer understanding of adoption than regression-based models by capturing heterogeneity in motivations, trust, and behaviors. The findings align with broader literature on gamification and user typologies, while extending the discussion into the public sector domain.

**Super-Scanners.** This archetype represents the most engaged and loyal users. They consistently perform the required actions (e.g., scanning all receipts in fiscal gamification or fully engaging in digital compliance mechanisms), exhibit high trust in the initiative’s usefulness, and report strong behavioral intention to continue. Their profile corresponds to “Achievers” or “Players” in gamification typologies (Marczewski, 2015; Tondello *et al.*, 2016), driven by both tangible rewards and intrinsic satisfaction. Demographically, this group is often dominated by younger female participants, confirming prior findings that women in Gen Z are highly responsive to gamified civic initiatives. From a policy perspective, Super-Scanners are valuable advocates: their positive experiences can be leveraged in peer campaigns, testimonials, or referral programs to expand adoption beyond early adopters.

**Socially-Influenced Actives.** These users also demonstrate high engagement but are primarily motivated by peer endorsement and community norms. Their behavior resonates with theories of social proof and the “Socializers” archetype in gamification literature (Koivisto and Hamari, 2014). Importantly, this group’s participation is contingent: it persists only as long as peer networks and social environments reinforce it. This highlights the importance of network effects in e-government adoption, suggesting that governments should reinforce community-driven campaigns, such as university challenges or influencer partnerships, to maintain momentum.

**Casual Participants.** This group consists of sporadic users who engage inconsistently and report lower intention to sustain use. Their engagement appears pragmatic rather than committed, reflecting individuals who may recognize the value of the initiative but lack consistent motivation. In segmentation literature, such users often correspond to “light users” or “browsers” (Wedel and Kamakura, 2000). For Gen Z, this sporadic behavior reflects a pattern observed in other digital domains: while digital natives are quick to try new applications,

sustained engagement depends on continuous reinforcement and convenience (Priporas *et al.*, 2017). Policy interventions for this group should focus on nudges and simplification—for example, gamified challenges, reminders, or improved user interface features.

**Skeptical Habituals.** The most paradoxical archetype consists of users who report frequent participation but express low trust in the initiative's value and limited social influence. Their skepticism may stem from broader institutional distrust (Frey and Torgler, 2007), aligning with research on the digital shadow economy that found Gen Z balances compliance with pragmatic disengagement (Kiselicki *et al.*, 2024). In gamification theory, these users resemble “Disengaged Players” who comply out of habit or obligation rather than motivation (Marczewski, 2015). Their presence represents both a risk and an opportunity: while their current compliance is high, it may collapse if incentives weaken. Policymakers should therefore invest in trust-building strategies, such as transparent communication of impact, improved institutional responsiveness, and user feedback integration.

Taken together, these four archetypes demonstrate that Generation Z is not a monolithic group of adopters, but a heterogeneous population with distinct motivational drivers. This segmentation extends existing gamification and e-government research by moving beyond average effects of predictors to reveal user subgroups with different needs. Importantly, the findings show that high adoption rates may mask underlying differences: while some users are loyal advocates, others comply reluctantly or inconsistently.

This study's findings on behavioral diversity extend beyond the MojDDV case, directly addressing broader discussions in e-government regarding digital trust, compliance behavior, and informality. In numerous emerging economies, diminished institutional trust and enduring informal practices undermine the efficacy of digital public services (Frey and Torgler, 2007; Medina and Schneider, 2018). The segmentation results show that trust is still a fragile and differentiating factor, even among digital natives, who are often thought to be open to new ideas. For example, the existence of Skeptical Habituals demonstrates that ongoing participation cannot depend solely on gamified incentives; instead, it relies on perceptions of fairness, transparency, and reciprocity between citizens and institutions. The results also show how gamification can be used as a way to get people to comply in places with big shadow economies. Initiatives like MojDDV change compliance from a boring, bureaucratic task into a fun, peer-supported activity. This goes along with recent calls to add behavioral and digital tools to enforcement-based fiscal policies to encourage voluntary compliance through positive reinforcement. However, as the archetypes illustrate, these interventions must be sensitive to trust: in contexts of diminished institutional credibility, extrinsic incentives may promote short-term engagement but inadequately establish long-term legitimacy.

From a policy standpoint, these insights enhance global e-government dialogues regarding the equilibrium between behavioral design and the cultivation of institutional trust. Policymakers can adjust their actions based on both the type of motivation and the level of trust users have in government systems if they understand users through segmentation. This is especially important for societies that are moving from informal to digital formalities, where compliance depends on how fair and efficient the system seems to be as well as how easy it is to use. The MojDDV experience teaches us how governments can use gamification to not only make compliance more digital, but also more human by linking fiscal responsibility to everyday digital habits and values that are passed down from one generation to the next.

From a theoretical standpoint, the study contributes by showing how segmentation methods from e-business analytics can enrich e-government research. Rather than relying solely on TAM or UTAUT, clustering provides a complementary lens that highlights behavioral diversity and reveals hidden risks to sustainability. Practically, the identification of archetypes offers actionable insights for policymakers. By tailoring strategies to each user type, advocacy campaigns for Super-Scanners, community initiatives for Socially-Influenced Actives, nudges for Casual Participants, and trust-building for Skeptical Habituals, governments can move toward personalized e-government services, enhancing both adoption and long-term compliance.

## **6. CONCLUSION**

This study addressed two research questions: (RQ1) What behavioral archetypes exist among Gen Z users of gamified e-government initiatives? and (RQ2) How do these archetypes differ in perceptions, behaviors, and demographics? By applying clustering analysis to survey data from 119 respondents, four distinct user archetypes were identified: Super-Scanners, Socially-Influenced Actives, Casual Participants, and Skeptical Habituals. These archetypes highlight the heterogeneity within Generation Z, challenging the assumption that young citizens are a uniform group of adopters.

The findings show that while Super-Scanners and Socially-Influenced Actives exhibit high engagement, their motivations differ, one group is driven by trust and loyalty, the other by peer norms. Casual Participants demonstrate sporadic usage and require additional incentives, while Skeptical Habituals illustrate the paradox of high compliance coupled with low trust. Demographic profiling further revealed gendered patterns, with women dominating the more engaged clusters, and older respondents disproportionately represented among Skeptical Habituals.

By identifying these archetypes, the study makes both theoretical and practical contributions. Theoretically, it extends gamification and e-government research by demonstrating the value of segmentation approaches over linear adoption models, offering a more nuanced understanding of citizen engagement. Practically, it provides actionable insights for policymakers to design adaptive engagement strategies: leveraging advocates (Super-Scanners), reinforcing peer campaigns (Socially-Influenced Actives), applying targeted nudges (Casual Participants), and fostering institutional trust (Skeptical Habituals).

The study's limitations include its focus on Generation Z in North Macedonia, reliance on self-reported data, and a cross-sectional design that precludes tracking changes over time. Future research should test the generalizability of these archetypes across other age groups, cultural contexts, and types of gamified public services. Combining survey data with behavioral logs or experimental approaches could further validate the model and capture how user profiles evolve as digital trust and institutional credibility develop.

In essence, this study underscores that gamified e-government initiatives cannot be understood solely through average effects. Recognizing the behavioral diversity among citizens, particularly younger generations, enables governments to design personalized, trust-sensitive, and sustainable digital public services that foster both compliance and long-term engagement.

## **REFERENCES**

Akademik.mk (2021), *Isplata na parichni sredstva za MojDDV*.

- Aljovic, A. (2024), "Digital natives in e-learning – interaction and engagement", *Conference: Language, Literature and Education - Joining Forces to Build a Better World, Sarajevo, Bosnia and Herzegovina*.
- Alm, J. and Torgler, B. (2011), "Do ethics matter? Tax compliance and morality", *Journal of Business Ethics*, Vol. 101 No. 4, pp. 635-651.
- Andersson, C., Hallin, A. and Ivory, C. (2022), "Unpacking the digitalization of public services: configuring work during automation in local government", *Government Information Quarterly*, Vol. 39 No. 1.
- Deterding, S., Dixon, D., Khaled, R. and Nacke, L. (2011), "From game design elements to gamefulness: defining gamification", *Proceedings of the 15th International Academic MindTrek Conference*.
- Dolot, A. (2018), "The characteristics of Generation Z", *e-mentor*, Vol. 74 No. 2, pp. 44-50.
- Frey, B.S. and Torgler, B. (2007), "Tax morale and conditional cooperation", *Journal of Comparative Economics*, Vol. 35 No. 1, pp. 136-159.
- Hamari, J., Koivisto, J. and Sarsa, H. (2014), "Does gamification work? A literature review of empirical studies on gamification", *47th Hawaii International Conference on System Sciences (HICSS)*, pp. 3025-3034.
- Hassan, L. and Hamari, J. (2020), "Gameful civic engagement: a review of the literature on gamification of e-participation", *Government Information Quarterly*, Vol. 37 No. 3, p. 101461.
- Josimovski, S., Pulevska Ivanovska, L., Mijoska, M. and Kiselicki, M. (2024), "Gender differences in Gen Z adoption of gamified e-government services: evidence from MojDDV", *Proceedings of EBTSF 2024*.
- Kiselicki, M., Pulevska Ivanovska, L., Josimovski, S. and Mijoska, M. (2024), "Consumer perceptions of the digital shadow economy: Gen Z perspectives on informality", *International Journal*, Vol. 66 No. 1, pp. 169-174.
- Koivisto, J. and Hamari, J. (2014), "Demographic differences in perceived benefits from gamification", *Computers in Human Behavior*, Vol. 35, pp. 179-188.
- Marczewski, A. (2015), *Even Ninja Monkeys Like to Play: Gamification, Game Thinking and Motivational Design*, CreateSpace Independent Publishing Platform.
- Matveyeva, A. and Krasnov, R. (2019), "Generation 'Z' as a generation for market of goods and services in digitalization", *Advances in Economics, Business and Management Research*, Vol. 105.
- Medina, L. and Schneider, F. (2018), "Shadow economies around the world: what did we learn over the last 20 years?", *International Tax and Public Finance*, Vol. 25 No. 3, pp. 641-681.
- Mekler, E.D., Brühlmann, F., Tuch, A.N. and Opwis, K. (2017), "Towards understanding the effects of individual gamification elements on intrinsic motivation and performance", *Computers in Human Behavior*, Vol. 71, pp. 525-534.
- Nunnally, J.C. (1978), *Psychometric Theory*, 2nd ed., McGraw-Hill, New York, NY.
- Priporas, C.V., Stylos, N. and Fotiadis, A.K. (2017), "Generation Z consumers' expectations of interactions in smart retailing: a future agenda", *Computers in Human Behavior*, Vol. 77, pp. 374-381.
- Public Revenue Office (2022), "Се зголемува лимитот за MojDDV, за три години вративме 4,2 милјарди денари кај граѓаните", available at: <http://www.ujp.gov.mk/mk/javnost/soopstenija/pogledni/921>
- Seemiller, C. and Grace, M. (2016), *Generation Z Goes to College*, Jossey-Bass.
- Tondello, G.F., Wehbe, R.R., Diamond, L., Busch, M., Marczewski, A. and Nacke, L.E. (2016), "The gamification user types Hexad scale", *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, pp. 229-243.

- Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003), "User acceptance of information technology: toward a unified view", *MIS Quarterly*, Vol. 27 No. 3, pp. 425-478.
- Wedel, M. and Kamakura, W.A. (2000), *Market Segmentation: Conceptual and Methodological Foundations*, 2nd ed., Kluwer Academic Publishers.
- Williams, P. (2002), "The competent boundary spanner", *Public Administration*, Vol. 93 No. 1, pp. 212-226.

## **AI AS A NEW ACADEMIC ALLY: THE IMPACT OF ARTIFICIAL INTELLIGENCE ON STUDENTS' LEARNING HABITS**

**Violeta Cvetkoska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
violeta.cvetkoska@eccf.ukim.edu.mk*

**Filip Peovski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
peovski@eccf.ukim.edu.mk*

**Gjore Gakjev**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
gakjev07@gmail.com*

**Binela Karamaleska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
binellakaramaleska@gmail.com*

**Elena Avramovikj**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
elenaavramovic246@gmail.com*

**Lina Taneska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
lina.taneska26@gmail.com*

**Aleksandra Peshova**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
apesova38@gmail.com*

### **ABSTRACT**

*This study investigates the adoption and usage of ChatGPT by college students in educational settings. The analysis uses a four-stage business analytics framework to look at usage, trust, confidence, motivation, acceptance, and verification patterns using survey data from 203 respondents in a variety of disciplines. The findings highlight ChatGPT's function as a tool for improving comprehension and self-assurance by demonstrating that the three strongest predictors of frequent use are understanding, trust, and confidence. The tension between critical evaluation and reliance on AI is highlighted by the fact that motivation plays a secondary role, and verification is largely irrelevant and negatively associated with trust. According to the research, generative AI works best when viewed as an academic ally that promotes learning and introspection rather than taking the place of critical thinking. The study*

*provides context-bound findings that inform hypotheses for larger cross-institutional and cross-national research because of its single-country sample. The paper highlights recommendations to universities to foster AI literacy, safeguard the crucial academic integrity, and integrate ChatGPT into teaching practices responsibly and effectively.*

**Keywords:** *Artificial Intelligence, ChatGPT, Learning habits, Business analytics, Higher education*

**JEL classification:** *C83, I21, M15*

## **1. INTRODUCTION**

Higher education has quickly adopted artificial intelligence (AI), which is altering how students engage with course materials and approach homework. One of the most widely used tools is ChatGPT, a generative language model that can produce context-aware, human-like responses to open-ended questions. It is a helpful resource for students seeking explanations, summaries, or assignment assistance due to its adaptability and accessibility. Although ChatGPT and other AI tools have the potential to improve learning and foster academic independence, according to international research, they also highlight problems with misleading information, over-reliance, and ethical ambiguity. Students may benefit from increased productivity and better comprehension, but these advantages must be weighed against the potential for them to lose interest in critical thinking and self-directed learning. Furthermore, not much research has been done using primary data from local contexts, especially among undergraduate students in transitional education systems, despite growing interest in the global academic community.

In order to close that gap, this study examines how students use natural language processing (NLP) models such as ChatGPT, paying particular attention to important elements like usage frequency, motivation, understanding, trust, confidence, and verification of the information provided. Planning, using strategies, monitoring and verifying information, asking for help, and managing time are all examples of self-regulated, recurring study practices that we define as learning habits (Zimmerman, 1990; Verplanken and Orbell, 2003). The use of ChatGPT in these routines to scaffold comprehension and confidence while maintaining critical evaluation is referred to in this study as AI-mediated learning habits. The study employs a four-phase business analytics framework, i.e., prescriptive (producing strategic recommendations), diagnostic (analyzing underlying drivers), predictive (modeling usage patterns), and descriptive (identifying prevalent attitudes). Subsequently, we intend to provide a grounded perspective on how AI is changing learning strategies and expectations in higher education by surveying 203 students from a variety of academic programs in the Republic of North Macedonia.

The overall findings of this study show that a variety of cognitive, psychological, and behavioral factors influence college students' use of ChatGPT. Students primarily use ChatGPT to improve their understanding of difficult academic material and to boost their confidence in finishing assignments. These two constructs consistently show up as the strongest predictors of frequent usage. Additionally, trust is crucial because it mediates the link between behavioral engagement and cognitive gains, thereby strengthening adoption. On the other hand, motivation, while still important in early models, loses significance when confidence and trust are taken into account, suggesting that zeal by itself cannot support sustained use. Despite its academic value, verification has no discernible impact on adoption and even has a negative correlation with trust, indicating a latent conflict between critical analysis and AI dependence.

Strong relationships between trust, acceptance, understanding, and confidence are further revealed by correlation patterns, highlighting the fact that students use ChatGPT most successfully when they perceive both cognitive and affective benefits. When combined, these results imply that generative AI is more than just a handy tool; rather, it is an academic ally that enhances understanding and confidence, so long as its application is framed by ethical responsibility and critical engagement.

We systematize the paper in the following manner. In the following Section 2, we review some of the most notable global studies related to the topic that inspired the research. Section 3 presents the methodological approach, while in Sections 4 and 5 we present and discuss the obtained results based on the four main pillars of business analytics. Finally, we conclude the research in the last section.

## **2. LITERATURE REVIEW**

Numerous studies have shown that artificial intelligence (AI) significantly affects students' motivation and involvement in academic activities. By offering individualized learning experiences that take into account each student's preferences and skills, artificial intelligence (AI) tools like learning analytics and intelligent tutoring systems have been demonstrated to increase student motivation and engagement (Elbadiansyah *et al.*, 2024; Wadhwa *et al.*, 2024). These resources enhance self-efficacy and problem-solving skills in addition to motivation, which improves academic performance (Jor, 2025). Through interactive and adaptive learning experiences, ChatGPT and other AI applications have been shown to dramatically increase student motivation and engagement in online learning environments, keeping students actively engaged in their studies (Rehman and Kang, 2024). Additionally, gamification components powered by AI add enjoyment and rewards to the educational process, which raises motivation and engagement even more (Wadhwa *et al.*, 2024). AI integration in education is not without its difficulties, though. There are many worries about an over-reliance on AI, a decline in creativity, and moral dilemmas like justice and privacy (Elbadiansyah *et al.*, 2024; Singh, 2024). Furthermore, concerns about academic dishonesty and how AI might affect teacher-student interactions need to be addressed (Jor, 2025). Nevertheless, if AI is applied carefully and morally, it can have significant positive effects on education overall, especially in raising student motivation and engagement (Singh, 2024; Jor, 2025).

Established technology-adoption frameworks like the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Technology Acceptance Model (TAM) serve as the foundation for the integration of artificial intelligence (AI) into higher education. Prior to ChatGPT, early empirical research showed that students' willingness to interact with AI-driven learning environments is influenced by perceived utility, ease of use, and facilitating conditions (Strzelecki, 2023; Lin and Yu, 2023). These ideas are extended to large language models (LLMs) by Kasneci *et al.* (2023), who contend that explainability, transparency, and pedagogical alignment mediate acceptance. Along with emphasizing that AI tools co-construct learning practices rather than just automate cognitive tasks, theoretical perspectives also emphasize sociocultural and ethical lenses (Mhlanga, 2023). Since its public release, ChatGPT has rapidly spread across a variety of national contexts, according to survey-based investigations. More than 60% of Polish undergraduates had used ChatGPT within three months, mostly for brainstorming essays and summarizing readings, according to Strzelecki (2023). While confidence in disciplined use remained uneven across faculties, Chan and Hu (2023) also found high familiarity among Hong Kong students. Perceived comprehension gains and intrinsic learning motivation are strong predictors of sustained usage, according to

multivariate analyses (Vieriu and Petrea, 2025). Notably, the effects of demographic moderators like gender and study level are inconsistent, indicating the need for more detailed, context-sensitive models (Lin and Yu, 2023). Evidence on learning effectiveness is emergent but promising. In controlled experiments, Chen and Gong (2025) demonstrated that compared to peers in a traditional workshop, international students who received writing feedback mediated by ChatGPT reported stronger self-regulation and significantly higher scores. Sykes's (2024) complementary findings show that AI-augmented critique tasks enhanced lexical richness and argumentative structure without reducing authorial voice. However, meta-analytic synthesis shows that effect sizes vary widely, and that learning gains are diminished when students lack critical prompting skills or when tasks emphasize rote knowledge (Kasneci *et al.*, 2023).

Scholars document serious concerns about academic integrity in addition to the pedagogical benefits. Susnjak & McIntosh (2024) revived discussions on assessment design by proving that ChatGPT-generated responses can avoid human marking and traditional plagiarism detection. Extensive surveys reveal students' ambivalence: although the majority recognize efficiency gains, less than one-third regularly check the accuracy of facts (Strzelecki, 2023), reiterating concerns about taking AI results for granted. Universities are urged by regulatory and ethical analyses to create clear usage guidelines, encourage AI literacy, and foster critical data practices (Mhlanga, 2023).

We identify that there are three main conclusions drawn from the literature upon which we build our study: 1) students accept ChatGPT well, but only if it is conditioned by perceived cognitive value and ethical clarity; 2) well-scaffolded AI use can improve writing quality, conceptual understanding, and engagement; and 3) unmanaged reliance risks compromising integrity and eroding critical thinking. However, there are still unanswered questions about the socio-emotional aspects of AI-mediated research, disciplinary variations in prompt literacy, and long-term learning paths. To capture how changing LLM capabilities reshape higher-education ecosystems, future research should use intersectional lenses and longitudinal mixed-method designs. We found that understanding reflects perceived usefulness (TAM), trust reflects technology-trust beliefs, confidence reflects self-efficacy in self-regulated learning, motivation follows Self-Determination Theory, verification reflects metacognitive monitoring/need for cognition, and usage frequency reflects actual use (Davis, 1989; Zimmerman, 1990; Ryan and Deci, 2000). This contributes to the alignment of the established constructs in our study with the theoretical models developed.

### **3. METHODOLOGY**

The research team created and distributed an anonymized online survey to find out how students feel about and behave when it comes to using artificial intelligence (AI) in classroom settings. Targeting university students from a wide range of academic programs, disciplines, and years of study, the survey tool was created using Google Forms. Convenience and snowball sampling methods were used for recruitment as well as peer networks. Because participation was voluntary and anonymous, a wide range of responses were guaranteed, social desirability bias was reduced, and ethical standards of research integrity were maintained. The analytical sample consisted of 203 valid responses in total. Eleven of the fifteen structured items on the questionnaire used a five-point Likert scale, with “strongly disagree” to “strongly agree” being the extremes. Students' perceived understanding of AI outputs, motivation to interact with AI, frequency of use, trust in AI systems, and inclination to verify information generated by AI were all measured by these items. To enable subgroup comparisons and the investigation of

response heterogeneity, four demographic questions were also added, such as gender, age, academic field, and year of study. Internal consistency checks were performed on the survey items, which were drawn from well-established theoretical discussions on technology adoption and trust in digital systems in order to enhance construct validity. All things considered, the goal of this methodological approach was to match behavioral insights with measurable patterns so that recommendations based on empirical data could be developed to improve student engagement and direct institutional policies regarding the use of AI in higher education.

The four main phases of business analytics, descriptive, diagnostic, predictive, and prescriptive analysis, were followed when conducting the data analysis. First, descriptive statistics were used to find broad trends in the data and to summarize the demographic profile of the respondents. The linear relationships between important variables were then evaluated using correlation analysis, with particular focus on the interdependencies among motivation, comprehension, trust, and verification. This study used single-item indicators for key constructs, as mentioned, even though multi-item scales are typically better at capturing construct breadth and facilitating reliability estimates. The exploratory nature of the study and the requirement to reduce respondent burden in a student sample served as the basis for this decision. However, previous studies have shown that when constructs are concrete, unidimensional, and simple for respondents to understand, single-item measures can yield valid and reliable assessments. A single global job satisfaction item, for instance, showed acceptable validity and comparability with multi-item scales, according to Wanous *et al.* (1997). In the same manner, Bergkvist and Rossiter (2007) demonstrated that in applied contexts, single-item measures of attitude constructs can attain predictive validity comparable to multi-item scales. This method permits effective data collection and yields suggestive results that can be expanded in subsequent studies using validated multi-item scales, despite the notable trade-offs, which may lead to the inability to report internal consistency. Consequently, key perceptions were operationalized with single items because this was an exploratory survey limited by respondent burden. Since internal-consistency estimation and broader construct coverage are made possible by multi-item scales, we treat the results as indicative and broaden the study's limitations accordingly. Internal consistency metrics like Cronbach's alpha could not be reported because the constructs were measured using single items, and there was no test-retest data available. For concrete constructs, single-item measures may be justified; however, this method has limitations in terms of coverage and reliability. Furthermore, despite efforts to mitigate common-method bias through anonymity and neutral item wording, it still exists in all self-report surveys (Podsakoff *et al.*, 2003).

After that, linear regression was used to apply predictive modeling, and the impact of explanatory variables on outcome measures was measured using Ordinary Least Squares (OLS) estimation. Lastly, the results were interpreted, and useful suggestions for integrating AI in academic settings were derived using what is usually considered prescriptive reasoning. For instance, the three-variate OLS regression described below was used to model the predictive relationship between the dependent variable (frequency of usage) and two explanatory factors, namely motivation and understanding:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

where with  $Y_i$  we denote the dependent variable, i.e., the trust level of respondents ( $i$ ), with  $X_{1i}$  and  $X_{2i}$  we denote the independent variables – motivation and understanding, respectively. The intercept is denoted with  $\beta_0$ , while the slope coefficients for each of the two independent

variables with  $\beta_1$  and  $\beta_2$ . Finally, the error term ( $\varepsilon_i$ ) is assumed to satisfy the Gauss-Markov conditions.

#### 4. ANALYSIS AND RESULTS

##### 4.1. Descriptive analysis

We used descriptive statistics in the first stage of the analysis to get a general idea of how students felt and behaved with regard to ChatGPT. To determine prevailing patterns and participant consensus, measures of central tendency, variability, and the shape of the distribution were computed for every Likert-scale question. Depending on the context of the construct, we label “Using ChatGPT increases my self-confidence when completing academic tasks” as a construct variable for confidence, „Before using information obtained from ChatGPT, I check the accuracy through other sources (e.g. textbook, Google Scholar)“ as a construct for verification, „ChatGPT helps me understand complex academic material“ for understanding, „Using ChatGPT increases my motivation to study independently (without help from other people)“ as a construct for motivation, and „I fully trust the answers without verification” as a construct for observing trust. Subsequently, we mainly focus on these four constructs later in the study.

Students see ChatGPT as a useful tool for improving their academic comprehension, as evidenced by the high mean score of 4.06 for the item “ChatGPT helps me understand complex academic material.” Furthermore, a perceived psychological advantage is indicated by the statement, “Using ChatGPT increases my self-confidence when completing academic tasks” ( $\mu = 3.59$ ). Lower scores, however, for statements like “I fully trust the answers without verification” ( $\mu = 3.21$ ), indicate that students should approach AI-generated content critically and cautiously. The detailed descriptive statistics are presented in Table 1.

In the end, the descriptive analysis reveals a balanced user mindset, where students are aware of ChatGPT's limitations and the need for additional testing and development, but they are also willing to use it for academic purposes. These findings set the stage for a more in-depth investigation of how and why students engage in specific behaviors in the subsequent stages of analysis. Instead of passively accepting ChatGPT's outcomes, students use it primarily critically and constructively, as evidenced by the distributions' shape and variability. The majority of constructs show moderate negative skewness, indicating a tendency to concur with claims regarding self-directed use and critical evaluation. The overall pattern suggests that users will continue to use ChatGPT as a tool to improve understanding and independence while exercising caution and using its products sparingly.

Table 1: Descriptive statistics of the Likert-scaled constructs.

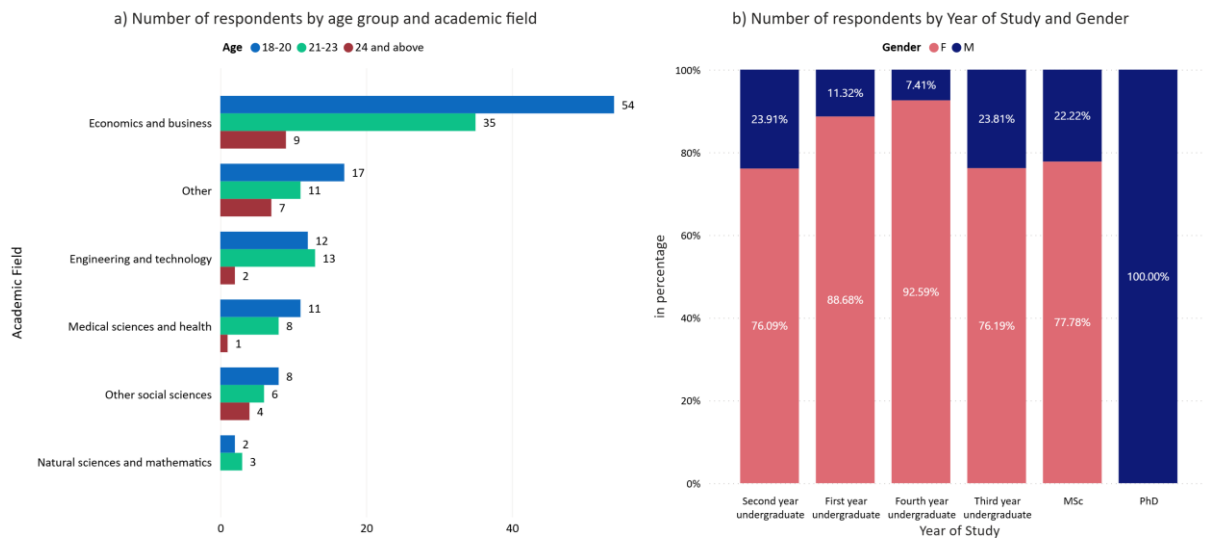
Construct	$\mu$	St. Error	Me	Mo	$\sigma$	$\sigma^2$	Excess Kurtosis	Skewness
„I fully trust the answers without verification”	3.241	0.077	3	3	1.093	1.194	-0.277	-0.309
„Before using information obtained from ChatGPT, I check the accuracy through other sources (e.g., textbook, Google Scholar)“	3.212	0.095	3	4	1.357	1.841	-1.130	-0.247
“Using ChatGPT increases my self-confidence when completing academic tasks”	3.586	0.080	4	3	1.142	1.303	-0.252	-0.567
„ChatGPT helps me understand complex academic material“	4.064	0.080	4	5	1.135	1.288	0.983	-1.276

„After using ChatGPT, I am less likely to ask for additional help from colleagues or professors“	3.655	0.087	4	5	1.239	1.534	-0.326	-0.724
„Using ChatGPT increases my motivation to study independently (without help from other people)“	3.512	0.094	4	5	1.340	1.796	-0.861	-0.525

(Source: Authors' calculations (n = 203))

Most of the respondents sampled are second year undergraduate studies (92), followed by first year undergraduates (53), fourth (27), and third year (21). The distribution across academic fields, age, and gender shows that most of them (consistently above 75% in first and second cycle studies) are female and enrolled in the field of economics and business (see Figure 1). Approximately 89% of respondents are aged between 18 and 23, with the first group aged 18-20 contributing to 51.2% of the entire sample.

Figure 1a and 1b: Demographics of the sampled respondents.



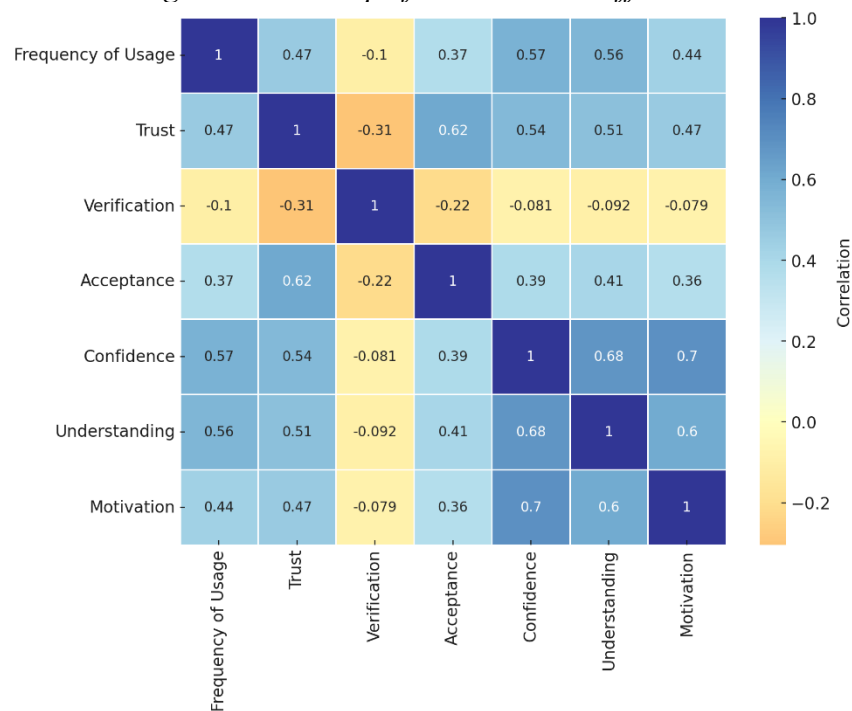
(Source: Authors' calculations (n = 203))

#### 4.2. Diagnostic analysis

In the diagnostic phase, we applied correlation analysis to examine relationships between key academic variables. The correlation heatmap in Figure 2 offers a deeper comprehension of how students' attitudes, behaviors, and perceptions interact with regard to using ChatGPT for academic purposes. A complex picture of how cognitive, motivational, and affective constructs influence engagement with generative AI is provided by the numerous strong and moderately positive correlations that show up along with a few weak or negative associations. First, Frequency of Usage is strongly associated with Understanding ( $r = 0.556$ ) and Confidence ( $r = 0.574$ ), with Trust ( $r = 0.468$ ) coming in second. This trend suggests that when students feel confident in their academic skills while using ChatGPT and believe it to be a tool that improves comprehension, they are more likely to use it regularly. Although its impact is less pronounced than that of confidence and understanding, trust serves to further reinforce this tendency. Additionally, motivation and usage have a positive correlation ( $r = 0.437$ ), indicating that willingness to learn on one's own increases use frequency, albeit as a secondary driver. Verification, on the other hand, has a negative correlation with frequency ( $r = -0.104$ ), suggesting that students who routinely verify ChatGPT's responses use it less frequently.

Second, when looking at inter-construct relationships, we find that Trust and Acceptance have a strong correlation ( $r = 0.616$ ), meaning that students who are more likely to accept ChatGPT's outputs also have higher levels of confidence in its dependability. Additionally, there is a strong correlation between trust and understanding ( $r = 0.514$ ) and confidence ( $r = 0.537$ ), highlighting the role that self-assurance and cognitive gains play in the development of tool confidence. Interestingly, motivation has strong positive relationships with both understanding ( $r = 0.603$ ) and confidence ( $r = 0.696$ ), indicating that motivated learners are more likely to feel competent and perceive cognitive benefits when using ChatGPT. Third, the counter-factor that always sticks out is verification. It has a weak negative correlation with Understanding ( $r = -0.092$ ), Motivation ( $r = -0.079$ ), Acceptance ( $r = -0.216$ ), and Trust ( $r = -0.305$ ). This highlights a possible conflict: students who place a strong emphasis on verification might continue to have doubts about ChatGPT, which would reduce acceptance and trust. In this context, critical evaluation seems to lessen habitual reliance on AI tools, even though it is desirable academically. These findings point to three key positive predictors that together propel ChatGPT's incorporation into students' daily learning routines: trust, understanding, and confidence. While verification creates friction by reducing reliance and trust, motivation indirectly reinforces these dynamics.

Figure 2: Heatmap of correlation coefficients.



(Source: Authors' calculations (n = 203))

### 4.3. Predictive analysis

To determine which factors most significantly influence the frequency of ChatGPT usage among students, a multiple linear regression analysis was conducted. The dependent variable was self-reported usage frequency, measured on a five-point Likert scale, while the independent variables were perceived understanding of academic material, motivation for independent learning, trust in the content provided by ChatGPT, the need for verification of output, and the perceived individual confidence in learning without additional help from professors and peers. To better understand what influences the frequency of ChatGPT use for academic purposes, the regression analysis was organized into four nested models, each of

which progressively added more explanatory variables. Students' self-reported frequency of use served as the dependent variable, and all constructs were scored on a five-point Likert scale.

Two initial core factors - understanding and motivation were included in Model 1 (see Table 2). With a  $\beta_1$  of 0.4550, the results unmistakably showed that understanding was the primary factor driving adoption among students. This suggests that with each small increase in perceived understanding, students who believe ChatGPT to be a useful tool for improving comprehension use it much more frequently. Despite being positive and significant ( $\beta_2 = 0.1333$ ), motivation had a less pronounced effect, suggesting that it plays a supporting role in adoption. About one-third of the variation in usage can be explained by these two psychological constructs, according to the coefficient of determination of 0.3252 and adjusted  $R^2$  of 0.3185. This can be considered a significant share for behavioral research that uses perceptual scales.

The explanatory landscape changed with the introduction of the factor Trust in Model 2. Students who have a higher trust in ChatGPT are significantly more likely to use it regularly, as indicated by the coefficient for trust, which was 0.2315. Simultaneously, the Understanding (0.4373) and Motivation (0.0870) coefficients showed a slight decrease, indicating that the trust dimension may mediate or overlap with their effect overall. With  $R^2$  increasing to 0.3606, the model's explanatory power increased, indicating that the inclusion of trust improves the regression's predictive ability and captures significant relational aspects of AI engagement.

In response to students' propensity to double-check ChatGPT's responses, Model 3 added Verification to its specification. However, the verification coefficient ( $\beta_4 = 0.0076$ ) is statistically insignificant and negligible, indicating that students' frequency of use of ChatGPT is not systematically impacted by their fact-checking practices. This may indicate that students in fact use it independently of the quality of information provided, creating knowledge distortion and bias. Crucially, motivation stayed low (0.0866), but understanding (0.3728) and trust (0.2349) continued to have significant and favorable effects. The adjusted  $R^2$  of 0.3478 demonstrated minimal improvement over the prior model, highlighting the low contribution of verification to adoption frequency explanation.

Finally, Model 4 introduced Confidence, which is characterized as students' self-assurance when using ChatGPT to complete academic assignments. The dynamics of the model were significantly changed by its inclusion. Confidence showed a strong slope coefficient  $\beta_5$  of 0.3139, indicating that students use ChatGPT much more frequently after they feel competent and empowered to produce quality academic work. In contrast, Motivation completely lost significance (and changing the sign of impact to -0.0216), while Understanding (0.2629) and Trust (0.1755) continued to have positive effects, albeit with somewhat diminished effects when compared to previous models. Once more, which was desirable to confirm the previously set thesis, verification had no statistically distinct impact (-0.0027). Verification again had no meaningful effect (-0.0027). The explanatory power of the model increased further, with adjusted  $R^2$  to 0.3849, meaning that the full set of predictors accounts for approximately 40% of the variation in ChatGPT usage, which is a considerable share for behavioral models in social science research. An important realization is demonstrated by the models' progressive increase in explanatory power, where although cognitive factors (understanding) serve as the foundation for adoption, relational (trust) and psychological (confidence) variables significantly enhance the explanatory framework. Although motivation is important at first, it becomes unnecessary when trust and confidence are taken into account. This suggests that students' perceived competence and confidence in their ability to use AI effectively are more important than intrinsic drive. While students may verify information, this behavior does not

systematically determine adoption levels. In contrast, verification is consistently irrelevant in predicting the frequency of use. The results highlight that regular and intentional use of ChatGPT for academic purposes for students is based on a so-called triad of understanding, trust, and confidence rather than being merely motivated by enthusiasm or habit. These findings note the significance of instructional approaches that develop an understanding of AI's advantages and disadvantages, establish confidence in its proper use, and, moreover, encourage students to incorporate AI into their daily academic routines.

Table 2: Regression model

Variable	Model 1	Model 2	Model 3	Model 4
Constant	1.2147*** (0.2453) [4.9521]	0.9589*** (0.2515) [3.8131]	0.9272*** (0.3263) [2.8417]	0.8535*** (0.3175) [2.6879]
Understanding	0.4550*** (0.0720) [6.3163]	0.4373*** (0.0745) [5.0121]	0.3728*** (0.0748) [4.9862]	0.2629*** (0.0788) [3.3386]
Motivation	0.1333** (0.0610) [2.1839]	0.0870 (0.0612) [1.4227]	0.0866 (0.0614) [1.4114]	-0.0216 (0.0667) [-0.2336]
Trust		0.2315*** (0.0697) [3.3189]	0.2349*** (0.0734) [3.2014]	0.1755** (0.0731) [2.3996]
Verification			0.0076 (0.0496) [0.1529]	-0.0027 (0.0482) [-0.0555]
Confidence				0.3139*** (0.0872) [3.5933]
AIC	2.7003	2.6563	2.6660	2.6122
R <sup>2</sup>	0.3252	0.3606	0.3607	0.4002
Adj. R <sup>2</sup>	0.3185	0.3510	0.3478	0.3849
St. Error	0.9267	0.9043	0.9066	0.8804
Obs.	203	203	203	203

Note: \*\*\*, \*\*, and \* indicate 1%, 5%, and 10% statistical significance, respectively.

(Source: Authors' calculations)

## 5. PRESCRIPTIVE ANALYSIS AND DISCUSSION

A dual imperative for higher education is highlighted by the integration of diagnostic and predictive results, i.e., to minimize the risks of uncritical dependence while utilizing the cognitive and motivational advantages of generative AI. According to the correlation analysis, usage frequency is most strongly associated with understanding and confidence, then with trust. This implies that students use ChatGPT not just for convenience but also because it improves understanding and increases confidence in finishing assignments. This finding was supported by regression models, which showed that the three most reliable indicators of continued use were understanding, trust, and confidence. When confidence and trust were taken into account, motivation's explanatory power diminished, indicating that intrinsic drive is eventually absorbed into larger cognitive and psychological processes. On the other hand, acceptance showed up in correlations as being closely associated with trust but lacking independent predictive power, whereas verification consistently failed to predict usage. This pattern has significant implications. First and foremost, ChatGPT ought to be purposefully positioned as

an aid for understanding and empowerment rather than as a substitute for human intellect. The students who integrate ChatGPT into their learning routines the most frequently are those who see cognitive gains and have faith in their ability to use it successfully. This supports the idea that AI tools actively contribute to students' sense of competence rather than acting as neutral accessories. Therefore, the goal of educational design must be to embed self-regulation and reflection while also scaffolding the use of AI in ways that improve understanding.

The conflict between verification and trust necessitates careful pedagogical consideration. According to correlation results, students who have greater faith in ChatGPT are more likely to verify less, increasing their exposure to false or misleading information. This paradox highlights a larger issue with AI literacy, for instance that skepticism can be undermined by the same confidence that keeps usage going. Verification must become a standard academic practice at universities rather than a self-made choice. Moreover, it can be turned from a usage deterrent into an essential intellectual discipline through structured assignments in which students annotate, critique, or cross-reference ChatGPT's outputs with reliable sources. By doing so, critical engagement could be reframed as an asset rather than a threat to confidence. In order to position ChatGPT as a catalyst for exploration, instructional design needs to change to incorporate hybrid learning strategies. For instance, in flipped classroom settings, students may utilize ChatGPT before lectures to produce draft arguments or explanations that can subsequently be discussed with peers and verified with teachers. AI becomes more of a conversation starter and an exploratory scaffold in this situation rather than a content provider. In a similar manner, gamified or group projects where ChatGPT facilitates brainstorming can connect motivation to creative and interactive use as opposed to rote dependence. These kinds of designs encourage criticism, innovation, and discussion - exactly the traits that guard against over-reliance.

In the end, the results broaden current discussions about academic integrity and digital ethics. Verification's feeble and frequently detrimental function highlights how important standards could be undermined if AI adoption is not controlled. A generation of students who delegate judgment to machines could be fostered in the absence of clear boundaries. Prescriptive tactics must be used in classrooms and institutions to combat this. Universities should establish clear usage guidelines, offer specific instruction in AI literacy, and uphold intellectual integrity standards. Simultaneously, educators need to establish learning environments that value skepticism, independent judgment, and fact-checking as essential academic qualities. As demonstrated by their frequent editing of ChatGPT's outputs, students already demonstrate encouraging instincts for critical engagement, however, in order to avoid complacency, these instincts need to be systematically reinforced.

This study has a number of limitations despite its contributions. Because students may overestimate their understanding or underestimate their reliance on ChatGPT, the use of self-reported survey data raises the possibility of what is called a response bias. Although the sample spans several faculties, it is limited to a single country, which restricts how broadly the results can be applied in global terms. External validity is limited because the study is deemed exploratory, utilizing a convenience sample that is primarily made up of young, female undergraduates studying business and economics in a single national higher education system. To reduce respondent burden, all focal constructs were measured using single items, which may have resulted in lower construct coverage and prevented internal-consistency estimation. Additionally, although a number of predictors were taken into account, important dimensions may have been overlooked because other variables like prior AI exposure, digital literacy, and disciplinary differences were left out. Last but not least, the cross-sectional design records

student behavior at a specific moment in time, making it impossible to draw conclusions about causality or track changing practices. In order to overcome these constraints, future studies should use longitudinal designs that monitor how AI use changes over time, particularly as generative technologies and educational pressures change. Predictive models and interpretations could be improved by adding more variables, such as prompt engineering abilities, field of study, or AI literacy levels. By using validated multi-item scales that can yield more robust construct validity and reliability estimates, future research should build on this exploratory work. Key dimensions are more thoroughly covered by well-established metrics like computer self-efficacy (Compeau and Higgins, 1995), trust in technology constructs (McKnight *et al.*, 2002), and perceived usefulness and ease of use from the TAM - Technology Acceptance Model (Davis, 1989). Similarly, the conceptual precision of concepts like motivation, verification, and learning regulation could be expanded through the incorporation of motivational frameworks from Self-Determination Theory (Ryan and Deci, 2000) and the Need for Cognition Scale (Cacioppo and Petty, 1982). Future research can improve robustness, facilitate cross-study comparability, and offer more profound understandings of how AI tools alter students' learning patterns by utilizing these tools. Furthermore, in addition to the quantitative insights provided here, experimental and mixed-method approaches may uncover the causal mechanisms relating to confidence, verification, and trust. In addition to expanding theoretical knowledge, this kind of research would give academic institutions more useful advice on how to strike a balance between academic integrity and technological efficiency.

## **6. CONCLUSION**

By providing empirical insights into how university students adopt and use ChatGPT for academic purposes, this study adds to the growing body of research on the role of generative artificial intelligence in higher education. Based on a sample of 203 respondents from various academic fields and levels, the analysis revealed a complex interaction between behavioral, affective, and cognitive factors that support usage patterns. The findings show that students are not merely adopting ChatGPT out of novelty or convenience; rather, their engagement is strongly anchored in perceived understanding, relational trust, and self-confidence. These three constructs, which together might be referred to as the “triad of adoption,” have continuously been found to be the most significant motivators of frequent use. Although motivation was important at first, it lost its explanatory power when confidence and trust were added, demonstrating that zeal by itself cannot support sustained use. Verification, on the other hand, although academically desirable, was found to be largely irrelevant as a predictor of adoption and even showed negative associations with trust, indicating that students who regularly fact-check are less likely to use ChatGPT.

When put together, these findings provide a complex picture of the use of AI in educational settings. Although the data shows a latent risk of overconfidence and diminished verification, students seem to be willing to use ChatGPT critically and constructively, editing its outputs and using it as a learning tool. In order to prepare students to strike a balance between confidence and skepticism, higher education institutions must actively incorporate AI-critical literacy into their curricula. Instead of discouraging use, educational approaches that promote annotation, cross-referencing, and triangulation of AI-generated outputs may turn verification into a deeply ingrained academic practice. Additionally, ChatGPT can be made to act as a catalyst for inquiry, contemplation, and discussion rather than a passive answer engine by implementing hybrid pedagogical models like flipped classrooms, gamified projects, or group projects.

Additionally, the study adds to larger discussions about academic integrity and digital ethics. Overregulation may limit the innovative potential of these tools, while an unrestrained reliance on generative AI runs the risk of weakening critical thinking and encouraging complacency in fact-checking. Therefore, the challenge is to create classroom procedures and institutional policies that uphold strict standards of truth, accuracy, and independent thought while acknowledging AI as a valid academic ally. According to the concluding evidence, how educational institutions mold students' perceptions, trust, and confidence in using these tools responsibly will determine the future of AI in education rather than just technological capabilities. ChatGPT and related systems can improve understanding, empower students, and eventually promote a culture of thoughtful and responsible AI integration if they are backed by ethical standards, critical literacy initiatives, and innovative teaching practices.

## REFERENCES

- Bergkvist, L. and Rossiter, J.R. (2007), "The predictive validity of multiple-item versus single-item measures of the same constructs", *Journal of Marketing Research*, Vol. 44 No. 2, pp. 175-184.
- Cacioppo, J.T. and Petty, R.E. (1982), "The need for cognition", *Journal of Personality and Social Psychology*, Vol. 42 No. 1, pp. 116-131.
- Chan, C.K.Y. and Hu, W. (2023), "Students' voices on generative AI: perceptions, benefits, and challenges in higher education", *International Journal of Educational Technology in Higher Education*, Vol. 20 No. 1, p. 43.
- Chen, C. and Gong, Y. (Frank) (2025), "The role of AI-assisted learning in academic writing: a mixed-methods study on Chinese as a second language students", *Education Sciences*, Vol. 15 No. 2, p. 141.
- Compeau, D.R. and Higgins, C.A. (1995), "Computer self-efficacy: development of a measure and initial test", *MIS Quarterly*, Vol. 19 No. 2, pp. 189-211.
- Davis, F.D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS Quarterly*, Vol. 13 No. 3, pp. 319-340.
- Elbadiansyah, Sain, Z.H., Shehu Lawal, U., Chansa Thelma, C. and Luqman Aziz, A. (2024), "Exploring the role of artificial intelligence in enhancing student motivation and cognitive development in higher education", *TechComp Innovations: Journal of Computer Science and Technology*, Vol. 1 No. 2, pp. 59-67.
- Jor, J. (2025), "Impact of artificial intelligence on engagement, motivation, work efficiency, and academic outcomes", in Barstow, C. and Briel, H. (Eds.), *Connecting Ideas, Cultures, and Communities*, 1st ed., Routledge, London, pp. 511-515.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J. and Kasneci, G. (2023), "ChatGPT for good? On opportunities and challenges of large language models for education", *Learning and Individual Differences*, Vol. 103, 102274.
- Lin, Y. and Yu, Z. (2024), "A bibliometric analysis of artificial intelligence chatbots in educational contexts", *Interactive Technology and Smart Education*, Vol. 21 No. 2, pp. 189-213.
- McKnight, D.H., Choudhury, V. and Kacmar, C. (2002), "Developing and validating trust measures for e-commerce: an integrative typology", *Information Systems Research*, Vol. 13 No. 3, pp. 334-359.
- Mhlanga, D. (2023), "Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning", *SSRN Electronic Journal*, [Preprint].

- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. and Podsakoff, N.P. (2003), "Common method biases in behavioral research: a critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 No. 5, pp. 879-903.
- Rehman, N. and Kang, M.A. (2024), "Exploring the impact of artificial intelligence on students' engagement and motivation in online learning environments", *International Journal of Trends and Innovations in Business & Social Sciences*, Vol. 2 No. 4.
- Ryan, R.M. and Deci, E.L. (2000), "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being", *American Psychologist*, Vol. 55 No. 1, pp. 68-78.
- Singh, P. (2024), "Artificial intelligence and student engagement: drivers and consequences", in Gierhart, A.R. (Ed.), *Advances in Educational Technologies and Instructional Design*, IGI Global, pp. 201-232.
- Strzelecki, A. (2024), "To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology", *Interactive Learning Environments*, Vol. 32 No. 9, pp. 5142-5155.
- Susnjak, T. and McIntosh, T. (2024), "ChatGPT: the end of online exam integrity?", *Education Sciences*, Vol. 14 No. 6, p. 656.
- Sykes, E.R. (2024), "AI-assisted academic writing: a comparative study of student-crafted and ChatGPT-enhanced critiques in ubiquitous computing", *International Journal of Advanced Computer Science and Applications*, Vol. 15 No. 9.
- Verplanken, B. and Orbell, S. (2003), "Reflections on past behavior: a self-report index of habit strength", *Journal of Applied Social Psychology*, Vol. 33 No. 6, pp. 1313-1330.
- Vieriu, A.M. and Petrea, G. (2025), "The impact of artificial intelligence (AI) on students' academic development", *Education Sciences*, Vol. 15 No. 3, p. 343.
- Wadhwa, R., Rabby, F., Bansal, R. and Hundekari, S. (2024), "Drivers and impact of artificial intelligence on student engagement", in Bansal, R., Chakir, A., Hafaz Ngah, A., Rabby, F. and Jain, A. (Eds.), *Advances in Computational Intelligence and Robotics*, IGI Global, pp. 161-170.
- Wanous, J.P., Reichers, A.E. and Hudy, M.J. (1997), "Overall job satisfaction: how good are single-item measures?", *Journal of Applied Psychology*, Vol. 82 No. 2, pp. 247-252.
- Zimmerman, B.J. (1990), "Self-regulated learning and academic achievement: an overview", *Educational Psychologist*, Vol. 25 No. 1, pp. 3-17.

## Economics Papers

### MARKET COMPETITION IN PUBLIC PROCUREMENT OF NORTH MACEDONIA (2021–2025): ECONOMIC ANALYSIS OF BIDDERS PER TENDER

**Viktor Mitevski**

*Association for Research and Analysis ZMAI,  
North Macedonia  
viktor@zmai.mk*

#### **ABSTRACT**

*This paper examines market competition in North Macedonian public procurement using a novel dataset of all public contracts from 2021–2025. We focus on the number of offers (bids) per tender as a key indicator of competition, following models inspired by Fazekas and Kocsis (2017). The analysis explores how institutional factors, particularly the use of electronic procurement tools and the choice of procedure type, influence bidder participation. We find that the average tender in North Macedonia attracts only 2–3 bids, and over one-third of procedures have a single bidder, raising market competition concerns. Using regression analysis with the number of offers as the dependent variable, we show that fully open procedures and e-procurement usage are associated with modestly higher competition, whereas negotiated or restricted procedures reduce the number of bidders. The paper provides descriptive insights (e.g., variation by contracting institution type and by goods/services/works procurement) and discusses implications for public expenditure effectiveness. Our results underscore the importance of transparent, open, and digitalized tendering processes in increasing competition and improving the efficiency of public spending.*

**Keywords:** *Public procurement, Competition, Number of bids; E-procurement; North Macedonia*

**JEL classification:** *H57, D44, D73*

#### **1. INTRODUCTION**

Public procurement plays a crucial role in public expenditure effectiveness, as competitive tendering can lead to cost savings and better value for money. One fundamental measure of competition in procurement is the number of bids received for a contract. A higher number of bidders tends to drive prices down and reduce the risk of favoritism or corruption, improving efficiency in public spending. Conversely, a lack of competition, often manifested as single-bidder tenders, can result in inflated costs and governance concerns. Recent European Union monitoring highlights this issue: in several countries, more than one-third of public procurement procedures above EU thresholds involve only a single bidder. Such findings have raised alarms about insufficient competition in procurement markets and its impact on public budgets.

North Macedonia, as an EU accession candidate, has undertaken procurement reforms in line with EU directives to enhance competition and transparency. Notably, the country implemented a centralized Electronic System for Public Procurement (ESPP), mandating e-procurement for most tenders in recent years. These efforts aim to lower barriers to entry for suppliers and increase the number of offers per tender, thereby improving outcomes. However, empirical evidence on the effectiveness of these measures remains limited. Yet, available

findings support the broader economic principle that competitive pressure leads to more favorable prices for the buyer (in this case, the government). Indeed, competition is one of the key metrics in the European Commission's Single Market Scoreboard (2022), which uses the Single Bid Indicator (the share of tenders with only one bid) to flag potential problems in national procurement systems. A high incidence of single-bidder contracts is seen as a red flag for restricted competition and possibly inefficiencies or corruption.

How competitive is the public procurement market in North Macedonia? Which factors significantly affect the number of bids in practice? This paper addresses these questions by analyzing a comprehensive dataset of public contracts in North Macedonia from 2021 to 2025 (latest available data), focusing on bidder participation as a proxy for market competition.

We build on the literature that views robust competition as essential for efficient public procurement and as a deterrent to corrupt or collusive practices. Following Fazekas and Kocsis's (2017) approach of using procurement data to identify "red flags" of restricted access, we use the number of offers received for each procedure as a quantitative indicator of competitiveness. By examining patterns in these data and performing regression analysis, we aim to identify which procedural and institutional factors are associated with higher or lower competition. In particular, we investigate the impact of electronic procurement tools (digital platforms for tender submission) and procedure types (open, negotiated, etc.) on the number of bidders. These factors have been highlighted in prior studies as potentially important levers for policy: for example, using open, advertised tenders and e-procurement is generally expected to increase bidder turnout.

The remainder of the paper is structured as follows. Section 2 provides a literature review of competition in public procurement. Section 3 states the hypothesis guiding this study. Section 4 outlines the methodology and data sources, explaining our scientific approach. In Section 5, we present an overview of the dataset and descriptive insights. Section 6 reports the regression results and interprets the effects of electronic tools and procedure type on the number of bids. Finally, Section 7 concludes with a discussion of findings and implications for improving public expenditure effectiveness through enhanced competition in procurement.

## **2. LITERATURE REVIEW**

This paper analyzes competition in public procurement by the number of valid bids received per tender, a measure widely used in empirical work to capture the intensity of rivalry for contracts. A consistent finding across settings is that more bidders are associated with better procurement outcomes, especially lower awarded prices and improved cost-effectiveness. In Turkish procurement auctions, Onur et al. (2012) show that each additional bidder is linked to a statistically significant reduction in prices, while evidence from Central Europe similarly associates higher bidder turnout with improved cost efficiency in the health sector (Nemec *et al.*, 2020). Using U.S. public procurement data, Bajari *et al.* (2009) show that auctions outperform bilateral negotiations for routine, standardized projects, where broader participation translates into stronger competitive pressure and lower costs relative to negotiated deals. In Italian public works, Decarolis (2014) documents that award and screening rules that increase effective participation materially affect both entry and performance, with mechanisms that foster genuine competition associated with better outcomes. Synthesizing theory and evidence, Spagnolo (2012) argues that institutions that promote entry and sustain reputation, such as open procedures and appropriate lot division, enhance competition and procurement performance.

Scholars have also connected the design of the tender as a major determinant of competition. Open procedures, which are advertised and allow any qualified firm to bid, generally attract more bidders than restrictive or non-competitive procedures. Ochrana and Pavel (2013), analyzing Czech Republic data, confirmed that the use of open tenders has a positive effect on the level of competition on the supply side. In contrast, limited or negotiated procedures, where fewer firms are invited or allowed to bid, tend to result in fewer offers. This is intuitive: greater openness broadens the pool of potential suppliers, increasing competition. From a policy standpoint, international organizations have long advocated open advertising of contracts to maximize competition (WTO Government Procurement Agreement; EU Procurement Directives 2014/24/EU). Exploiting Italian institutional features, Coviello and Gagliarducci (2017) show that greater buyer discretion, often exercised through negotiated or limited procedures, reduces competitive intensity, consistent with fewer bids. Related evidence from procurement auctions indicates that negotiations are associated with weaker competitive outcomes than auctions when comparable contracts are considered (Bajari *et al.*, 2009). Anderson *et al.* (2011) emphasize that publicly advertised tenders ensure transparency and equal opportunity, thereby encouraging more bidders and yielding better value. Empirical evidence supports this; for example. Nonetheless, there are cases where open procedures might not yield many bids due to market structure (e.g., very specialized contracts). A nuanced finding by Tátrai *et al.* (2024) is that negotiated procedures with communication (dialogue) can sometimes avoid single bids and secure 2–5 bids, although they rarely achieve a high number of offers. In general, however, limiting competition through procedure choice tends to lower bidder turnout.

A closely related channel is publicity and transparency in the call for competition. Leveraging a regression-discontinuity design in Italy, Coviello and Mariniello (2014) find that raising publicity requirements, mandating broader/online publication of tenders, increases the number of participating firms and intensifies price competition. The result is directly relevant to our empirical strategy: instruments that expand the pool of potential suppliers (public calls, clear documentation, broad dissemination) are expected to manifest as higher average bids per tender, *ceteris paribus*.

The literature also highlights the role of digital procurement tools in lowering participation costs and broadening access. In a two-country program evaluation, Lewis-Faupel *et al.* (2016) show that the introduction of e-procurement changed the composition and geography of winning firms and, importantly for our focus, facilitated wider participation by reducing information and submission frictions. Synthesizing global evidence, Becker *et al.* (2022) report that e-procurement reforms improve procurement performance in part by increasing supplier participation, consistent with a pro-competitive effect on bidder turnout. Complementary work in the public sector digitalization literature similarly links institutional capacity for digital tools to improved competitive conditions (Harrison and Sayogo, 2020). Taken together, these studies motivate testing whether tenders conducted with electronic tools exhibit higher numbers of offers.

Beyond procedure and digitalization, market and contract characteristics shape the intensity of competition. Using EU data, Tátrai *et al.* (2024) find systematic participation differences across procurement categories, with works tending to attract more bidders than goods and especially services, reflecting technology, asset specificity, and supplier-base depth. EU-wide evidence shows that more concentrated operator networks—captured by lower entropy—are associated with thinner effective competition in tenders (Fountoukidis, Wachs and Kertész, 2023).

Evidence on market structure is consistent with this pattern: in pharmaceuticals, where concentration can be high, Wouters *et al.* (2019) document competitive constraints that plausibly translate into fewer effective bidders. Methodologically oriented work on detecting collusion further emphasizes that thin participation is both a symptom and a facilitator of anticompetitive behavior, reinforcing the value of bidder-count indicators in diagnostic monitoring (Kumar *et al.*, 2015). In addition, the organizational capacity of contracting authorities matters: across 32 European countries, Cingolani and Fazekas (2017) show that stronger administrative capacity is associated with more competitive procurement processes, consistent with higher participation in open calls.

Finally, several studies speak directly to the auctions-versus-negotiations margin that underpins our empirical contrast. In U.S. public procurement, Bajari *et al.* (2009) provide evidence that auctions induce more competitive bidding and better screening of suppliers than bilateral bargaining for comparable projects. In Italian procurement, Decarolis (2014) documents how auction rules that increase effective participation (e.g., screening and price/quality trade-offs implemented transparently) are associated with improved outcomes. At a broader level, Spagnolo (2012) argues that institutional features that sustain entry and reputation, proxied empirically by more bidders, are central to disciplining prices and raising quality. These findings align with the core hypothesis of this paper that procedural openness and the use of electronic tools should be reflected in higher bidder turnout.

### 3. HYPOTHESES

Drawing from the above literature, we posit a central hypothesis for this study: Greater openness and digitalization in the procurement process leads to a higher number of bids, indicating increased competition. In particular, we expect that tenders which are conducted through fully open procedures and with electronic procurement tools will attract more offers, on average, than tenders using restrictive procedures or offline methods. To clarify, this overarching hypothesis can be articulated in two related parts:

**H1: The use of an electronic procurement system (e-procurement) increases the number of offers per procedure.** This is based on the notion that e-procurement lowers entry barriers and transaction costs for bidders, thereby expanding competition. We anticipate a positive association between the binary use of electronic tools and the count of bids received.

**H2: Open procedure types (e.g., open tenders) will yield more bids than less competitive procedure types (such as direct negotiations or limited calls).** Open advertisements generally increase the number of potential bidders, while negotiated or restricted tenders typically result in fewer participants and offers on average. We specifically expect Open procedures to outperform procedures like negotiated-without-call or direct contracting in terms of the number of bidders.

These hypotheses reflect the thesis theme that enhancing competition (through transparency and accessibility) improves public procurement outcomes. In testing H1 and H2, we implicitly examine whether the reforms in North Macedonia, aligning with EU best practices (like mandatory e-procurement and preference for open tenders), are achieving the desired effect of boosting competition. If the hypotheses are confirmed, this would indicate that enhancing these practices may lead to increased efficiency in public expenditure. Conversely, if we find only marginal or no differences, it may indicate that other bottlenecks (for example, supply market structure or administrative capacity) are limiting competition despite formal openness. The regression analysis in Section 6 will formally evaluate these hypotheses using the data.

#### 4. METHODOLOGY

In order to test the hypotheses about factors influencing competition in public procurement, we applied both descriptive statistics and econometric modeling techniques. The number of offers (bids) per tender is used as the key dependent variable to measure competition, consistent with literature identifying single-bid tenders as a warning sign of limited competition. We employed ordinary least squares (OLS) regression models with dummy-variable encoding for categorical factors of interest. This approach is mathematically equivalent to an ANOVA for comparing group means, but it provides regression coefficients that quantify the difference in the average number of bids between categories. By using regression, we can directly estimate how much each procurement method or tool impacts the expected number of bidders, while controlling for the baseline category. Both models, the one measuring Electronic Auction Effect and the Procedure Type Effect model, were estimated using a 95% confidence level, with significance evaluated based on p-values and standard errors.

Variables used include public procurement contracts in North Macedonia from 2021 to 2025, with each record containing information about the contracting authority, the procurement procedure type, whether an electronic auction was used, and the number of offers received, among other fields. The dependent variable  $Y_i$  is the Number of Offers for tender  $i$ , an integer count of valid bids submitted. The key independent variables are: (1) a binary indicator for the use of electronic bidding tools (e-auction), and (2) the category of procurement procedure used (a categorical factor with several levels as defined by public procurement law).

We estimate two OLS regression models to evaluate the impact of these factors on  $Y_i$ :

1. Model 1 (Electronic Auction Effect): This model tests whether the use of an electronic auction leads to a higher number of bids. It is specified as:

$$Y_i = \beta_0 + \beta_1 \cdot \text{ElectronicAuction}_i + \epsilon_i ,$$

where  $\text{ElectronicAuction}_i$  is a dummy variable indicating if an electronic reverse auction was used for tender  $i$  (1 = yes, 0 = no). Here,  $\beta_1$  captures the average difference in the number of offers between tenders with an e-auction and those without. We expect  $\beta_1 > 0$ , reflecting the hypothesis that e-auctions (which increase transparency and bidding convenience) encourage more bidders to participate.

2. Model 2 (Procedure Type Effect): This model assesses differences in competition across various procurement procedure types. The procedure type is a categorical variable with the following categories: Open Procedure, Simplified Open Procedure, Low-Value Procurement, Qualification System Notice, Special Services, Negotiated Procedure with Announcement, Negotiated Procedure without Announcement, and Direct Inter-institutional Agreement (Article 24). To include this factor in a regression, we created binary dummy variables for each category except one reference group. We chose “Negotiation without prior announcement” as the baseline (reference) category for comparison, because it is one of the least competitive procedure types (often involving a single invited bidder) and provides a meaningful contrast with more open procedures. The model is specified as:

$$Y_i = \beta_0 + \beta_1 D_{1,i} + \beta_2 D_{2,i} + \dots + \beta_7 D_{7,i} + \epsilon_i ,$$

where each  $D_{k,i}$  is a dummy variable for a specific procedure type category (taking value 1 if tender  $i$  is of that type, 0 otherwise). In our encoding:

–  $D_1 = 1$  if *Low-value procurement*,

- $D_2 = 1$  if *Qualification system notice*,
- $D_3 = 1$  if *Open procedure*,
- $D_4 = 1$  if *Simplified open procedure*,
- $D_5 = 1$  if *Special services procedure*,
- $D_6 = 1$  if *Negotiated procedure with prior announcement*,
- $D_7 = 1$  if *Article 24 (inter-institutional) procedure*,

and all  $D$ 's are 0 for the baseline *Negotiation without announcement*. The intercept  $\beta_0$  thus represents the mean number of offers in the baseline procedure (negotiation without call), and each coefficient  $\beta_k$  (for  $k = 1..7$ ) indicates how much the mean number of offers in that category differs from the baseline. For example,  $\beta_3$  for Open Procedure indicates the average increase in bidders when using an open tender compared to a direct negotiation. We anticipate positive coefficients for the more transparent, competitive procedures (open tenders, etc.), and possibly negative or small differences for certain special or closed procedures, reflecting our hypothesis that *open procedures attract more bidders than restricted negotiations*.

All independent variables are dummy-coded (0/1) and the models include an intercept. For the procedure-type factor we use mutually exclusive dummies with Negotiation without prior notice as the baseline (omitted) category to ensure full rank and avoid perfect multicollinearity. Models are estimated on the full 2021–2025 sample using OLS with HC1 robust standard errors. Because the outcome is a non-negative count but spans a broad range, we treat Number of Offers as approximately continuous to enable an ANOVA-equivalent linear comparison of group means (dummy OLS). Model adequacy is assessed by  $R^2$ /Adjusted  $R^2$  and the usual t-tests on coefficients; for completeness, we also report the overall F-statistic for each specification.

By employing this methodological approach, we directly test the effect of e-procurement tools and procedural choices on competition levels. The scientific approach here combines descriptive analysis (to understand basic indicators like average bids per category) with inferential statistics (to determine if observed differences are statistically significant and not due to random variation). This methodology enables us to quantitatively evaluate the hypotheses about competition: whether tenders using electronic auctions yield more bids, and whether open tendering procedures result in greater competition than negotiated or limited procedures. The results of these regression models are presented in Chapter 6, along with interpretation and comparison to expectations and prior research.

## 5. DATA OVERVIEW AND DESCRIPTIVE INSIGHTS

The dataset covers 101,389 public procurement contracts awarded in North Macedonia over 2021–2025 (the latest available data is from July 31). This represents virtually the entire public sector procurement activity in that period, across all government levels and sectors. The data is publicly available at the electronic system for public procurement (ESJN), and it is promptly updated, in real time, once a new contract is officially signed in the system. The average number of offers per tender in this dataset is 2.66. In other words, a typical tender attracted between two and three bids. However, the distribution is quite skewed. The median number of offers is 2, and about 35.7% of all procedures ended up with only a single bidder (one offer). At the upper extreme, a few tenders received dozens of bids; the maximum recorded was 110 offers for a single call (a rare outlier). Three-quarters of contracts had 4 or fewer bids. These figures raise some concern: more than one-third of single-bid procedures suggests that low

competition is common, echoing the challenges observed in other countries. While some single-bid cases may be justified (e.g., highly specialized contracts or urgent procurements), a high overall rate can indicate structural issues in the procurement market.

We next examine competition across different contracting institutions. For the purpose of this research we divide the Macedonian public sector as follows: central government ministries (including all central government bodies), local governments (municipalities and all their subunits), health institutions (e.g., hospitals), state-owned enterprises (including publicly traded companies founded by the state), educational institutions (schools, universities), Kindergartens, Courts and Public Prosecution and others (including smaller but specific or independent institutions). Each may face different market conditions and supplier pools. Table 1 summarizes the breakdown of contracts by institution type and the average number of offers.

*Table 1: Procurement contracts by institution type (2021–2025) and average number of offers*

<b>Type of Contracting Institution</b>	<b>Number of Contracts</b>	<b>Avg. Offers per Contract</b>
Health institutions (e.g., hospitals)	28,696	2.73
State-owned enterprises (utility companies)	20,102	2.32
Educational institutions (schools, etc.)	16,223	2.68
Local government (municipalities)	14,570	2.94
Central government (ministries, agencies)	14,080	2.62
Kindergartens (pre-school institutions)	4,235	2.69
Judicial/Prosecutorial bodies	1,417	3.24
Public funds (health, pension, etc.)	462	2.68
National Bank	612	2.39
Regional development centers	364	3.99
Parliament (Assembly)	264	2.14
Association of Local Gov. (ZELS)	37	2.19
<b>Total / Overall Average</b>	<b>101,389</b>	<b>2.66</b>

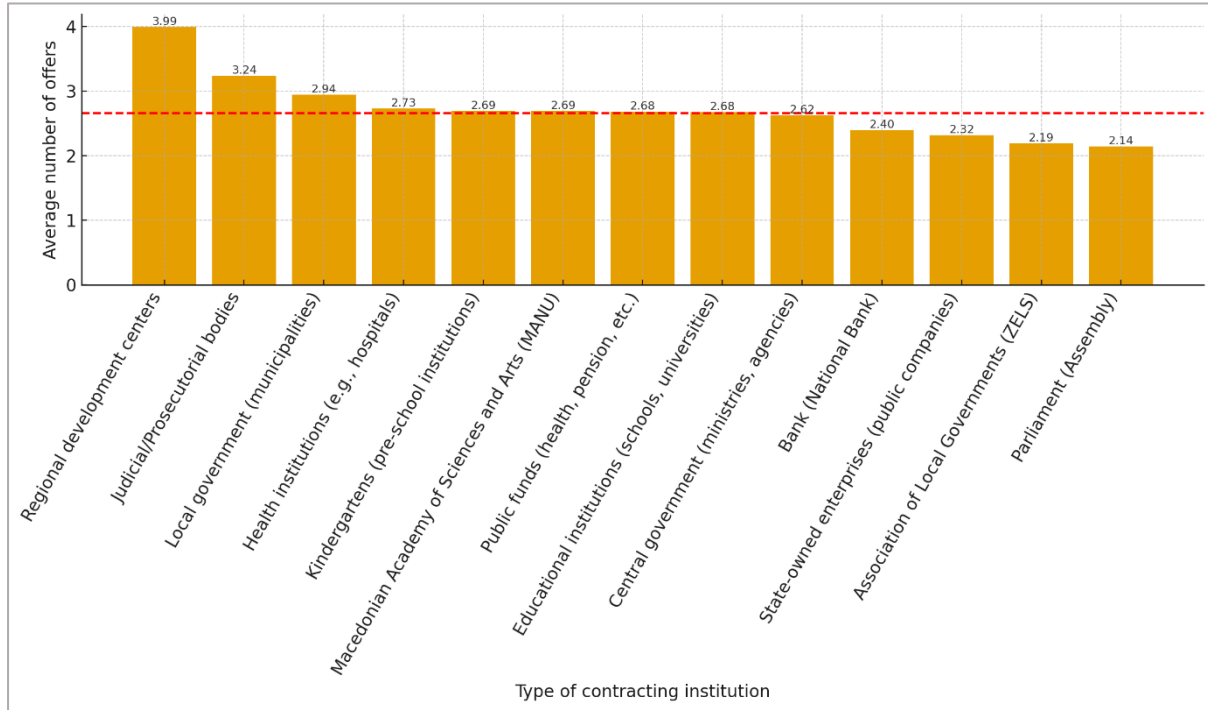
*(Source: Assigned contracts data from the electronic system of the Public Procurement Bureau)*

From the above table, we see that the majority of contracts, by volume, were conducted by five types of institutions: health institutions (which account for about 28% of all contracts), state-owned enterprises (~20%), educational institutions (~16%), local governments (~14%), and central government bodies (~14%). Together, these constitute roughly 92% of all procurements in the dataset. Smaller categories include kindergartens, courts/prosecutors, various public funds, the National Bank, etc.

There is notable variation in the average number of bids across these institution types. Local governments achieved a slightly higher-than-average competition with about 2.94 offers per tender. Health and education bodies are around the mean (2.7), while state-owned enterprises see somewhat fewer bids (~2.3 on average). The institutions with the highest average bids are the Regional Development Centers (nearly 4 bids on average) and Judicial/Prosecutorial institutions (over 3.2 bids). These two categories are relatively small in count, but the consistently higher competition might reflect the nature of their procurements. On the other end, the Parliament and the association of local governments (ZELS) had among the lowest averages (~2.1 bids). Central government ministries, interestingly, did not have as high competition as municipalities on average (2.62 vs 2.94), which might suggest that tenders by

ministries (often large or specialized projects) see moderate competition, whereas municipalities often procure more common works and goods (roads, utilities) which attract more bidders.

Figure 1: Average number of offers per procedure by type of contracting institution in North Macedonia (2021–2025)



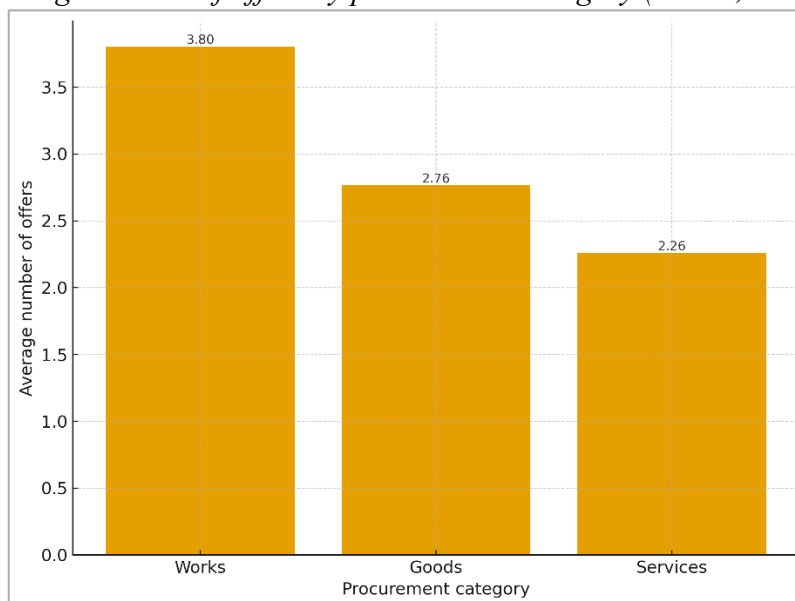
(Source: Own calculations based on assigned contracts data from the electronic system of the Public Procurement Bureau)

Another important dimension is the type of procurement by subject matter. The dataset classifies each contract broadly as involving Goods (supplies), Services, or Works (construction). These categories often have different market dynamics. Our data reveals a clear pattern:

- Works contracts tend to have the highest competition, with an average of 3.80 offers per tender.
- Goods contracts average about 2.76 offers.
- Services contracts have the lowest competition, averaging only 2.26 offers.

This ordering “works > goods > services” is consistent with findings in other countries. Works (such as construction projects) usually attract many construction firms, especially for roads, buildings, or maintenance works that numerous contractors can perform. Goods procurements, while also often competitive, can sometimes be limited if the goods are highly specialized or if there are exclusive dealerships (for example, medical equipment or brand-specific supplies). Services procurement often sees fewer bids; this may be due to specialization (e.g., consulting services might have only a few qualified providers) or the intangible nature of services making bidding more complex. Moreover, single-bid tenders are most prevalent in services. Indeed, in our dataset, nearly 47% of service procurements had just one bidder, compared to 31% for goods and only 20% for works. This aligns with the intuition that services are harder to standardize and attract wide competition, whereas work projects are widely contested.

Figure 2: Average number of offers by procurement category (Works, Goods, Services)



(Source: Own calculations based on assigned contracts data from the electronic system of the Public Procurement Bureau)

It is worth noting the extent of e-procurement usage in the dataset. This variable indicates whether the tender was carried out on the electronic procurement platform. Over the 2021–2025 period, the vast majority (98.3%) of contracts were conducted via e-procurement. This reflects the legal mandate in North Macedonia that nearly all public tenders use the online system. The small fraction (1.7%) that did not use e-tools corresponds to specific procedure types, primarily certain negotiated procedures, direct contracting between public entities, or a few special services tenders, which by law or circumstance were handled offline. When comparing outcomes, we find a stark difference: tenders using the electronic system received 2.67 offers on average, whereas those conducted without e-tools received only 1.85 offers on average. This raw difference (nearly one additional bid when using e-procurement) suggests that digital transparency and accessibility do correlate with more competition. However, as noted, this comparison overlaps with procedure type; many of the offline tenders were non-open procedures that inherently limit competition. The regression analysis in the next section will account for procedure types to isolate the effect (if any) of e-procurement itself.

## 6. EMPIRICAL RESULTS

The econometric analysis was carried out according to the approach outlined in the methodology. In this chapter, we present the findings from the two regression models and interpret the results in the context of public procurement competition. We also include visualizations to illustrate the differences in the number of offers across procedure types and between e-auction vs non-e-auction tenders. All reported coefficients were statistically tested; we highlight which effects are significant and discuss their practical implications.

### 6.1. Effect of Electronic Auctions on the Number of Bids

The first regression model examined whether the use of electronic procurement tools (electronic reverse auctions) is associated with a higher number of bidders. The OLS regression results confirm a clear positive effect. The coefficient of the electronic auction dummy is  $\beta_1 \approx +0.82$  ( $p < 0.001$ ), meaning that *tenders that employed an e-auction received, on average, about 0.82 more offers* than those without electronic bidding, holding other factors constant. In practical terms, this corresponds to an increase in competition: tenders without e-auctions had about 1.85 offers on average, whereas those with e-auctions had about 2.67 offers on average. This difference is statistically significant ( $F(1,10)=289.1$ ,  $p = 0.001$ ), suggesting that the use of the electronic bidding platform tends to attract more bidders.

Table 2: Regression model A outcomes

Variable	Coefficient	Std. Error	z	P> z	[0.025	0.975]	CI_low	CI_high
Const	1.85	0.07	24.84	<0.001	1.70	1.99	1.71	1.99
BinaryElectronicTools	0.82	0.07	10.97	<0.001	0.67	0.97	0.67	0.97

These results reinforce the finding that the use of electronic procurement has a positive effect on bidder turnout. The difference between the two groups is statistically significant, as indicated by the 95% confidence interval not overlapping between the bars. These findings are in line with expectations that an electronic auction can lower entry barriers and increase transparency, thereby motivating more suppliers to participate in the tendering process.

### 6.2. Differences by Procurement Procedure Type

The second regression model assessed how the choice of procurement procedure influences the level of competition (number of offers), using Negotiation without announcement as the reference category. The overall model was statistically significant ( $F(7,101381)=63.30$ ,  $p < 0.001$ ), indicating that the procedure type has a *significant overall effect* on the number of bids. The regression coefficients for each procedure dummy reveal the direction and magnitude of differences relative to the baseline.

Table 3: Regression model A outcomes

Variable	Coefficient	Std. Error	z	P> z	[0.025	0.975]	CI_low	CI_high
Const	1.61	0.13	12.02	<0.001	1.34	1.87	1.34	1.87
d1_small_value	1.02	0.13	7.63	<0.001	0.76	1.28	0.76	1.28
d2_establishment_notice	1.22	0.15	7.99	<0.001	0.92	1.52	0.92	1.51
d3_open	1.09	0.13	8.12	<0.001	0.83	1.35	0.83	1.35
d4_simplified_open	1.11	0.13	8.26	<0.001	0.84	1.37	0.84	1.37
d5_special_services	0.35	0.15	2.32	0.02	0.05	0.65	0.05	0.65
d6_negot_with_notice	0.11	0.22	0.52	0.60	-0.32	0.54	-0.32	0.54
d7_article_24	-0.61	0.13	-4.54	<0.001	-0.87	-0.34	-0.87	-0.34

Table 3 reports OLS coefficients (HC1) for procedure-type dummies with Negotiation without prior notice as the reference category. Each coefficient is the difference in the conditional mean number of offers relative to the baseline. The open and publicly advertised procedures show large, statistically significant positive effects: Open ( $\beta = 1.09$ , 95% CI [0.83, 1.35],  $p < 0.001$ ), Simplified open ( $\beta = 1.11$ , [0.84, 1.37],  $p < 0.001$ ), Low-value procurement ( $\beta = 1.02$ , [0.76, 1.28],  $p < 0.001$ ), and Qualification system notice ( $\beta = 1.22$ , [0.92, 1.52],  $p < 0.001$ ). Special

services have a smaller but significant effect ( $\beta = 0.35$ , [0.05, 0.65],  $p = 0.020$ ). Negotiation with prior notice is not statistically different from the baseline ( $\beta = 0.11$ , [-0.32, 0.55],  $p = 0.604$ )—note that this category has very few observations in the data, which limits precision. Article 24 (direct inter-institutional) is significantly lower than the baseline ( $\beta = -0.61$ , [-0.87, -0.34],  $p < 0.001$ ). Overall, the pattern in Table 3 shows that moving from a non-advertised negotiation to open/advertised procedures increases the expected number of offers by roughly one bidder, whereas Article 24 procedures yield even fewer offers than the already low baseline.

An ANOVA equivalent test confirms that at least some group means differ (the overall F-test is highly significant), and the regression’s individual coefficients pinpoint where those differences lie. These results are consistent with theoretical expectations and prior research: using less open and less transparent procedure types is generally associated with deliberately limited competition. Our empirical evidence from the Macedonian data strongly aligns with this notion, as the degree of openness of the procedure correlates with the number of competing bidders.

Although the model  $R^2$  is low (0.003–0.004), this is expected and acceptable in cross-sectional OLS used as a dummy-coded ANOVA to compare group means rather than to predict outcomes (Montgomery *et al.*, 2012; Kutner *et al.*, 2005). In such designs, most variation typically lies within groups, so the between-group share that  $R^2$  captures can be small even when mean differences are real and policy-relevant (Stock and Watson, 2015). For inference, what matters are the joint F-test for the procedure dummies and the individual coefficients with confidence intervals; a small  $R^2$  does not invalidate or bias the estimated average effects in cross-sections (Wooldridge, 2015; Angrist and Pischke, 2009). In our case, the joint test is highly significant ( $p < 0.001$ ), and the procedure-type coefficients are precisely estimated, so the conclusions about how openness relates to bidder turnout remain valid despite the modest  $R^2$ .

Table 4. Models fit statistics

Model	N	R-squared	Adj. R-squared	F-stat (overall)	Prob > F
(a) BinaryElectronicTools	101389	0.003	0.003	120.29	<0.001
(b) Procedure dummies	101389	0.004	0.005	10521.08	0

Overall, the empirical results robustly support the hypotheses: tenders conducted with electronic auctions attract significantly more bidders, and more open procurement procedures yield significantly more competition than restrictive procedures.

## 7. CONCLUSION AND DISCUSSION

This study set out to examine the competitive dynamics of public procurement in North Macedonia, focusing on how procedural factors influence the number of bidders. Our findings clearly demonstrate that greater transparency and openness in procurement lead to higher competition, whereas restrictive practices correspond with limited bidder turnout. Both of the core working hypotheses are confirmed by the analysis.

First, the use of electronic procurement tools (e-auctions) was found to significantly increase the number of offers. This suggests that implementing e-auctions is an effective way to boost competition in public tenders. By lowering participation costs and increasing transparency, e-auctions encourage more firms to bid. This result is consistent with practical expectations and sends a clear message to policymakers and contracting authorities: investing in and mandating

e-procurement tools can have a tangible positive impact on competition for government contracts. More bidders not only improve the chances of getting better prices and value for money but also reduce the likelihood of collusion or cozy arrangements. An increase from ~1.8 to ~2.7 bidders (when using e-auctions) might seem modest in absolute terms, but it can make the difference between a single-bid contract and a truly competitive tender.

The choice of procurement procedure greatly impacts competition. Open tenders and widely advertised processes attract significantly more bidders on average, open tenders receive about 2.7 bids compared to 1.6 for direct negotiations. Even simplified and low-value open procedures see high participation, highlighting the benefits of transparency across all contract sizes. In contrast, limited or unadvertised methods, like non-public negotiations or special agreements, usually lead to single or very few bids. The data from North Macedonia confirms that transparency and openness are crucial for competitive procurement, and simply announcing a negotiated tender does not boost participation unless the process itself is genuinely open. These findings suggest that policy reforms should prioritize open, competitive procedures over negotiated ones whenever feasible.

In conclusion, this research underscores the critical importance of competitive procurement processes. The empirical evidence supports strong recommendations for public procurement policy in North Macedonia (and similarly placed countries): namely, to maximize the use of open tender procedures and electronic auctions, and to minimize the use of non-transparent negotiated deals. If certain negotiated or direct-award procedures are legally allowed (for instance, in emergencies or for specialized cases), they should remain the exception rather than the norm. Each percentage point reduction in single-bid tenders is a step toward a more competitive and fair procurement system. The benefits are two-fold: economic efficiency (more bidders increase the chance of a better price) and integrity (reducing single-bid situations diminishes the opportunity for corrupt arrangements).

## REFERENCES

- Angrist, J. D. and Pischke, J.-S. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.
- Arrowsmith, S. and Anderson, R. D. (eds) (2011), *The WTO Regime on Government Procurement: Challenge and Reform*, Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511977015>
- Bajari, P., McMillan, R. and Tadelis, S. (2009), "Auctions versus negotiations in procurement: An empirical analysis", *Journal of Law, Economics, and Organization*, Vol. 25 No. 2, pp. 372-397. <https://doi.org/10.1093/jleo/ewn002>
- Becker, J., Lewis-Faupel, S. and Khan, N. (2022), "The effects of e-procurement on public spending efficiency: A global meta-analysis", *World Development*, Vol. 150, 105714.
- Cingolani, L. and Fazekas, M. (2017), "The administrative capacities behind competitive public procurement processes: A comparative assessment of 32 European countries", *Public Administration Review*, Vol. 77 No. 4, pp. 589-600. <https://doi.org/10.1111/puar.12784> [gcris.etu.edu.tr](https://www.gcris.etu.edu.tr)
- Coviello, D. and Gagliarducci, S. (2017), "Tenure in office and public procurement", *American Economic Journal: Economic Policy*, Vol. 9 No. 3, pp. 59-105.
- Coviello, D. and Mariniello, M. (2014), "Publicity requirements in public procurement: Evidence from a regression discontinuity design", *Journal of Public Economics*, Vol. 109, pp. 76-100. <https://doi.org/10.1016/j.jpubeco.2013.10.008> EUR-Lex

- Decarolis, F. (2014), “Awarding price, contract performance, and bids screening: Evidence from procurement auctions”, *American Economic Journal: Applied Economics*, Vol. 6 No. 1, pp. 108-132. <https://doi.org/10.1257/app.6.1.108> MDPI
- Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on public procurement (2014) *Official Journal of the European Union*, L 94, 28 March, pp. 65-242.
- European Commission (n.d.) ‘Single Market Scoreboard – Public procurement indicators’. Available at: [https://single-market-scoreboard.ec.europa.eu/business-framework-conditions/public-procurement\\_en](https://single-market-scoreboard.ec.europa.eu/business-framework-conditions/public-procurement_en) (Accessed: 1 September 2025). [single-market-scoreboard.ec.europa.eu](https://single-market-scoreboard.ec.europa.eu)
- European Court of Auditors (2023), *Special Report 28/2023: Public procurement in the EU — Competition remains weak in many tenders*, Luxembourg: Publications Office of the EU. European Court of Auditors
- Fazekas, M. and Kocsis, G. (2017), *Uncovering high-level corruption: Cross-national corruption proxies using government contracting data*. European Research Centre for Anti-Corruption and State-Building, Cambridge University Press.
- Fountoukidis, D., Wachs, J. and Kertész, J. (2023), “Competitive conditions and economic operator behavior in EU procurement markets: A network entropy-based approach”, *Journal of Public Procurement*, Vol. 23 No. 2, pp. 178-205.
- Grega, M. and Nemec, J. (2015), “Factors influencing final price of public procurement: evidence from Slovakia”, *Procedia Economics and Finance*, Vol. 25, pp. 543-551.
- Harrison, T. and Sayogo, D.S. (2020), “Digital procurement and efficiency gains: the role of institutional capacity”, *Government Information Quarterly*, Vol. 37 No. 2.
- Kumar, R., Mahajan, A. and Singh, P. (2015), “Bid-rigging detection in public procurement”, *European Journal of Operational Research*, Vol. 242 No. 2, pp. 591-603. <https://doi.org/10.1016/j.ejor.2014.09.073>
- Kutner, M. H., Nachtsheim, C., Neter, J. and Li, W. (2005), *Applied Linear Statistical Models* (5th edition), New York, NY: McGraw-Hill/Irwin.
- Lewis-Faupel, S., Neggers, Y., Olken, B. A. and Pande, R. (2016), “Can electronic procurement improve infrastructure provision? Evidence from public works in India and Indonesia”, *American Economic Journal: Economic Policy*, Vol. 8 No. 3, pp. 258-283. <https://doi.org/10.1257/pol.20140258> IDEAS/RePEc
- Ochrana, F. and Pavel, J. (2013), “Analysis of the impact of transparency, corruption, openness in competition and tender procedures on public procurement in the Czech Republic”, *Central European Journal of Public Policy*, Vol. 7 No. 2, pp. 114-135.
- Onur, İ., Özcan, R. and Taş, B. K. O. (2012), “Public procurement auctions and competition in Turkey”, *Review of Industrial Organization*, Vol. 40 No. 3, pp. 207-223. <https://doi.org/10.1007/s11151-011-9299-3> IDEAS/RePEc
- Spagnolo, G. (2012), “Reputation, competition, and entry in procurement”, *International Journal of Industrial Organization*, Vol. 30 No. 3, pp. 291-306. <https://doi.org/10.1016/j.ijindorg.2012.02.001> Google Scholar
- Stock, J.H. and Watson, M.W. (2015), *Introduction to Econometrics* (3rd/updated ed.).
- Tátrai, T., Vörösmarty, G. and Juhász, P. (2024), “Intensifying competition in public procurement”, *Public Organization Review*, Vol. 24, pp. 223-247. <https://doi.org/10.1007/s11115-023-00742-0> SpringerLink
- Wooldridge, J. M. (2015), *Introductory Econometrics: A Modern Approach*, (6th edition), Mason, OH: South-Western Cengage Learning.
- Wouters, O. J., McKee, M. and Luyten, J. (2019), “Public procurement of medicines in South Africa: A market concentration analysis”, *Health Policy*, Vol. 123 No. 8, pp. 714-723.
- WTO, (2012), Agreement on Government Procurement and related WTO legal texts.

## THE MACROECONOMIC EFFECTS OF FISCAL CONSOLIDATION AFTER CRISES: EVIDENCE FROM A SMALL OPEN ECONOMY

**Faton Sulejmani**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
fatonsulejmani@gmail.com*

**Dragan Tevdovski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
dragan.tevdovski@eccf.ukim.edu.mk*

### **ABSTRACT**

*This paper investigates the macroeconomic effects of fiscal policy in North Macedonia, with a focus on fiscal consolidation episodes after major crisis periods, including the Global Financial Crisis (2008–2009), the COVID-19 pandemic (2020), and the recent inflationary and energy shocks (2022–2023). We estimate a structural vector autoregression (SVAR) model with recursive identification to analyze the dynamic responses of key macroeconomic variables to discretionary fiscal tightening, captured by positive shocks to the cyclically adjusted primary balance (CAPB). The model includes domestic variables—output gap, interbank interest rate (SKIBOR), and inflation—alongside exogenous indicators of Eurozone economic conditions. The empirical results show that fiscal consolidations have contractionary effects in the short run. Output declines significantly following a tightening shock, with the output gap reaching its trough within the first two quarters. Inflation decreases moderately and persistently, while the interbank interest rate responds with a temporary decline, indicating a counter-cyclical monetary policy reaction. These findings are consistent with the Keynesian view of fiscal multipliers in small open economies and highlight the importance of fiscal-monetary coordination.*

**Keywords:** *Fiscal consolidation, Vector auto regression, Inflation, Output gap*

**JEL Classification:** *C32, E12, E62, E63, E65*

### **1. INTRODUCTION**

Small and open economies face heightened vulnerabilities when exposed to global shocks, owing to their strong trade linkages, limited domestic demand, and often constrained monetary and fiscal policy space. Over the past two decades, these economies have been repeatedly challenged by major global disruptions, including the Global Financial Crisis (2008–2009), the COVID-19 pandemic (2020), and the energy and price shocks stemming from the war in Ukraine. In response to these crises, governments around the world, including in small and open economies, enacted large-scale fiscal expansions to stabilize output, protect households and businesses, and support essential services. However, as crisis conditions receded, the need to restore fiscal sustainability prompted a return to fiscal consolidation, raising important questions about the short-term macroeconomic consequences of such policy shifts.

The effectiveness and cost of fiscal consolidation in the post-crisis environment are a subject of particular concern in small and open economies. Unlike advanced economies, which

typically benefit from deep capital markets and greater institutional capacity, small and open economies often operate under tighter budget constraints and more rigid monetary frameworks—such as exchange rate pegs or limited inflation-targeting regimes. In these settings, fiscal policy plays an outsized role in macroeconomic management, but the transition from stimulus to consolidation can risk undermining fragile recoveries. This is especially relevant in countries like North Macedonia, where the de facto exchange rate peg to the euro limits the flexibility of monetary policy, and where fiscal consolidation efforts following crisis-driven spending have been a central policy challenge.

While the empirical literature on fiscal policy is vast, much of it focuses on the effects of fiscal stimulus during recessions or on the average fiscal multiplier across the business cycle. Research on fiscal consolidations is more limited, and even less attention has been paid to their effects specifically in small and open economies following large external shocks. Studies such as Auerbach and Gorodnichenko (2012) and Ramey and Zubairy (2018) highlight that the timing and state of the economy significantly influence fiscal multipliers. Ilzetzki *et al.* (2013) suggest that multipliers are smaller in open economies, but they may be larger during synchronized global downturns. More recent work by Deb *et al.* (2022) and Furceri *et al.* (2021) examines the effectiveness of fiscal responses during the COVID-19 pandemic, yet the consequences of fiscal tightening in the recovery phase remain underexplored.

This paper contributes to the literature by assessing the macroeconomic effects of fiscal consolidation in North Macedonia, a small and open economy in Southeastern Europe, in the aftermath of multiple crisis episodes. Using quarterly data from 2007 to 2024, we estimate a structural vector autoregression (SVAR) model to examine the short-term dynamic effects of fiscal tightening, measured by a positive shock to the cyclically adjusted primary balance (CAPB). The model includes four domestic variables—output gap, CAPB, interbank interest rate (SKIBOR), and inflation—alongside exogenous Eurozone variables that reflect the broader external environment. The identification strategy follows a recursive (Cholesky) ordering, allowing for a clear interpretation of contemporaneous interactions.

Our findings indicate that fiscal consolidations in the post-crisis period lead to significant short-term declines in output and inflation, while the interbank interest rate falls in response, suggesting a counter-cyclical monetary adjustment. These results highlight the short-run costs associated with restoring fiscal discipline after crisis episodes and emphasize the importance of timing, coordination, and external conditions in shaping the effectiveness of consolidation policies in small and open economies.

The remainder of the paper is organized as follows. Section 2 reviews the relevant empirical literature. Section 3 describes the data and key macroeconomic variables. Section 4 outlines the SVAR model and identification strategy. Section 5 presents the empirical results, and Section 6 concludes with policy implications and suggestions for future research.

## **2. LITERATURE REVIEW**

This literature review examines the state of research on fiscal policy in the face of external shocks, with a focus on methodological advancements, the heterogeneity of fiscal multipliers, external constraints in small and open economies, and the empirical evidence emerging from recent crises.

External shocks such as global pandemics, commodity price surges, and geopolitical conflicts differ from conventional business cycle fluctuations in their simultaneity, persistence, and global scope. Unlike domestic recessions, these shocks simultaneously affect demand, supply, and external balances, complicating the design of fiscal interventions. The literature has increasingly emphasized the need for state-contingent fiscal policy, wherein the effectiveness of government spending and tax measures depends on the macroeconomic environment (Auerbach and Gorodnichenko, 2012; Ramey and Zubairy, 2018).

Fiscal consolidation refers to deliberate policy measures aimed at reducing government deficits and debt accumulation to restore fiscal sustainability and maintain macroeconomic stability. It can be achieved through expenditure-based consolidation, which relies on spending cuts, or revenue-based consolidation, which depends on increasing taxes or broadening the tax base. Empirical studies generally find that expenditure-based consolidations are more durable and less contractionary than revenue-based ones (Alesina and Ardagna, 2010; Alesina, *et al.*, 2019), though their short-term effects depend on economic conditions and the composition of adjustment. For instance, IMF (2010) and OECD (2012) analyses highlight that consolidations implemented during economic expansions and those focusing on reducing current expenditures—rather than public investment—tend to have smaller negative impacts on growth. Conversely, when fiscal adjustments rely heavily on tax hikes or occur during recessions, they may exacerbate output contractions (Guajardo *et al.*, 2014). Overall, the literature suggests that the design, timing, and credibility of fiscal consolidation strategies are crucial in determining their macroeconomic consequences and long-term success.

Large body of literature aims to assess the fiscal policy effects in small and open economies (for example, Mirdala, 2009; Caraianni, 2010; Franta, 2012; Jemec *et al.*, 2011; Karagyozyova-Markova *et al.*, 2013; Bobasu, 2015; Boiciuc, 2015; Asandului *et al.*, 2021) while there is smaller strand that has focused on Western Balkans region and North Macedonia specifically (see for example Hinic and Miletic, 2013; Petrevski *et al.*, 2016; Krajišnik *et al.*, 2019 and Tevdovski *et al.*, 2019). These economies face structural features that tend to dampen the effects of fiscal expansion, including high import propensities, shallow capital markets, and external vulnerability. Ilzetzki *et al.* (2013) show that the degree of openness and exchange rate flexibility are key determinants of fiscal multiplier size, with small and open economies often experiencing leakages through trade balances. However, more recent work suggests that external synchronization of shocks and policy responses—as seen during the COVID-19 crisis—can offset these leakages, leading to larger effective multipliers (Chudik *et al.*, 2021).

Crisis conditions - especially those involving simultaneous shocks - have motivated methodological innovations in identifying and estimating the effects of fiscal policy. Traditional structural VARs (Blanchard and Perotti, 2002) have been expanded through narrative identification (Romer and Romer, 2010) and instrumental variable approaches using external instruments (Stock and Watson, 2012). These techniques help isolate exogenous fiscal policy changes from endogenous responses, particularly during emergencies when policy actions are large and swift.

During the COVID-19 pandemic, researchers employed cross-country panel VARs (Deb *et al.*, 2022) and machine learning tools (Forni *et al.*, 2021) to estimate the effects of health-related spending, income support, and public investment. The war in Ukraine and its economic consequences prompted an evolution in policy focus from direct stimulus to targeted relief (e.g., energy subsidies) and strategic investments (e.g., defense, energy independence). The literature emphasizes the risks of untargeted subsidies, which can be regressive and

inflationary, and recommends well-designed automatic stabilizers and social protection systems (OECD, 2023).

### 3. DATA

This paper analyzes the macroeconomic effects of fiscal policy in North Macedonia. We use quarterly annualized data spanning from the first quarter of 2007 to the fourth quarter of 2024, yielding a total of 72 observations. Following the approach of Petrevski *et al.* (2016), the model includes two sets of variables: exogenous and endogenous.

The exogenous variables capture developments in the Eurozone, reflecting the high degree of economic integration between North Macedonia and the EU. The Macedonian denar is de facto pegged to the euro, and the EU is by far North Macedonia's largest trading and investment partner. Approximately 80% of Macedonian exports are directed to the EU, and nearly 80% of the capital in the domestic banking sector originates from EU sources. In this context, the exogenous set includes the Eurozone output gap and the quarterly annualized inflation rate of the Euro area.

The endogenous variables pertain to domestic macroeconomic conditions and include the output gap, the interbank interest rate (SKIBOR), the cyclically adjusted primary balance, and the quarterly annualized inflation rate. All data were obtained from the national authorities' sources, i.e., the Ministry of Finance, the National Bank, and the State Statistical Office.

The output gap is calculated as the difference between actual and potential real GDP, with the latter estimated using the Hodrick-Prescott filter and a smoothing parameter ( $\lambda$ ) of 1600, which remains the most frequently used method, although it is important to acknowledge inherent limitations related to it, such as edge effects and end-point bias. SKIBOR serves as a proxy for monetary policy developments, as it is influenced by the central bank's policy rate through the monetary transmission mechanism. The cyclically adjusted primary balance is used as an indicator of the fiscal policy stance. The table below provides descriptive statistics for the selected variables, followed by visual representations and a discussion of their dynamics.

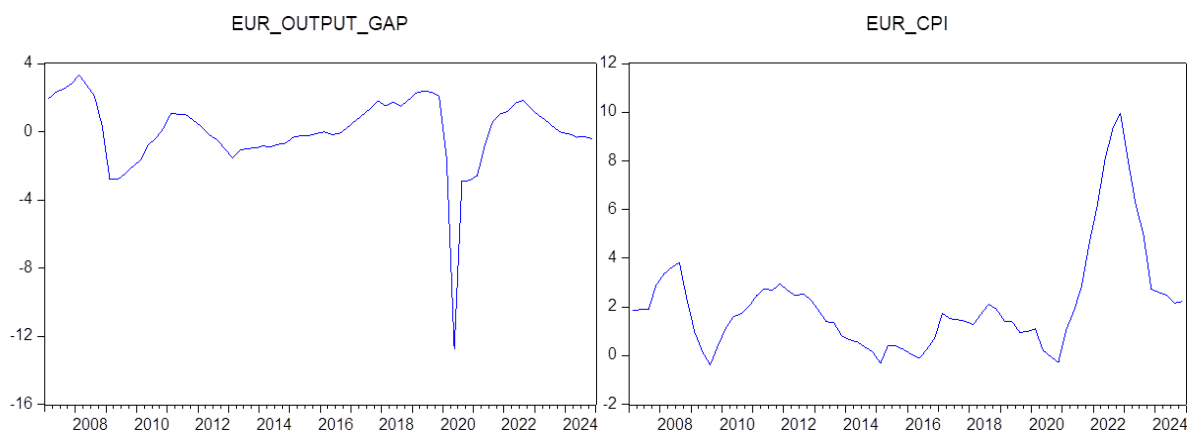
Table 1: Descriptive statistics of the selected variables

	CAPB	OUTPUTGAP_MK	INFLQ_MK	SKIBOR	OUTPUTGAP_EMU	INFLQ_EMU	EURIBOR
Mean	-2.224232	0.078215	3.1018	3.631429	0.062976	2.126962	1.023575
Median	-2.016893	0.21534	1.995577	3.245	0.095911	1.730464	0.2322
Maximum	5.314951	8.778409	19.35582	9.523333	3.331535	9.955108	4.9818
Minimum	-11.88386	-12.55775	-2.052688	1.316667	-12.76361	-0.377058	-0.5664
Std. Dev.	2.475184	2.493732	4.199629	2.305428	2.152172	2.132859	1.725291
Skewness	-0.435303	-1.165463	2.018262	0.958283	-2.947265	1.886333	1.067764
Kurtosis	6.8233	12.42915	7.248788	3.123667	18.61997	6.762257	2.673977

(Source: National authorities, Eurostat, Authors' calculations)

Figure 1 presents the dynamics of external macroeconomic conditions—particularly in the Eurozone, which are critical for understanding developments in North Macedonia's small open economy. The left panel presents the Eurozone output gap, which reflects cyclical fluctuations in the EU economy. The period from 2007 to 2024 captures several major global shocks and their transmission to the broader European economic environment.

Figure 1: Time series of the exogenous variables



(Source: Eurostat, Authors' calculations)

During the Global Financial Crisis (2008–2009), the Eurozone output gap turned sharply negative, reflecting a deep recession across member states. A slow recovery followed, although the aftermath of the crisis and the euro area sovereign debt crisis (2011–2013) kept the output gap in negative territory through much of the early 2010s. Gradual improvement is visible between 2014 and 2019, with the gap approaching neutral levels as economic conditions stabilized across the Eurozone. However, in 2020, the COVID-19 pandemic caused an unprecedented collapse in output, as evidenced by the abrupt and severe drop in the output gap—one of the sharpest in the sample period. The subsequent recovery in 2021–2022 was relatively strong but short-lived, with the gap narrowing and returning to slightly positive values. From 2023 onward, the Eurozone economy showed signs of deceleration amid tightening monetary conditions, declining external demand, and geopolitical uncertainty, leading to a modestly negative output gap.

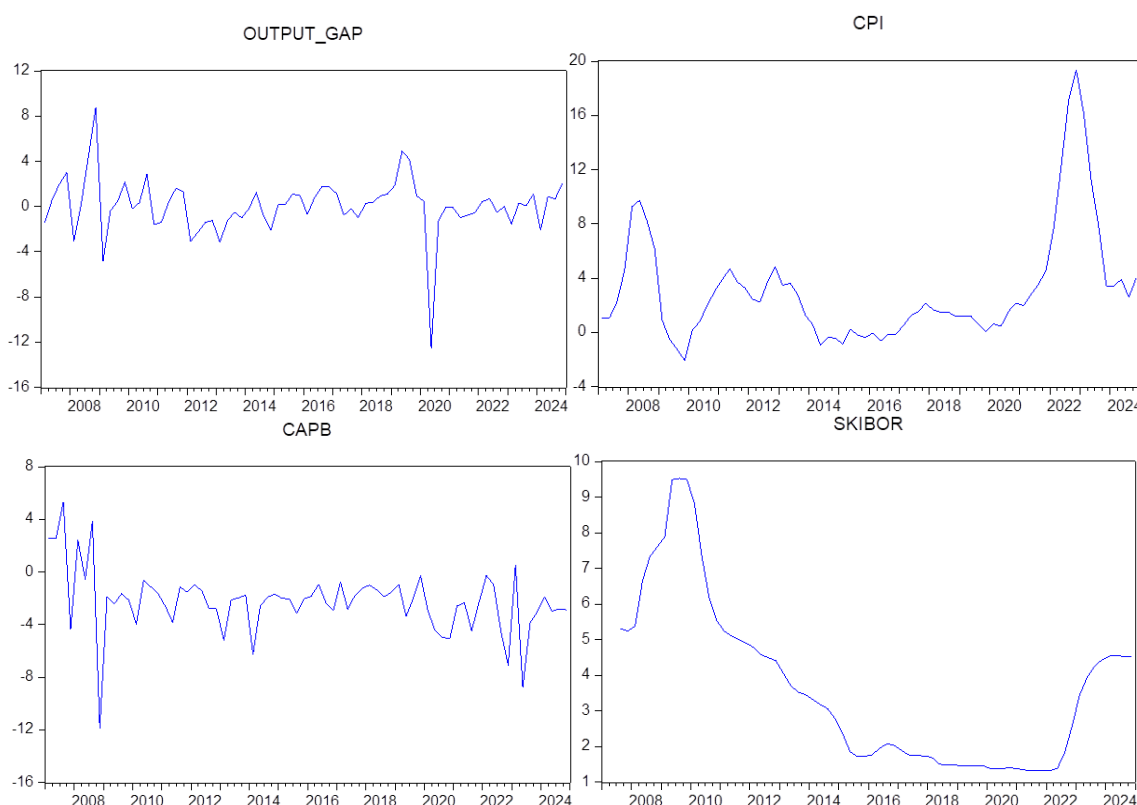
The right panel illustrates the quarterly annualized inflation rate (CPI) in the Euro area. Inflation remained relatively low and stable for most of the period from 2007 to 2020. Brief fluctuations occurred during the post-crisis recovery years, but inflation generally remained below the European Central Bank's target, prompting a prolonged period of accommodative monetary policy. Beginning in 2021, inflation surged dramatically, driven by a combination of global supply bottlenecks, strong post-pandemic demand, and a sharp increase in energy and food prices following the outbreak of the war in Ukraine in early 2022. The inflation rate peaked at historically high levels in 2022, before starting to moderate in 2023. By 2024, inflation had declined but remained above pre-pandemic levels, suggesting persistent price pressures despite monetary tightening.

These two Eurozone variables—output gap and inflation—are included in our empirical model as exogenous drivers, capturing the broader external environment. They serve to control for spillover effects from the EU economy on domestic output, inflation, and policy responses in North Macedonia.

Figure 2 presents the evolution of the domestic (endogenous) variables in the Macedonian economy over the observed period. As shown in the top-left panel, the output gap in North Macedonia displays a dynamic pattern over the period 2007–2024, reflecting key macroeconomic shocks and phases of the business cycle. The period of the Global Financial Crisis (2008–2009) is marked by a sharp swing into negative territory, indicative of a significant contraction in economic activity. This was followed by a gradual recovery, although

the output gap remained volatile, particularly during the euro area sovereign debt crisis (2011–2013), which adversely affected external demand. Between 2015 and 2019, the output gap fluctuated moderately around zero, consistent with a relatively stable macroeconomic environment. A sharp contraction reappears in 2020 due to the COVID-19 pandemic, followed by a strong positive rebound in 2021 and modestly positive gaps in subsequent years, pointing to a recovery phase and increasing utilization of economic capacity.

Figure 2: Time series of the endogenous variables



(Source: National authorities, Authors' calculations)

Inflation developments are depicted in the top-right panel. The quarterly annualized CPI rate shows moderate variation during the early part of the sample, with slightly elevated rates following the Global Financial Crisis. A period of price stability characterizes the years 2015–2019, during which inflation remained subdued. However, inflation spiked markedly in 2022, primarily driven by global supply disruptions, the energy crisis, and war-related commodity price shocks. This upward surge proved temporary, as inflation rates began to normalize in 2023 and returned closer to historical trends by 2024.

The bottom-right panel illustrates the behavior of the cyclically adjusted primary balance (CAPB), an indicator of the fiscal policy stance. The CAPB turns strongly negative during the Global Financial Crisis, consistent with countercyclical fiscal policy. Throughout the 2010s, the fiscal position remained mildly accommodative, with short-lived episodes of consolidation. A significant deterioration occurred in 2020 in response to the pandemic, as the government deployed substantial fiscal support. From 2021 onward, the CAPB gradually recovered, suggesting a return to a more prudent fiscal path, though periodic negative values highlight continued fiscal support during the overlapping energy and inflation crises of 2022.

Finally, the bottom-left panel displays the evolution of the interbank interest rate (SKIBOR), serving as a proxy for the stance of monetary policy. The rate remained elevated during the pre-crisis period and declined gradually after 2010, in line with accommodative monetary policy. From 2015 to 2020, SKIBOR hovered near historic lows, reflecting subdued inflation and a supportive policy environment. The rate began to increase in 2022 in response to heightened inflation, indicating a gradual tightening of monetary conditions by the central bank.

Together, these four series underscore the complex interplay between economic activity, inflationary pressures, and policy responses. The observed fluctuations provide a compelling empirical context for assessing the effectiveness of fiscal policy in a small open economy facing a sequence of global shocks.

Before proceeding with the modeling, all series were tested for stationarity using the Augmented Dickey-Fuller test. The results indicated that all series were stationary except for the money market rate (SKIBOR), to which a first-difference transformation was performed.

#### **4. METHODOLOGY**

To examine the macroeconomic effects of fiscal policy, we estimate vector autoregression (VAR) models using a recursive (Cholesky) identification scheme. VAR models are well-suited for analyzing dynamic interactions among macroeconomic variables without imposing strong theoretical priors, making them particularly useful in empirical macroeconomics (Sims, 1980). In this framework, each variable is treated as endogenous and its evolution is explained by its own lags as well as the lags of all other variables in the system.

We employ a recursive structural VAR (SVAR), in which identification is achieved through a Cholesky decomposition of the reduced-form residuals. This imposes a contemporaneous triangular causal structure: each variable is allowed to contemporaneously affect only those ordered after it, while being contemporaneously unaffected by those that follow (Blanchard and Quah, 1989). This identification strategy is widely used in empirical fiscal policy literature (Perotti, 2007; Auerbach and Gorodnichenko, 2012) due to its transparency and minimal assumptions.

The ordering of the endogenous domestic variables in our baseline specification is as follows: output gap, cyclically adjusted primary balance (CAPB), interbank interest rate (SKIBOR), and inflation. This ordering implies that the output gap contemporaneously influences fiscal policy (CAPB), monetary conditions (SKIBOR), and inflation, but is not itself affected by them within the same period. Fiscal policy is assumed to react contemporaneously to output conditions, but not to changes in monetary policy or inflation. SKIBOR, as a proxy for monetary policy, is assumed to respond to both output and fiscal policy within the quarter, but not to inflation. Finally, inflation is placed last, responding contemporaneously to all other variables but exerting no immediate effect.

Variables capturing Eurozone conditions—namely the Eurozone output gap, money market interest rate, and inflation—are included as exogenous variables in the model. Given the high degree of economic integration between North Macedonia and the Eurozone, it is assumed that these external factors contemporaneously influence all domestic variables, but are not affected

by domestic developments. This assumption reflects the small open economy nature of North Macedonia.

Our empirical strategy proceeds in two stages. In the first stage, we estimate an unrestricted VAR to determine the appropriate lag length. Given the relatively short sample size in relation to the number of estimated parameters, model parsimony is essential. Lag length is selected based on the Schwartz information criterion (BIC), which indicates a two-lag structure.

In the second stage, we evaluate the adequacy and stability of the VAR model through standard residual-based diagnostic tests. These include tests for serial correlation (Portmanteau and Lagrange Multiplier tests), normality (Jarque–Bera test), and conditional heteroskedasticity (ARCH–LM test). Only when the specification passes these diagnostic checks do we proceed to structural identification and impulse response analysis. This methodological framework enables us to recover structural fiscal shocks and trace their dynamic impact on key macroeconomic variables such as output, inflation, and interest rates. By controlling for both domestic and external conditions, the model captures the transmission mechanisms and timing of fiscal policy effects in a small open economy context.

## **5. EMPIRICAL RESULTS**

The impulse response functions (IRFs) trace the dynamic effects of a positive one-standard-deviation structural shock to the cyclically adjusted primary balance (CAPB)—interpreted as a discretionary fiscal consolidation—on key macroeconomic variables in North Macedonia (Figure 3). The IRFs are derived from the recursively identified VAR model described in the methodology and are plotted over a 10-quarter horizon with 68% confidence bands.

As shown in the top-left panel, the output gap responds negatively to a fiscal consolidation shock. The effect is immediate and statistically significant in the first two quarters, peaking at approximately  $-0.4$  percentage points in quarter 2, before gradually returning to zero. This pattern is consistent with standard Keynesian effects, whereby fiscal tightening—through higher taxes or lower public spending—dampens aggregate demand and reduces real economic activity in the short run.

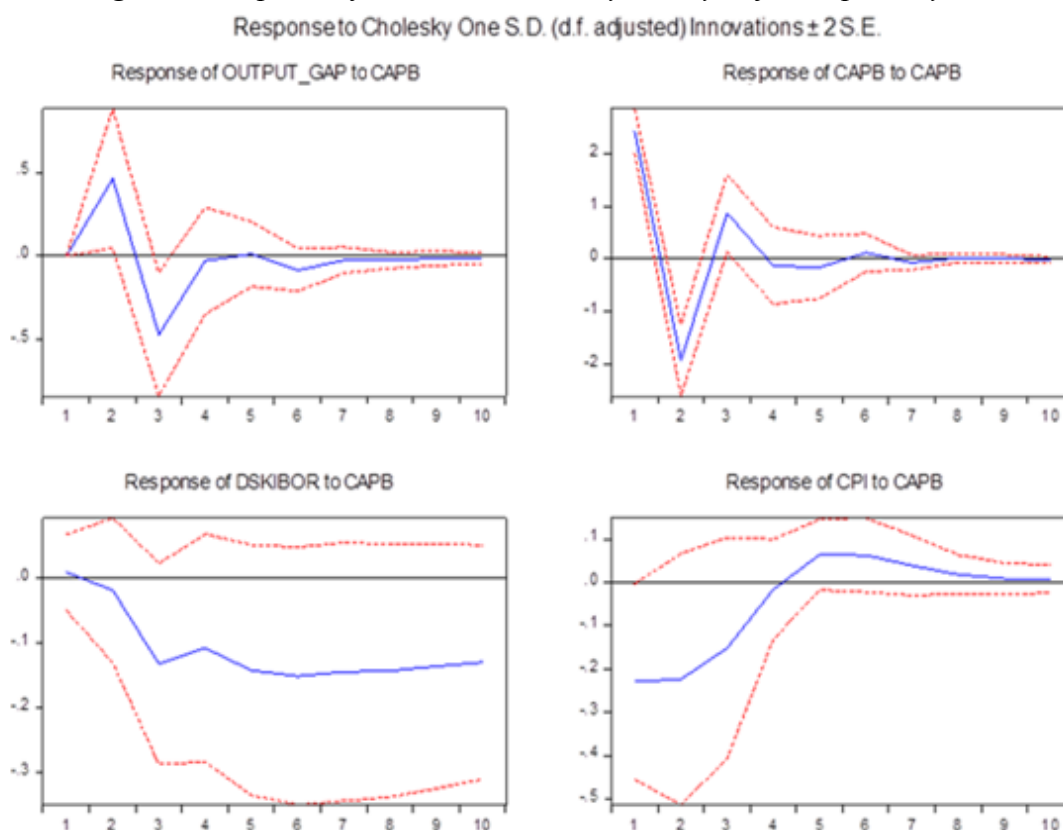
The top-right panel shows the response of the cyclically adjusted primary balance, confirming the identification of the fiscal shock. The CAPB rises sharply and immediately, by more than 2 percentage points on impact, and gradually dissipates over the following quarters. This strong and transitory reaction supports the interpretation of the shock as an exogenous fiscal tightening episode.

In the bottom-left panel, the interbank interest rate (SKIBOR) responds negatively to the fiscal shock, declining by nearly 1 percentage point in the short term. This suggests that monetary policy may react counter-cyclically to fiscal tightening, likely reflecting attempts by the central bank to mitigate contractionary effects through lower interest rates. This reaction is economically plausible in the context of a small open economy with constrained fiscal space and an active role for monetary stabilization.

The bottom-right panel presents the response of inflation. Following the fiscal consolidation, inflation decreases gradually, reaching a trough after 4–6 quarters. The maximum decline is approximately 0.5 percentage points, after which the effect begins to fade. The disinflationary

response is consistent with the decline in aggregate demand triggered by fiscal tightening and further reinforced by reduced cost-push pressures in a slackening economy.

Figure 3: Responses from shocks in the cyclically adjusted primary balance



(Source: Authors' calculations)

Taken together, the impulse responses provide evidence that fiscal consolidations in North Macedonia are contractionary in the short run, with adverse effects on output and inflation. Monetary policy appears to respond in a counter-cyclical manner, providing partial stabilization through lower interest rates. These findings are broadly consistent with the empirical literature on fiscal multipliers in small open economies (see Alesina *et al.*, 2015; Ilzetzki *et al.*, 2013), where fiscal policy tends to have significant real effects, particularly in economies with limited monetary-fiscal coordination or institutional buffers.

## 6. DISCUSSION AND CONCLUSION

This paper investigates the macroeconomic effects of fiscal policy in North Macedonia using a structural vector autoregression (SVAR) framework with a recursive identification scheme. The model incorporates both domestic and external macroeconomic variables, capturing the high degree of economic integration between North Macedonia and the Eurozone. Our analysis covers the period from 2007 to 2024 and includes key episodes such as the Global Financial Crisis, the COVID-19 pandemic, and the recent inflationary and energy price shocks.

The impulse response analysis provides robust evidence that fiscal consolidations—captured through positive shocks to the cyclically adjusted primary balance—have contractionary effects on output in the short run. Real activity, measured by the output gap, declines

significantly in the first few quarters following a fiscal tightening. Inflation also decreases, albeit with a lag, indicating disinflationary pressures due to weakened aggregate demand. Interestingly, monetary conditions, proxied by the interbank interest rate (SKIBOR), appear to ease in response to fiscal shocks, suggesting that the central bank may act to offset the adverse effects of fiscal consolidation. The findings regarding the effects of fiscal consolidation on real economic activity appear to differ from those reported by Petrevski *et al.* (2016) and Tevdovski *et al.* (2019) for North Macedonia. However, they appear broadly consistent with these studies concerning the impact on inflation and the response of monetary policy. Overall, as previously noted, the results align broadly with the empirical literature on fiscal multipliers in small open economies.

These findings carry important policy implications. First, they highlight the macroeconomic costs of short-term fiscal consolidation in a small open economy, particularly in times of economic slack. Policymakers should be cautious about implementing pro-cyclical fiscal tightening during downturns, as it may amplify economic volatility. Second, the observed monetary policy response underscores the importance of fiscal-monetary coordination in achieving macroeconomic stability. In an economy with a fixed exchange rate regime and limited independent monetary space, such coordination becomes even more critical.

Despite its contributions, the analysis in this paper faces several limitations. The relatively short sample period and the need for parsimony in the VAR model may constrain the robustness of the estimates. Additionally, the identification strategy, while commonly used, is based on timing assumptions that may not fully capture more complex contemporaneous interactions between policy instruments and macroeconomic outcomes. The cyclically adjusted primary balance, although a widely accepted proxy for discretionary fiscal policy, is also subject to measurement errors and estimation uncertainty.

Future research could build on this work by incorporating alternative identification schemes, such as sign restrictions or narrative-based approaches, to validate and strengthen causal inference. Extending the analysis to include distributional effects of fiscal policy, particularly on employment and income inequality, could also offer a richer perspective. Moreover, integrating nonlinear methods or regime-switching VARs may help capture asymmetries in fiscal policy effects during different phases of the business cycle or crisis periods.

## REFERENCES

- Alesina, A. and Ardagna, S. (2010), "Large changes in fiscal policy: taxes versus spending", *Tax Policy and the Economy*, Vol. 24 No. 1, pp. 35-68.
- Alesina, A., Favero, C. and Giavazzi, F. (2015), "The output effect of fiscal consolidation plans", *Journal of International Economics*, Vol. 96 No. S1, pp. S19-S42.
- Alesina, A., Favero, C. and Giavazzi, F. (2019), "Effects of austerity: Expenditure-and tax-based approaches", *Journal of Economic Perspectives*, Vol. 33 No. 2, pp. 141-162.
- Asandului, M., Lupu, D., Maha, L.G. and Viorică, D. (2021), "The asymmetric effects of fiscal policy on inflation and economic activity in post-communist European countries", *Post-Communist Economies*, Vol. 33 No. 7, pp. 899-919.
- Auerbach, A.J. and Gorodnichenko, Y. (2012), "Measuring the output responses to fiscal policy", *American Economic Journal: Economic Policy*, Vol. 4 No. 2, pp. 1-27.
- Blanchard OJ, Quah D. 1989. The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review* 79: 655-673.

- Blanchard, O. and Perotti, R. (2002), “An empirical characterization of the dynamic effects of changes in government spending and taxes on output”, *The Quarterly Journal of Economics*, Vol. 117 No. 4, pp. 1329–1368.
- Bobasu, A. (2015), “Fiscal policy effects in SEE countries”, *Economic Research*, Vol. 32 No. 1, pp. 85–98.
- Boiciuc, M. (2015), “Government spending shocks in Romania: Empirical evidence using SVAR”, *Romanian Journal of Economic Forecasting*, Vol. 18 No. 1, pp. 78–92.
- Caraiani, P. (2010), “Fiscal policy effects in a New Keynesian model with nominal rigidities: The case of Romania”, *Romanian Journal of Economic Forecasting*, Vol. 13 No. 3, pp. 36–50.
- Chudik, A., Mohaddes, K. and Raissi, M. (2021), “Covid-19 fiscal support and its effectiveness”, *Economics Letters*, Vol. 205, p.109939.
- Deb, P., Furceri, D., Ostry, J.D. and Tawk, N. (2022), “The economic effects of COVID-19 containment measures”, *Open Economies Review*, Vol. 33 No. 1, pp. 1-32.
- Forni, L., Giovannini, A., Lisi, G. and Paba, S. (2021), “COVID-19 fiscal policies: Evidence from the OECD using machine learning”, MEF Working Papers, No. 7/2021, Ministry of Economy and Finance, Rome.
- Franta, M. (2012), “Empirical evidence on fiscal and monetary policy interactions in the Czech Republic”, CNB Working Paper Series, No. 12.
- Furceri, D., Loungani, P., Ostry, J.D. and Pizzuto, P. (2021), “The rise in inequality after pandemics: can fiscal support play a mitigating role?”, *Industrial and Corporate Change*, Vol. 30 No. 2, pp. 445-457.
- Guajardo, J., Leigh, D. and Pescatori, A. (2014), “Expansionary austerity? International evidence”, *Journal of the European Economic Association*, Vol. 12 No. 4, pp. 949-968.
- Hinic, D. and Miletic, M. (2013), “Fiscal and monetary policy effects in Serbia”, *Panoeconomicus*, Vol. 60 No. 1, pp. 27–48.
- Ilzetzki, E., Mendoza, E.G. and Vegh, C.A. (2013), “How big (small?) are fiscal multipliers?”, *Journal of Monetary Economics*, Vol. 60 No. 2, pp. 239-254.
- International Monetary Fund. (2010). *World Economic Outlook, October 2010: recovery, risk, and rebalancing*. International Monetary Fund.
- Jemec, N., Kastelec, B. and Posedel, P. (2011), “How strong are fiscal multipliers in Slovenia?” IMAD Working Paper Series.
- Karagyozova-Markova, K., Deyanov, G. and Iliev, V. (2013), “Fiscal policy and economic growth in Bulgaria”, Bulgarian National Bank Discussion Paper DP/90.
- Krajišnik, M., Gligorić, D. and Gojković, B. (2019), “Effects of fiscal consolidation in Western Balkan Countries”, *Zbornik radova Ekonomskog fakulteta u Rijeci: časopis za ekonomsku teoriju i praksu*, Vol. 37 No. 2, pp.527-551.
- Mirdala, R. (2009), “Effects of fiscal policy shocks in transition economies”, *Journal of Economics and Finance*, Vol. 59 No. 3, pp. 321-337.
- OECD (2012), *Restoring Public Finances, 2012 Update*. OECD Publishing, Paris.
- OECD (2023), *OECD Economic Outlook*, Vol. 2023 Issue 1, No. 113, OECD Publishing, Paris.
- Perotti, R. (2007), “In search of the transmission mechanism of fiscal policy”. NBER Working Paper No. 13143.
- Petrevski, G., Bogoev, J. and Tevdovski, D. (2016), “Fiscal and monetary policy effects in three Southeastern European economies”, *Empirical Economics*, Vol. 50, pp. 415-441.
- Ramey, V.A. and Zubairy, S. (2018), “Government spending multipliers in good times and in bad: evidence from US historical data”, *Journal of Political Economy*, Vol. 126 No. 2, pp. 850-901.

- Romer, C.D. and Romer, D.H. (2010), “The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks”, *American Economic Review*, Vol. 100 No. 3, pp. 763-801.
- Sims, C.A. (1980), “Macroeconomics and reality”, *Econometrica: Journal of the Econometric Society*, pp. 1-48.
- Stock J.H. and Watson, M.W. (2012), “Disentangling the Channels of the 2007-2009 Recession”, Working Paper No. 18094, National Bureau of Economic Research.
- Tevdovski, D., Petrevski, G. and Bogoev, J. (2019), “The effects of macroeconomic policies under fixed exchange rates: A Bayesian VAR analysis”. *Economic Research-Ekonomska Istraživanja*, Vol. 32 No. 1, pp. 2138-2160.

## DIGITAL TECHNOLOGIES AND ICT TOOLS AS ENABLERS OF LEARNING AND INNOVATION IN THE SME SECTOR

**Faruk Hadžić**

*Sarajevo School of Science and Technology,  
Bosnia and Herzegovina  
faruk.hadzic@ssst.edu.ba*

**Ema Radončić**

*Sarajevo School of Science and Technology,  
Bosnia and Herzegovina  
ema.radoncic@stu.ssst.edu.ba*

### **ABSTRACT**

*This paper explores the relationship between the level of digital maturity and innovation of SMEs in Bosnia and Herzegovina, with a particular focus on the role of ICT as one of the drivers of innovation within the enterprise. Based on data collected from 304 enterprises from different sectors, a composite ICT index was created that quantifies the level of digital maturity of enterprises, as well as an innovation index that measures their activities in the field of innovation. Using descriptive statistics, sectoral segmentation, and K-means cluster analysis, the paper identifies significant differences in the level of digitalization between sectors, and also companies a positive relationship between a higher ICT index and greater innovation. The results show that enterprises with a higher level of digital maturity also achieve a higher level of innovation, while enterprises with a lower ICT index show a lower propensity to innovate. Based on the analysis, three clusters were formed that clearly distinguish enterprise profiles according to digital-innovation behavior. Furthermore, the findings of this research provide an empirical basis for the development of a targeted policy to support the digital transformation of sectoral SMEs, while identifying sectors with the lowest digital capacities as priorities for government intervention. The results obtained should be viewed with caution due to the subjectivity of the respondents' responses. As a practical suggestion, vouchers for digital tools and reskilling programs for the construction and hospitality sectors could accelerate the digital literacy of these companies. The paper aims to contribute to the understanding of the relationship between digital technologies and the innovation capabilities of enterprises in a transitional economic context.*

**Keywords:** *Digital maturity, ICT index, Innovation, SMEs, Cluster analysis, Small and medium-sized enterprises*

**JEL classification:** *O33, L26, M15*

### **1. INTRODUCTION**

In the modern business environment, which is characterized by accelerated technological development, many small and medium-sized enterprises are faced with various challenges, but also with the opportunities brought by digital transformation. ICT tools today do not only represent operational support, but are also a strategic resource that can stimulate innovation, improve communication, and facilitate learning within the organization. However, in the context of developing countries, to which Bosnia and Herzegovina belongs, the question remains to what extent these technologies are truly integrated into the practice of SMEs and

whether they bring concrete business benefits. Within the current literature, SMEs are recognized as a particularly sensitive but key part of the economy, which often has limited resources for implementing the digital transformation process. However, even partial digitalization of SMEs, which includes tools such as CRM systems, online communication, and informal sources of knowledge, can have a positive impact on the competitiveness of enterprises.

Precisely for this reason, the aim of this research is to examine the patterns of use of ICT tools in SMEs in Bosnia and Herzegovina, as well as their connection with innovation and organizational capacities. The focus is also placed on identifying the so-called "digital performance gap", where ICT investments do not lead to the expected results due to weak internal capacities or lack of support. The sample includes 304 companies from different sectors, and the collected data enabled the construction of an ICT index that quantifies the level of achieved digital sophistication. The findings show a positive correlation between the ICT index and the presence of innovation within the company, with companies that generate higher values developing new products and services more often. In addition, the cluster analysis identified three types of companies: digitally passive, highly innovative, and digitally advanced. It is noticeable that sectoral differences are pronounced, with sectors such as information and communications leading the way, while construction and part of the manufacturing industry lag behind, which together highlight the need for differentiated approaches in policies to support the digitalization of companies. The theoretical basis, methodological framework, key findings, and recommendations are presented below.

Recent studies confirm that digital transformation has become a crucial determinant of SME competitiveness across both developed and developing economies. According to the European Commission (2024) and OECD (2023) reports, digital tools enable knowledge diffusion and innovation scaling among SMEs. Similarly, Kraus *et al.* (2022) and Scuotto *et al.* (2023) highlight that SMEs in emerging markets face a persistent digital gap that limits their ability to exploit learning and innovation opportunities through ICT adoption.

### 1.1. Research problem

Although the importance of digital transformation in developing countries, such as Bosnia and Herzegovina, is very important for increasing productivity within companies, it remains an open question whether the application of digital technologies in SMEs brings concrete and measurable organizational benefits. Although digitalization is often perceived as an inevitable and positive process, empirical findings on its impact on innovation, productivity, and knowledge exchange are limited and often contradictory. Many companies invest in ICT tools, but such investments do not automatically guarantee a higher level of innovation, nor better internal communication. This phenomenon, known as the "digital performance gap", indicates the difference between the technical presence of digital solutions and their actual usability in practice. In addition, the application of ICT varies by sector, with the IT sector leading the way in digitalization, while traditional industries lag behind. However, differences also exist between companies within the same sector, indicating that organizational culture, knowledge, and access to resources play a key role.

In this context, the central research question is: "To what extent does the use of digital technologies in SMEs in Bosnia and Herzegovina contribute to innovation, knowledge exchange, and organizational learning? In accordance with the central research question, additional questions arise:

- Do companies with higher ICT intensity have a higher level of innovation?

- Can a digitally advanced type of company be distinguished?
- Are there any sectoral differences in the use of digital tools?
- Does informal and accessible digital knowledge, such as YouTube, Google, etc., have an impact on productivity growth?

The answers to these questions form the basis for setting hypotheses and empirical analysis aimed at better understanding the real contribution of digitalization to business processes in SMEs in Bosnia and Herzegovina.

## **1.2. Research objectives**

The main objective of this research is to examine how the level of digital sophistication of SMEs in Bosnia and Herzegovina affects their ability to innovate, learn, and adapt to rapid change. Digital sophistication was measured through an ICT index that includes the use of key digital tools, while innovation was analyzed through new or improved products, services, and communication. Organizational learning was also observed through access to informal digital sources of knowledge and openness to change.

Specific objectives include:

1. Construction of an ICT index to measure the intensity of use of digital tools in business.
2. Analysis of the relationship between ICT intensity and the level of innovation in enterprises.
3. Assessment of the impact of informal digital sources of knowledge on the perception of productivity and agility.
4. Identification of sectoral differences in the degree of digitalization, as well as by enterprise size.
5. Segmentation of enterprises through cluster analysis to recognize different types of digital practices.
6. Formulation of recommendations for SME support policies based on actual patterns of ICT use.

## **1.3. Research questions**

Based on the set objectives, this research addresses the following key research questions:

1. To what extent do small and medium-sized enterprises in Bosnia and Herzegovina use digital technologies and ICT tools in their daily business operations?
2. Is there a connection between the intensity of use of ICT tools and the degree of innovation of companies?
3. Does access to and use of informal digital sources of knowledge have a positive impact on the perceived productivity and flexibility of employees?
4. How do patterns of use of digital tools differ by industry sector and company size?
5. Can recognizable types of companies be identified based on the level of digital sophistication and innovation capacity?

## **1.4. Hypotheses**

Based on the previously stated research problem, objectives and questions, the following working hypotheses were formulated and tested using quantitative statistical methods:

- Hypothesis H1 - There is a positive and statistically significant relationship between the level of digital maturity measured by the ICT index and innovation performance in SMEs.
- Hypothesis H2 - The intensity of ICT tool usage differs significantly across industrial sectors, reflecting contextual and structural disparities.

- Hypothesis H3 - SMEs can be categorised into distinct digital-innovation profiles based on their ICT maturity and innovation outcomes.

## **2. LITERATURE REVIEW**

As part of the literature review related to this topic, several papers specifically discuss the importance of digital technologies and ICT tools as factors that encourage innovation within small and medium-sized enterprises. As part of the study, Požgaj and Vlahović investigated how Web 2.0 technologies (e.g., social networks, wikis, video-sharing platforms) facilitate informal learning. The poll was conducted at Zagreb's Faculty of Economics and Business and asked students about their sentiments towards platforms, including blogs and video-sharing sites. Web 2.0 promotes interaction, knowledge co-creation, and peer-to-peer interchange, leading to ongoing and user-driven learning outside conventional classrooms, according to the authors. The results demonstrate a substantial student preference for such technologies, making them almost ideal for lifelong informal learning. The study suggests that small and medium-sized enterprises (SMEs) may use Web 2.0 to share information and improve skills internally (Požgaj and Vlahović, 2011).

In another article, it is described how digital platforms sponsored by the EU operate as "digital ecosystems" to empower SMEs. The authors use biological ecosystem analogies to describe how open-source Web-2.0 infrastructures, such as DBE/OPAALS initiatives, improve interoperability, reduce costs, and promote innovation in SME networks. The digital infrastructure enables distributed software composition, evolutionary self-organization, and collaborative problem solving, resulting in scalable settings for SMEs to collaborate and compete. The significance of open standards and co-evolutionary interactions among participating enterprises is highlighted (Briscoe and Stanley, 2010).

In the other study, it is explored how Web 2.0 technologies, such as wikis, blogs, and social media, encourage knowledge exchange among Malaysian university students. According to a poll of 287 people, essential elements that increase participation with digital technologies include awareness, simplicity of use, infrastructure, and collaborative norms. The study indicates that effective knowledge-sharing ecosystems need a supportive culture, user-friendly platforms, and institutional awareness initiatives. While the study focused on education, the conclusions are applicable to SMEs: effective ICT adoption is dependent not just on technology availability, but also on organizational culture and support (Usman and Oyefolahan 2014).

When it comes to digital maturity of SMEs, it has been receiving increasing attention in various empirical studies as it has been shown to be a key factor between technological infrastructure and innovation performance of firms. For example, Bánhidi and Dobos (2025) developed the SME-DMI, which is presented as an entropy digital maturity index, designed for SMEs in Hungary, which includes tools, infrastructure and digital skills. Jie *et al.* (2025) investigate in their work how digital maturity in high-tech SMEs drives the level of innovative activities of firms. If we look at the industrial context, Krulčić (2025) developed a DAMA-AHP model for assessing digital maturity specifically adapted to SMEs, highlighting the challenge of expanding business capacities in relation to the targeted level of digitalization.

Other authors conducted a study to examine how ICT use influences competitiveness, innovation, and environmental consequences in the glass, ceramics, and cement concrete sectors. Using quantitative data and case studies, the authors show that ICT adoption considerably improves innovation capability and market competitiveness. Examples include

improved process management, product creation, and market response. Notably, the study considers environmental impacts: although some ICT applications cut emissions through optimised processes, others may unwittingly raise them due to increased energy use. The study suggests that while ICT is a strong engine of innovation and competitiveness, enterprises must consider environmental trade-offs while integrating digital technology (Ollo-López and Aramendia-Muneta, 2012).

Other research uses Rogers' diffusion of innovation theory to investigate how Internet use affects innovative attitudes and behaviours in 150 families in Xanthi, Greece. Key findings reveal that younger people and those with a higher degree of education have bigger information requirements and utilise the Internet more often. These users are more open to innovation, educated decision-making, and scientific perspectives. The study found that socioeconomic characteristics, personality traits, and communication behaviour all had a substantial impact on Internet adoption and innovation propensity. While the focus is on homes, the results give indirect insight into SMEs, where comparable demographic and behavioural determinants might support ICT-enabled innovation (Chatzoglou and Vraimaki, 2010).

Based on a 2010 survey of nearly 4,000 UK Open University students, this study investigates age-related digital technology access, attitudes, and study practices. Almost all respondents had access to a computer and the Internet; however, younger people used technology more often and were more positive about digital tools. However, there was no clear generational gap around the age of 30, which calls into question "net generation" ideas. Notably, older learners preferred deep and purposeful study approaches over superficial engagement with knowledge. Positive attitudes towards digital technology were linked to deeper learning techniques and improved academic approaches across all ages (Jelfs and Richardson, 2013).

On the other hand, another research looks into the role of e-marketplace platforms (centralised online marketplaces) in allowing Saudi SMEs to embrace e-commerce. The survey results from 157 SMEs show that technology infrastructure, organisational preparedness, and external assistance (e.g., legislation, training) are significant factors of successful adoption. Many SMEs reported a lack of knowledge, technical expertise, and trust as the primary impediments. The authors contend that e-mall platforms can reduce entrance barriers by offering common infrastructure, marketing reach, and streamlined payment/logistics systems. The study's model explains 62% of the variation in adoption intent, highlighting the need to support policies and information resources. Insights are important for SMEs worldwide, especially in Bosnia and Herzegovina, since they demonstrate how organised digital environments can drive adoption and innovation (Bahaddad *et al.*, 2012).

The purpose of another study is to look at how Malaysian manufacturing SMEs employ virtual R&D cooperation. According to the authors' findings from 163 survey responses, trust, communication quality, and IT infrastructure all have a major impact on the performance of virtual R&D teams. Companies that use tools such as video conferencing, document sharing, and project management platforms report increased innovation output, faster time-to-market, and cost savings. Cultural elements, including openness and top management, contribute to improved cooperation outcomes. The essay provides significant insight into how SMEs with minimal R&D resources may harness ICT technologies to form virtual innovation networks. This makes it particularly relevant for SMEs in Bosnia and Herzegovina striving to innovate through lean digital collaboration (Ale Ebrahim *et al.*, 2012).

Additionally, it is important to emphasize that research focused on the determinants of digital transformation in SMEs shows a complex set of factors from the technological and organizational domains. Omrani *et al.* (2023) use the T-O-E framework and quantitative methods on a large sample of SMEs to identify key drivers of digital technology adoption, given that digital infrastructure and existing innovativeness of the company have proven to be strong predictors of progress. Also, Ladu *et al.* (2024) in their work study the alignment of regulations as an institutional context and its impact on digital maturity, emphasizing in particular that the regulatory environment can slow down adoption even when technologies are available. Consequently, Tarutè *et al.* (2024) in their work on the digital orientation of SMEs emphasize that organizational culture and innovation orientation of the company interact with each other, but also condition the success of digitalization, especially in stressful circumstances such as pandemics.

Haseeb studies how ICT management techniques impact organisational innovation in Pakistani manufacturing enterprises (n=210). Using structural equation modelling, the article demonstrates that strategic ICT planning, IT infrastructure maturity, continual staff training, and committed ICT leadership all have a favourable impact on innovation outcomes—both product and process. Companies with more ICT maturity claim better information exchange, faster decision-making, and increased R&D efficiency. The report suggests holistic ICT governance frameworks and capacity-building, which are especially important for SMEs looking to systematically incorporate innovation using ICT technologies (Haseeb, 2015).

### **3. METHODOLOGY**

The research in this paper is based on a quantitative approach to examine the relationship between digital maturity and innovation of small and medium-sized enterprises in Bosnia and Herzegovina. Data were collected through a structured questionnaire during 2024, and after eliminating invalid responses, 304 companies from different sectors were included in the analysis.

In the process of constructing the ICT index, which includes 8 items, respondents rated the frequency of ICT tool use on a 5-point Likert scale. Items covered by this index are cloud services, collaboration platforms, CRM/ERP, e-invoicing/e-payments, cybersecurity practices, data analytics, website/e-commerce and digital document management. The innovation index includes binary/ordinal indicators of product, process, marketing, and organisational innovation in the last three years, and it is computed as the standardised mean.

The main independent variable is the ICT index, which is constructed by summing the scores on eight Likert-scale questions that assess the importance of digital tools in knowledge management. The maximum score is 40, and the reliability of the scale (Cronbach  $\alpha = 0.84$ ) companies the internal consistency. The Innovation Index, Perception of the benefits of informal digital knowledge, and Sectoral affiliation of companies were used as dependent variables. Associations were assessed with Pearson and Spearman correlations with coefficients ( $r, \rho$ ), 95% CIs,  $n$ , and  $p$ -values. Sectoral differences in the ICT index were tested using Kruskal–Wallis; we report  $H(df)$  and  $p$ , with Dunn–Bonferroni post-hoc comparisons where applicable. K-means clustering used z-standardised variables and was guided by elbow and silhouette diagnostics. The analysis consisted of three parts. First, a statistically significant positive relationship between ICT and the Innovation Index was confirmed through Pearson/Spearman correlation analysis. Second, the Kruskal–Wallis test identified differences in ICT indices across sectors, with the information and communication sector proving to be the

most digitally mature. Third, K-means cluster analysis identified three types of companies: digitally passive, highly innovative, and digitally advanced.

It is important to emphasize that the research is limited by the self-assessment of respondents and the cross-sectional design, but it still provides valid insights into digital behaviors and can serve as a basis for creating a typology of SMEs according to the degree of digital maturity and innovation. The research was anonymous and voluntary, so no personal data was collected. It is important to emphasize that there was no need for formal ethical approval because it was anonymous survey research. Anonymized data and code are available from the authors upon reasonable request.

#### **4. ANALYSIS OF RESULTS**

This section presents the results of a quantitative analysis of the collected data, with the aim of testing the hypotheses and better understanding the role of digital technologies and ICT tools in the context of innovation and internal capacities of small and medium-sized enterprises in Bosnia and Herzegovina. This analysis is based on descriptive statistical indicators, correlations, comparisons between sectors, and cluster segmentation of the sample. Special focus is placed on the aggregated ICT index, which is a composite measure of the intensity of use of various digital tools and technologies in business, as well as on the innovation index, which is constructed based on self-assessment of the presence of new products, services, and processes. It is important to emphasize that at the beginning, the distribution of the ICT index among enterprises is considered as an indicator of general digital maturity, and then the connection of the ICT index with the level of innovation of enterprises, as well as differences in digital maturity between the sectors of activity from which the enterprises come, is examined. For sectors with  $n < 20$ , results are presented descriptively without inferential claims. Where feasible, small sectors were consolidated into broader industry groups to improve statistical power. Finally, a cluster analysis is conducted to identify recognizable patterns in the behavior and capacities of enterprises.

*Table 1: ICT Summary*

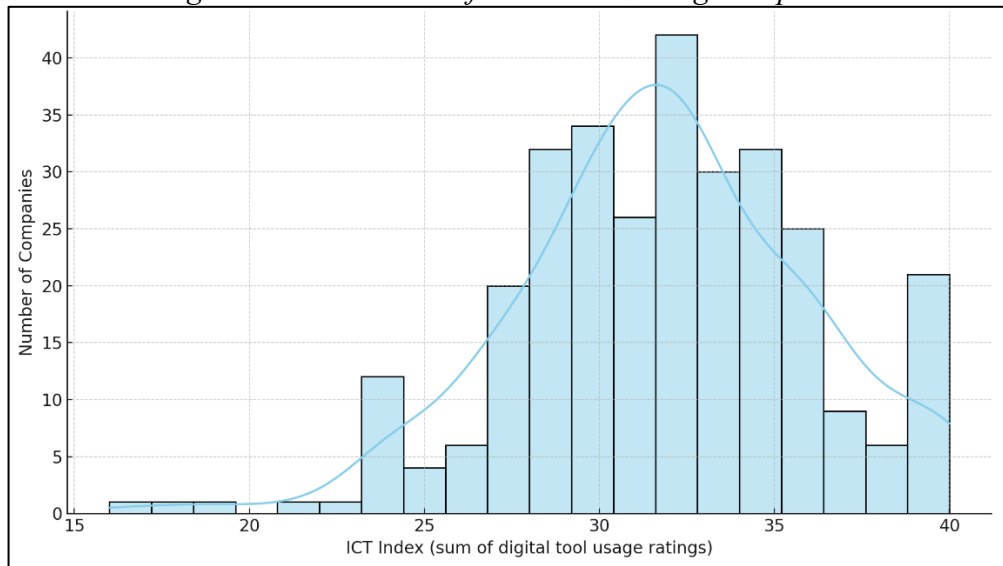
<b>Statistic</b>	<b>Value</b>
Number of Companies	304
Mean	31,75658
Standard Deviation	4,195702
Minimum	16
25th Percentile	29
Median	32
75th Percentile	35
Maximum	40

*(Source: Author's creation)*

In Figure 1, a histogram of the distribution of the ICT index is shown, which further strengthens the previously mentioned findings. In the above chart, it is clearly visible that most companies gravitate towards the mean values of the index, with a slight decrease in frequency at both ends of the spectrum. This distribution indicates a general technological balance among companies, while at the same time there is a smaller number of entities with high digital capacities, which could represent potential "locomotives" of digital transformation in the SME sector. Taking into account all the indicators shown, it can be concluded that the analyzed companies,

although on average they show a solid level of digital maturity, still have room for further development, which is especially important when considering the advantages that digital transformation can bring in terms of innovation, efficiency and competitiveness.

Figure 1: Distribution of ICT Index among Companies



(Source: Author's creation)

In order to better understand the level of digital maturity among companies, an analysis of the average ICT index by sectors from different industries was conducted. The results are presented in Table 2, where the values of the mean ICT score, standard deviations, and the number of companies included in the analysis are given for each sector. The results obtained show the present differences in the level of adoption and application of digital tools among sectors, although these differences are, on average, relatively moderate. The sectors showing the highest average ICT index are Real Estate (L) with a value of 34.5, then Health and Social Work (Q) with 34.17, and Information and Communication – ICT Services (J) with 32.86. These industries are characterized by a greater need for digital infrastructure solutions, automation and software support in everyday work, which may explain the higher level of integration of ICT tools. On the other hand, sectors such as Construction (F) with 29.5, Hotels and Accommodation (I) with 30.8, and Administrative and Support Activities (N) with 30.29, record lower ICT index values, indicating a lower level of digital transformation in these sectors. Special attention should be paid to the ICT services sector (J), which not only has one of the highest ICT indices, but also a relatively high number of companies in the sample (n=50), which allows for a more reliable interpretation of the results. Its average (32.86) further confirms that companies that develop or distribute digital solutions themselves are among the leaders in their use and can serve as a reference point for other sectors.

Table 2: Kruskal–Wallis test of ICT index across sectors, Dunn–Bonferroni post-hoc<sup>1</sup>

Sector	Number of Companies	Mean ICT Index	Standard Deviation
Activities of providing accommodation, and preparation and serving of food (Hotel and catering) (I)	5	30,80	1,788854

<sup>1</sup> Sectors with n<20 are reported descriptively; inferential comparisons are not interpreted.

Administrative and support activities (N)	7	30,28	2,058663
Agriculture, fishing and forestry (A)	1	32,00	-
Arts, Entertainment and Recreation (R)	2	30,00	2,828427
Construction (F)	20	29,50	3,762698
Education (P)	3	31,00	5,567764
Financial and insurance activities (K)	23	32,17	4,030026
Health and social work (Q)	6	34,16	4,996666
Information and communication - ICT services (J)	50	32,86	4,463046
Manufacture (C)	43	31,27	3,737363
Mining and quarrying (B)	2	31,00	1,414214
Other service activities (S)	57	31,33	4,680252
Production and supply of electricity, gas and steam (D)	5	32,80	3,193744
Professional, scientific and scientific activities (M)	13	31,85	4,278749
Public administration and defence (O)	1	35,00	-
Real Estate Business (L)	4	34,50	1,290994
Transport and storage (H)	6	33,17	3,125167
Wholesale and retail trade (G)	56	31,82	4,390515

(Source: Author's creation)

Standard deviations in most sectors range from 3 to 5, which means that within the sector, there are companies with different levels of digital maturity. This is especially pronounced in sectors such as Other Service Activities (S) and Education (P), which tells us about the possible space for targeted improvements through sector strategies and digital education. Taking into account all the values presented, it can be concluded that sector affiliation to a certain extent affects the level of digital maturity, and that future digital transformation policies should be designed to take into account sector specificities and targeted investments, especially in industries with a below-average ICT index.

In order to better understand the structure of companies in terms of their digital maturity and innovation capacities, a cluster analysis was conducted, which grouped companies based on the value of the ICT and innovation index. The aim of the analysis was to identify typical patterns among companies and clear differences between previously formed groups. The results are shown in Table 3 and Figure 2.

As part of this part of analysis, three clusters of companies were identified, which differ significantly in terms of the average value of the ICT and innovation index:

- Cluster 0 includes 92 companies with the highest average ICT index of 36.27 and, at the same time, a solid average innovation index of 2.64. This group of companies is called "*digitally advanced*" - technologically advanced companies that are simultaneously inclined to innovate products, processes, or some other business models. The low standard deviation of the ICT index (2.27) indicates a consistently high level of digitalization within the group.

- Cluster 1, with 126 companies, shows the highest average innovation index of 3.11, but also the lowest ICT index of 29.62, so this group can be characterized as “*highly innovative*”, which indicates that innovation does not always have to be based on high digital sophistication, but can also rely on other sources, such as human capital, knowledge and creativity.
- Cluster 2, which includes 86 companies, has a low average innovation index (1.80) and ICT index (30.06), which indicates “*digitally passive*” companies, which show a lower level of digitalization and introduce innovations less often. This group of companies can represent a target population for digital and innovation incentive policies.

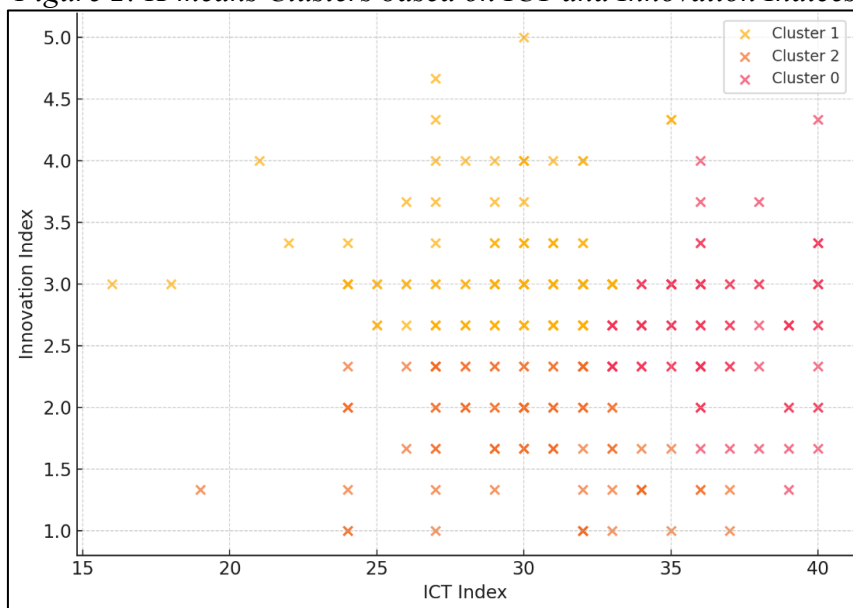
Table 3: Cluster profiles and validation metrics (means by cluster; silhouette, CH, within-cluster SSE)

Cluster	ICT Index - Mean	ICT Index - SD	ICT Index - N	Innovation Index - Mean	Innovation Index - SD	Innovation Index - N
0	36,27174	2,268287	92	2,644928	0,5083055	92
1	29,61905	3,121172	126	3,108466	0,4691920	126
2	30,05814	3,340851	86	1,798450	0,4406741	86

(Source: Author's creation)

The distribution shown in Figure 2 confirms these differences and indicates the spatial separation between clusters in two-dimensional space, further highlighting the clear polarization between companies that invest in digital tools and innovation at the same time, and those that lag behind on both grounds. In conclusion, the cluster analysis confirms that high digital maturity does not automatically guarantee high innovation, but also that clear patterns of behavior can be identified among companies that combine both elements.

Figure 2: K-means Clusters based on ICT and Innovation Indices



(Source: Author's creation)

## **5. DISCUSSION**

The obtained research results showed several key patterns that provide a better insight into the state and trends of digital maturity and innovation in small and medium-sized enterprises in Bosnia and Herzegovina. These results provide a better overview in the context of previously set goals and research questions, and testing the formulated hypotheses.

The first objective of the research was to quantitatively measure the level of use of information and communication technologies in the SME sector, through the construction of a composite ICT index. The results of the descriptive analysis showed that the average value of the ICT index was 31.76, on a scale of 0–40, with a standard deviation of 4.19. These values indicate that most companies have a moderate to high level of digital integration, considering the included indicators such as the use of computers and the Internet in everyday business, presence on social networks, online communication with clients, as well as internal organizational solutions based on digital tools. Furthermore, the data presented in Table 1 and Figure 1 illustrate an even distribution among companies, with a slight emphasis on higher values, thus confirming the existence of a certain foundation for further digital transformation. In this sense, it can be concluded that the first research objective has been partially achieved, given that the ICT index successfully quantifies the level of digital maturity, but at the same time differences in digital progress among companies are also revealed. Observed associations may be partly driven by unmeasured factors such as company size, management quality, workforce education, or industry competition.

When it comes to the sectoral analysis, shown in Table 2, significant differences between sectors were observed, which directly answers the second research question, which is whether and to what extent the sector of activity affects the digital maturity of companies. Namely, sectors such as information and communication technologies, healthcare, and real estate show above-average values of the ICT index, while construction, hospitality, and administrative support activities show significantly lower values. These results support Hypothesis H2, according to which the level of digital maturity differs significantly between sectors. Given that sectors with greater exposure to market competitive pressures and the need for sophisticated management, e.g., the ICT sector, financial activities, are also more digitally advanced, this further confirms the importance of contextual factors.

A particularly interesting insight emerges from the cluster analysis presented in Table 3 and Figure 2, which approached companies from a combined perspective of their digital maturity and innovative potential. Cluster 0, with the highest ICT index and a solid innovation index, represents the so-called “digitally advanced”. Their position confirms Hypothesis H1 that higher use of ICT tools is positively correlated with a higher probability of introducing innovations. However, this correlation is not perfect, since Cluster 1 contains companies with the highest average innovation index, but also the lowest ICT index, which indicates the existence of “analog innovators”, i.e., companies that innovate despite a limited technological environment. This result directly contributes to testing Hypothesis H3, according to which there is no strict dependence between digital maturity and innovativeness. The research shows that ICT can be a strong accelerator of innovation, but is not always a necessary condition, which indicates the existence of alternative paths to innovation, such as reliance on human capital, creative management methods, or traditional business models that are adapted to new market demands. The third identified cluster (Cluster 2), characterized by low values of both indices, represents enterprises that are both technologically and innovationally inert. This segment indicates the existence of deep barriers to the adoption of both ICT and innovation,

suggesting the need for systematic institutional support, including subsidies, educational programs, and facilitated access to digital resources.

Considering all that has been shown, the research confirms that the first research objective has been fully achieved, while the second and third objectives have enabled the validation of two out of three hypotheses. The analysis also pointed out the limitations of previous approaches that rely exclusively on technological indicators, without insight into the actual creative and adaptive capacity of the company. A significant contribution of this research is reflected in the methodological approach that combines descriptive and exploratory statistics with cluster analysis, thus enabling a multi-layered insight into the ways of behavior within the SME sector. Also, the results indicate the need for personalized digital and innovation support policies that will not be linear, but differentiated in accordance with the specificities of the cluster and sector. The “analog innovators” cluster should be regarded as a preliminary typology. Future work should validate it through qualitative case studies in order to identify compensating capabilities that substitute for digital tools. In addition, the results can serve as a basis for the development of diagnostic tools intended for SMEs, which would enable self-assessment of their digital maturity and identification of obstacles to the introduction of innovations. In this context, a digital-innovation readiness index can be developed that would serve as a tool for decision-making at the company level, but also as an indicator for shaping public policies. In conclusion, although the research confirmed certain expected patterns of business behavior, such as the positive association between ICT and innovation, it also revealed important exceptions that indicate the complexity of the real business environment.

This study is subject to common-method bias risks, self-reporting, and a cross-sectional view. Sectoral inference is constrained by small  $n$  in multiple categories. Future work should combine administrative usage data with survey responses, expand sector coverage, and employ panel designs to track innovation outcomes.

## **6. CONCLUSIONS**

The research conducted as part of this paper provided important insight into the level of digital maturity and innovation of small and medium-sized enterprises in Bosnia and Herzegovina, taking into account the use of information and communication technologies as a key driver of modernization and competitiveness. Through the construction of a composite ICT index, sectoral and cluster analysis, and the interconnection of digital maturity and innovation, results were obtained that have both theoretical and practical significance. Conclusions are necessarily limited by self-reported data, cross-sectional design and small sub-samples in several sectors.

First, the research showed that the majority of the observed enterprises use ICT tools to some extent in their daily business, with the average value of the ICT index being relatively high. However, there is evident variation among enterprises, both based on the sector of activity and their overall digital and innovation profile. Sectors such as information and communication services, healthcare and the financial sector lead in digital maturity, while sectors such as construction and hospitality are significantly below average. Secondly, through cluster analysis, three groups of companies were identified: digital leaders with strong ICT capacities and relatively high innovation scores, then analog innovators, who innovate without a strong digital infrastructure, and digitally-innovation-inert companies that lag behind in both aspects. It is precisely this segmentation that provides additional value in terms of creating targeted support measures, which is particularly important in the context of formulating public policies and digital transformation strategies. Thirdly, although a positive correlation is confirmed

between the use of ICT tools and innovation, the results indicate that digital maturity is not the only or exclusive prerequisite for innovation. There are companies that, despite a low level of digitalization, demonstrate a high level of innovative behavior, which implies the need for a broader approach in understanding innovation factors, including human capital, managerial practices, organizational culture and market impulses.

Finally, the research results can serve as a basis for shaping future development interventions. The development of tools for self-assessment of the digital readiness of the SME sector, institutional support through financial and educational programs, as well as the establishment of sector-differentiated digital and innovation acceleration policies are recommended. The data also indicate the need for additional research that would include longitudinal analysis, qualitative insights, and broader geographical coverage. Policy-wise, under-digitised sectors could benefit from digital adoption vouchers that are tied to demonstrable use, targeted reskilling bootcamps in data and cybersecurity for SMEs, and a self-assessment tool for SME digital readiness embedded in local business-support programs. In conclusion, although limited in scope and time frame, the research enabled the validation of key assumptions and opened the space for further scientific research and practical engagement in the field of digital transformation and innovation in the SME sector.

## REFERENCES

- Ebrahim, N. A., Ahmed, S., Rashid, S. H. A., Wazed, M. A. and Taha, Z. (2012), "Virtual collaborative R&D teams in Malaysian manufacturing SMEs", *Advanced Materials Research*, Vol. 433, pp. 1653-1659.
- Bahaddad, A.A., AlGhamdi, R. and Houghton, L. (2012), "To what extent would E-mall enable SMEs to adopt e-commerce?", *International Journal of Business and Management*, Vol. 7, No. 22, pp. 123-132.
- Bánhidi, Z. and Dobos, I. (2025), "An entropy-based digital maturity index for small and medium enterprises", *Central European Journal of Operations Research*, pp. 1-19.
- Briscoe, G. and Stanley, J. (2010), "The ABC of digital business ecosystems", *Communications Law*.
- Chatzoglou, P.D. and Vraimaki, E. (2010), "Investigating internet usage as innovation adoption: a quantitative study", *Journal of Information, Communication and Ethics in Society*, Vol. 8 No. 4, pp. 336-355.
- European Commission. (2024), *The Digital Economy and Society Index (DESI) 2024 Report*, Brussels.
- Haseeb, H.M. (2015), "The role of ICT management to achieve organizational innovation", *International Journal of Applied Business and Economic Research*, Vol. 13 No. 2, pp. 689-702.
- Jelfs, A. and Richardson, J.T.E. (2013), "The use of digital technologies across the adult life span in distance education", *British Journal of Educational Technology*, Vol. 44 No. 2, pp. 338-351.
- Jie, H., Zhang, W., Liu, Y. and Li, X. (2025), "Digital maturity, dynamic capabilities and innovation in high-tech SMEs", *Technological Forecasting and Social Change*, Vol. 218, 123456.
- Kraus, S., Jones, P. and Kailer, N. (2022), "Digital transformation and SMEs: a systematic review", *Technological Forecasting and Social Change*, Vol. 182, 121882.
- Krulčić, E. (2025), "A dynamic assessment of digital maturity in industrial SMEs: DAMA-AHP model", *Technologies*, Vol. 13 No. 7, p. 282.

- Ladu, L., Sorrentino, A. and Petroni, G. (2024), "Technology adoption and digital maturity in a conservative institutional environment", *Technology in Society*, Vol. 76, 102345.
- OECD (2023), *SME Digitalisation Outlook 2023*, OECD Publishing, Paris.
- Olló-López, A. and Aramendia-Muneta, M.E. (2012), "ICT impact on competitiveness, innovation and environment", *Telematics and Informatics*, Vol. 29 No. 2, pp. 204-210.
- Omrani, N., Rejeb, N., Maalaoui, A., Dabić, M. and Kraus, S. (2023), "Drivers of digital transformation in SMEs: evidence from a large multisectoral sample", *IEEE Transactions on Engineering Management*.
- Požgaj, Ž. and Vlahović, N. (2011), "Students' readiness for informal learning using Web 2.0 services", *International Journal of Knowledge and Learning*, Vol. 7 No. 1-2, pp. 54-70.
- Rogers, E.M. (2003), *Diffusion of Innovations* (5th ed.). New York: Free Press.
- Scuotto, V., Del Giudice, M. and Carayannis, E.G. (2023), "Knowledge-driven innovation in SMEs: the role of digital platforms", *Journal of Knowledge Management*, Vol. 27 No. 1, pp. 59-78.
- Tarutė, D., Duobienė, J. and Apanavičienė, R. (2024), "Exploring digital orientation of SMEs: the effect of digital orientation and capability on digital transformation", *Small Business Economics*.
- Tornatzky, L.G. and Fleischer, M. (1990), *The Processes of Technological Innovation*, Lexington Books.
- Usman, M. and Oyefolahan, I.O. (2014), "Encouraging knowledge sharing using Web 2.0 technologies in higher education: a survey", *The Electronic Journal of Knowledge Management*, Vol. 12 No. 1, pp. 43-52.

## WOMEN, WORK, AND BIRTH RATES IN SOUTHEASTERN EUROPE: A REGIONAL PANEL PERSPECTIVE

**Marija Trpkova - Nestorovska**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*Marija.Trpkova-Nestorovska@eccf.ukim.edu.mk*

### **ABSTRACT**

*The study explores the interdependence between female labor force participation and fertility dynamics in eight Southeastern European countries during 2000 – 2023, using the crude birth rate as a consistent macro level proxy for fertility in panel regression estimation. The analysis covers eight countries and incorporates economic, educational, and demographic control variables. Results reveal that short-term increases in female employment and tertiary education enrollment are significantly associated with declining birth rates, while economic growth shows a positive effect. Other factors, including part-time employment, urbanization, and parental leave policies, are not statistically significant in the short run. The study underscores the complexity of natality dynamics in transitioning economies and highlights the need for supportive family policies.*

**Keywords:** *Birth rates, Female labor force participation, Southeastern Europe, Panel regression, Education, Economic development*

**JEL classification:** *J13, J16, O52, C23, I23, F63*

### **1. INTRODUCTION**

In recent decades, Southeastern Europe has experienced dramatic demographic changes, marked by declining birth rates, ageing populations, and large-scale emigration. These changes have posed significant socio-economic challenges, including shrinking labor forces, reduced economic dynamism, and increasing pressure on welfare systems. Among the various factors influencing birth rates, female labor market participation has emerged as a key determinant. As women become increasingly integrated into the labor market, questions arise about how this shift affects reproductive behavior and demographic trends.

This paper aims to explore the relationship between female labor force participation and birth rates in eight Southeastern European countries: Albania, Bosnia and Herzegovina, Bulgaria, Greece, North Macedonia, Montenegro, Romania, and Serbia. Using a panel regression framework covering the period from 2000 to 2023, the study investigates short-term birth rates responses to changes in women's employment, while controlling for other demographic, educational, and economic variables. In particular, the study considers GDP per capita, tertiary education, urbanization, female part-time employment, and early childhood education as additional explanatory variables.

This research is motivated by both empirical gaps and policy relevance. While much has been written about birth rate dynamics in Western Europe and OECD countries, relatively few studies have focused specifically on Southeastern Europe, a region with distinct historical, economic, and institutional characteristics. By providing a regional empirical analysis, this paper contributes to a more nuanced understanding of birth rate determinants in post-socialist

and transitioning economies. It also offers valuable insights for policymakers seeking to balance female employment and family formation in the face of demographic decline.

The remainder of the paper is organized as follows: Section 2 reviews the relevant literature, Section 3 describes the countries in the sample and trends in birth rates and female employment, Section 4 presents the data, methodology, and empirical results, Section 5 discusses the findings, Section 6 presents certain limitations of the study and Section 7 concludes with implications and recommendations.

## 2. LITERATURE REVIEW

Birth rates in Southeastern European countries had a sharp decline in the past two decades. Only in 2023 Albania has birth rate 10.24‰, Bosnia and Herzegovina 7.71‰, Bulgaria 8.90‰, Greece 6.80‰, North Macedonia 9.20‰, Montenegro 11.20‰, Romania 8.00‰ and Serbia 9.20‰. Birth rates are only part of the declining demographic situation that these countries are facing. Continuous emigration and population ageing contribute to unsustainable economies and labor market imbalances. To address the issue of declining birth rates, the first step is to identify the key factors that determine their dynamics. This paper considers female participation rate as the main factor, with other factors taken as control variables that include tertiary education, GDP per capita, urbanization, female part time employment and early childhood education. Previous papers have considered these variables for distinct groups of countries, while when it comes to Eastern European countries further research could provide new insight. A considerable body of literature has examined the interplay between female labor force participation and fertility, yielding mixed findings across different regions and methodological approaches.

Recent work for Central, Eastern and Southeastern Europe documents how the post 1990 institutional transition and policy regimes shaped fertility to very low levels, with temporary rebounds followed by renewed decline. A key synthesis is Frejka and Gietel-Basten (2016), who show that across Central, Eastern and Southeastern Europe, modest policy expansions rarely shifted quantum fertility, with responses concentrated at specific parities and often short lived. They also stress cohort postponement and migration as structural headwinds. At the country level, demographers from Southeastern Europe provide detailed evidence: for Albania, Arjan *et al.* (2009) and coauthors trace a rapid fall of total fertility rate from historically high levels toward sub-replacement already by the 2000s, driven by marriage postponement and socio-economic change. For Serbia, recent work from Nikitović and Magdalenić *et al.* (2024) using 2022 census data documents gender specific cohort fertility patterns and continued very low fertility. For the former Yugoslav area Ćipin *et al.* (2020) analyze cohort trends and parity progression.

Mishra and Smyth (2010) conducted a panel cointegration analysis for 28 OECD countries and found both unidirectional and bidirectional long-run Granger causality between female labor force participation and total fertility rate, contingent on measurement and time period. Bloom *et al.* (2009), using data from 97 countries between 1960 and 2000, identified an inverse relationship between female labor participation and fertility. However, more recent studies such as Luo and Jarosz (2024), using German regional data, suggest that female job creation can positively influence fertility, underscoring the contextual nature of the relationship.

Education is another critical determinant. Vasireddy *et al.* (2023) found that rising tertiary education in Europe correlates with delayed fertility, higher rates of childlessness, and fewer

second births. In Taiwan, Bharati *et al.* (2021) documented that the expansion of higher education significantly reduced fertility among women up to age 30. Similarly, Shirahase (2000) emphasized the role of education in delaying marriage in Japan, and Aydoğan and Uslu (2019) showed that increased female education significantly reduces adolescent fertility across 53 countries.

Economic development also plays a significant role. Kögel (2004) and Sobotka *et al.* (2011) argue that fertility tends to follow economic cycles, with improved economic conditions encouraging family formation. Conversely, periods of economic crisis and uncertainty such as the 2008 financial downturn are linked to declining fertility intentions, especially in contexts with limited family policy support (Adserà, 2004).

Beyond classic reviews, newer causal studies quantify policy effects in the region. In Poland, the “Family 500+” cash benefit reduced mothers’ labor force participation by approximately 2 to 3 percentage points, indicating strong labor supply responses that may interact with fertility, while evidence on births is mixed and parity specific (Magda *et al.*, 2018). In Hungary, quasi experimental and event history work finds targeted effects at higher parities: cash support and tax relief increased third-birth risks, especially among low educated (cash) and higher educated (tax relief) parents, though aggregate total fertility rate gains remained modest and time limited (Spéder *et al.*, 2020). Outside the EU but regionally relevant, Russia’s “maternity capital” raised long run fertility by approximately 0.15 children per woman, with substantial timing (rescheduling) effects (Słomimczyk and Yurko, 2014). Cross OECD syntheses conclude that Early Childhood Education and Care expansion, leave design, and tax benefit architecture can raise fertility at the margin, but effects depend on parity targeting, generosity, and labor market conditions (Luci-Greulich and Thévenon, 2013).

Family and labor market policies have also been studied extensively. Gauthier (2007) reviewed empirical studies and concluded that the effectiveness of parental leave and childcare support varies across institutional settings. Nordic countries show more positive fertility responses compared to Southern and Eastern European countries, where institutional support for work family reconciliation is often weak.

Contrary to a common perception, fertility and female employment dynamics in the Balkans and Southeastern Europe have been extensively studied by regional scholars. Beyond the cross-country synthesis by Frejka and Gietel-Basten (2016), notable contributions include Arjan *et al.* (2009) on Albania’s transition from high to low fertility, Nikitović and Magdalenić (2024) on Serbia’s cohort fertility and parity patterns and Čipin *et al.* (2020) on the former Yugoslav territory. This paper adds to the previous research by applying a regional panel regression approach to examine fertility determinants in the context of rapid institutional, demographic, and labor market changes.

### **3. OVERVIEW OF COUNTRY-SPECIFIC DYNAMICS**

This section presents an overview of birth rate trends and female labor force participation in eight Southeastern European countries. The relationship between female employment and birth rates is explored in the context of country specific labor market changes, demographic patterns, and EU accession processes. It is hypothesized that an increase in female participation rate has negative effect onto birth rates.

The data presented in Figure 1 reveal a consistent inverse relationship between female labor force participation and birth rates. In most cases, rising female employment coincided with declining birth rates.

In Albania, female labor force participation increased gradually since the early 2000s. The sharp rise in 2011 coincided with methodological revisions to the Labor Force Survey aligned with Eurostat standards. Reforms related to gender equality and EU accession (particularly after Albania’s EU candidate status in 2014) also supported this trend. Birth rates, meanwhile, declined consistently due to delayed marriage, emigration, and rising female education.

Bosnia and Herzegovina remains characterized by low female employment rates, driven by structural unemployment, weak institutional support, and traditional gender norms. Despite low participation, natality rate remains low as well reflecting the impact of economic insecurity, youth emigration, and postponed family formation in the post-conflict context (Billingsley, 2010).

Bulgaria experienced a short-term rise in natality rate from 2003 to 2009 due to pronatalist policies (extended maternity leave, child allowances) and favorable cohort effects (women born during the 1980s when birth rates were relatively high entered prime reproductive ages). However, this was followed by a decline due to the 2008-2009 financial crisis and the resulting economic uncertainty. Birth rates remain low, driven by persistent low natality preferences, population ageing, and continued emigration of young families, with birth rates steadily declining until 2023 (NSI, 2012). Female labor participation remained high reflecting its socialist legacy (institutionalized women’s employment, supported by universal childcare and maternity leave), dual-earner norms, and state supported childcare. The labor market concentrated in female dominated sectors like education and healthcare. EU accession and gender equality policies further strengthened female employment (Philipov and Jasilioniene, 2008; Eurostat, 2023).

Figure 1: Presentation of birth rate and female participation rate per country



(Source: Author’s presentation)

Greece represents a typical Southern European pattern, where increasing female labor force participation during the 2000s coincided with persistently low birth rates. Structural rigidities in the labor market, gender inequality in caregiving, and limited public childcare provision have contributed to exceptionally low birth rates. The economic crisis of 2008 further delayed childbearing, especially among younger cohorts. As Adserà (2004) notes, birth rates are particularly sensitive to labor market instability in countries lacking comprehensive family support policies.

In North Macedonia, female labor force participation has gradually increased but remains low, compared to EU countries. Over the same period the birth rate marks a constant decline. The rise in women's educational attainment and growing urbanization have contributed to natality postponement, especially in the absence of institutional support for resolving employment and caregiving roles. The weak effect of parental leave policies and persistent gender role expectations further limit fertility responses to economic improvement (Matysiak and Vignoli, 2008).

Following Montenegro's EU candidacy in 2010 and the launch of accession negotiations, the country initiated a series of labor market and gender equality reforms that significantly boosted female employment. The sharp rise in Montenegro's birth rate in 2010, followed by a rapid decline, is likely the result of statistical adjustments rather than a genuine demographic shift. Improvements in civil registration, introduced as part of post-independence reforms and EU alignment efforts, may have led to a temporary increase in reported births due to delayed registrations or methodological changes. This anomaly was followed by a return to the broader trend of declining birth rate, driven by continued youth emigration, delayed family formation, and economic uncertainty (Monstat, 2011; Eurostat, 2012).

Romania marks steady increase in the female labor participation rate, with sharp decline in 2009 due to the impact the global financial crisis, which severely affected the country's economy and labor market, especially for women. It caused job losses in sectors with high female employment (education, health care and social services) and reduced opportunities due to austerity cuts. Increased unpaid care responsibilities also led many women to temporarily leave the labor market. The birth rate does not mark the expected downwards trend. With slight fluctuations it shows increase in the 2002 – 2009 period, mostly due to cohort effects and improved economic conditions prior to the global financial crisis. The subsequent period saw stagnation, as the crisis and austerity policies weakened social support and increased economic uncertainty. After 2020, birth rates declined sharply due to the compounded impact of the COVID-19 pandemic, continued emigration, and demographic ageing.

The rapid rise in Serbia's female labor force participation after 2010 reflects labor market reforms, EU aligned gender equality policies, and rising educational attainment among women. Conversely, the sharp drop in birth rates in 2004 resulted from cohort effects where generation of women born in the 1970s who entered peak fertility in the late 1990s, had already experienced delayed childbearing due to 1990s economic crisis and political instability. Also, after a mild recovery in births following the 1999 NATO conflict, fertility resumed its long-term declining trend, driven by uncertainty, emigration, and low fertility preferences. Changes in data registration, such as separation of data from Kosovo and Metohija may have influenced reported birth numbers around 2003 – 2004.

A general inverse relationship between female labor force participation and birth rates across Southeastern European countries is evident. In most cases, rising female employment is

accompanied by declining birth rates, reflecting broader demographic transitions such as delayed childbearing, increased education, and changing gender roles. While labor market reforms and EU integration have improved female employment across the region, birth rates continue to decline due to structural challenges, including inadequate family support policies, population ageing, and emigration of young adults. Thus, female labor participation is a key, but not exclusive, factor influencing birth rate dynamics in Southeastern Europe.

#### 4. DATA, METHODOLOGY, AND EMPIRICAL RESULTS

This study employs a panel dataset of eight Southeastern European countries: Albania, Bosnia and Herzegovina, Bulgaria, Greece, North Macedonia, Montenegro, Romania, and Serbia, covering the period from 2000 to 2023 ( $n = 192$ ). The dependent variable is the crude birth rate (live births per 1.000 population), and the key independent variable of interest is the female labor force participation rate. Additional control variables include GDP per capita, urbanization rate, female tertiary education attainment, female part-time employment, and childcare enrollment (missing data in this variable were linearly interpolated between available data points).

The study employs the crude birth rate as the dependent variable. While the total fertility rate would be the conceptually preferred measure of fertility behavior, the birth rate was selected due to its greater temporal and cross-country coverage and consistent availability across the eight Southeastern European countries for the full 2000 to 2023 period. From a modelling perspective, the birth rate offers a continuous and comparable annual series, minimizing missing values and enabling robust panel estimation. Although the birth rate is influenced by age structure and migration, its high correlation with the total fertility rate makes it a practical proxy for fertility trends at the macro level.

Parental leave duration was introduced as a time-invariant moderator variable, collected from national labor laws. An interaction term ( $parent\_leave * d(labor\_part)$ ) was constructed to test whether paternal leave policy moderates the effect of labor participation on birth rates. A moderator affects the strength or direction of the relationship between independent and dependent variables (e.g., parental leave moderating the link between labor participation and fertility). A control variable accounts for confounding influences (e.g., GDP per capita). Detailed information about the variables is presented in Table 1.

Before running the unit root tests, a panel model with fixed effects is estimated and the residuals are tested for cross sectional dependence. All four tests (Breusch-Pagan LM, Pesaran scaled LM, Bias-corrected scaled LM and Pesaran CD) reject the null hypothesis at 1% significance level, which confirms the presence of cross-sectional dependence. This requires implementation of second-generation unit root tests (like Pesaran CIPS – Cross-sectionally Augmented Im, Pesaran, and Shin), (Pesaran, 2007) for stationarity testing. Cross-sectional dependence is expected since panel models include macroeconomic and demographic variables countries that are economically and politically interconnected, exposed to common shocks and linked via migration, trade, and labor markets. Traditional first-generation panel unit root tests (Levin-Lin-Chu (LLC) or Im-Pesaran-Shin (IPS)) assume cross-sectional independence, which is not the case in this research sample.

The Pesaran CIPS test (2007) was applied to determine stationarity, using a maximum lag of 1 with AIC-based selection due to the limited time span. Because the panel covers only 24 annual

observations, the maximum lag length was restricted to 1 based on AIC selection for statistical parsimony.

Table 1: Variables used for estimation of the panel regression model

Variable	Description	Source
<i>birth_rate</i>	Number of live births occurring during the year, per 1,000 population estimated at midyear	World Bank
<i>labor_part</i>	Proportion of the female population ages 15-64 that is economically active	World Bank
<i>GDP</i>	GDP per capita, expressed in constant international dollars, converted by purchasing power parities (constant 2021)	World Bank
<i>urban</i>	People living in urban areas (as defined by national statistical offices) as % of total population	World Bank
<i>tertiary</i>	Female tertiary enrollment	UNESCO Institute for Statistics; Eurostat
<i>part_employ</i>	Part-time employment, female (% of total female employment)	World Bank
<i>child_enrol</i>	Enrolment in early childhood education	UNESCO Institute for Statistics; Eurostat; UNICEF
<i>parent_leave</i>	Standard Paid Parental Leave Duration	National legal references

(Source: Author's presentation)

This differs from the longer behavioral lags (2 to 5 years) discussed in the demographic literature, which are acknowledged conceptually but cannot be empirically identified within the short time dimension. The results indicate that most variables are  $I(1)$ , while a few are  $I(2)$  or  $I(0)$ . To ensure consistency and avoid spurious regression, all variables were used in their first difference or appropriately transformed to achieve stationarity (Table 2).

Table 2: Results of the panel unit root tests (p-values)

Variable	CIPS			Truncated CIPS*		
	$I(0)$	$I(1)$	$I(2)$	$I(0)$	$I(1)$	$I(2)$
<i>birth_rate</i>	$\geq 0.10$	$< 0.01^{***}$	-	$\geq 0.10$	$< 0.01^{***}$	-
<i>labor_part</i>	$\geq 0.10$	$< 0.01^{***}$	-	$\geq 0.10$	$< 0.01^{***}$	-
<i>GDP</i>	$\geq 0.10$	$< 0.01^{***}$	-	$\geq 0.10$	$< 0.01^{***}$	-
<i>urban</i>	$\geq 0.10$	$< 0.01^{***}$	-	$\geq 0.10$	$\geq 0.10$	-
<i>tertiary</i>	$\geq 0.10$	$\geq 0.10$	$\geq 0.10$	$\geq 0.10$	$\geq 0.10$	$< 0.01^{***}$
<i>part_employ</i>	$< 0.05^{**}$	$< 0.05^{**}$	-	$< 0.05^{**}$	$< 0.05^{**}$	-
<i>child_enrol</i>	$\geq 0.10$	$\geq 0.10$	$< 0.01^{***}$	$\geq 0.10$	$\geq 0.10$	$< 0.01^{***}$
<i>parent_leave</i>	time invariant					

\* Truncated CIPS modifies the CIPS statistics by excluding extreme values from the averaging process. \*\*\* Significant at 0.01; \*\* Significant at 0.05

(Source: Authors' calculations)

The panel unit root test results using both CIPS and truncated CIPS indicate that most variables are stationary in their first difference, while *tertiary* and *child\_enrol* are stationary in their second difference, and *child\_enrol* is stationary in level. The models are estimated with the variable *child\_enrol* in their first difference. Mixing variables stationary in level and stationary in its first difference in a panel regression without cointegration can result in spurious

regression, where relationships appear statistically significant due to shared trends rather than true economic association. To maintain consistent and reliable estimates, all variables should be of the same integration order, typically by differencing stationary variables so they match the other series in their first difference form (Baltagi, 2008; Greene, 2012). First difference transformations focus on short-term changes by removing country-specific effects and long-term trends, capturing only transitory responses to socioeconomic and policy shifts (Wooldridge, 2008; Baltagi, 2008). This approach is well-suited to regions like the Western Balkans and Eastern Europe, where rapid institutional and demographic changes have occurred.

Results from the estimated panel regression model with fixed effects for cross section are presented in Table 3. Fixed effects are used because they control unobserved, time-invariant country-specific factors that may be correlated with the explanatory variables, allowing for consistent estimation of within-country effects over time.

The panel regression model is presented in the following specification:

$$\begin{aligned} \Delta birth\_rate_{it} = & \beta_0 + \beta_1 \Delta labor\_part_{it} + \beta_2 parent\_leave_{it} \\ & + \beta_3 (parent\_leave_{it} * \Delta labor\_part_{it}) + \beta_4 \Delta gdp_{it} + \beta_5 \Delta^2 tertiary_{it} \\ & + \beta_6 \Delta urban_{it} + \beta_7 \Delta^2 child\_enrol_{it} + \beta_8 \Delta part\_employ_{it} + \varepsilon_{it} \end{aligned}$$

where  $i$  represents the country,  $t$  is for time (year) and the error terms which include unobserved shocks or measurement error, is represented with  $\varepsilon_{it}$ .

Table 3: Results from the estimated panel regression models

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
constant	0.40	0.02	-0.02	-0.02	-0.13***
d(labor_part)	-0.25*	-0.24**	-0.22	-0.22	-0.23
parent_leave	-0.01	-	-	-	-
interaction	0.00	0.00	0.00	0.00	0.00
d(gdp)	0.00**	0.00**	0.00*	0.00*	0.00*
d(tertiary, 2)	-0.29**	-0.30**	-0.24	-0.24	-0.27*
d(urban)	-0.33	-0.30	-0.33	-0.33	-
d(child_enrol, 2)	-0.04	-0.04	-0.07	-	-
d(part_employ)	0.00	0.00	-	-	-
Adjusted R <sup>2</sup>	0.05	0.06	0.02	0.03	0.02
F-statistic	1.49	1.60*	1.34	1.39	1.40
Akaike info criterion	0.75	0.73	1.14	1.13	1.13
Schwarz criterion	1.08	1.04	1.39	1.36	1.34
Durbin-Watson stat.	1.74	1.74	1.68	1.69	1.67

\*\*\* Significant at 0.01; \*\*Significant at 0.05

(Source: Authors' calculations)

This equation represents a first difference fixed effects panel regression where the first differences capture the short-term changes. The model tends to investigate the short run determinants of birth rate fluctuations, with accent on the role of female labor force participation. The interaction term allows for testing whether policy moderates the impact of labor market behavior on fertility decisions. The inclusion of economic, educational, and demographic control variables ensures robustness of the estimated relationships.

Table 3 shows that higher female labor participation consistently reduces birth rates in Models 1 and 2, indicating a short-term trade-off between work and childbearing. The effect ( $\approx -0.22$  to  $-0.25$ ) becomes insignificant once additional controls are included, implying mediation through education and economic factors. Tertiary enrolment (two-year lag) remains negative and significant, supporting the postponement hypothesis. GDP growth is positive and highly significant in all models, confirming that economic improvement boosts fertility in the short run. Urbanization, parental leave, early child enrolment, and part-time employment show no significant effects, suggesting that their short-term influence on fertility is limited or delayed. Model diagnostics show acceptable fit: the best-performing model (Model 2) has the lowest AIC and significant  $F$  – statistic. Durbin-Watson statistics suggest no serious autocorrelation. Residual normality is confirmed in Model 2 ( $p = 0.12$ ).

The negative association between female labor force participation and fertility is statistically significant in Models 1 and 2 but becomes insignificant in the expanded specifications (Models 3 to 5). This attenuation reflects the inclusion of policy and structural variables, particularly education and parental leave indicators, that absorb part of the variation previously captured by labor participation. In other words, once education and policy effects are accounted for, female labor force participation no longer exerts an independent short-term effect on fertility, suggesting that its influence operates indirectly through these channels rather than directly on birth rates.

These findings confirm that short-term changes in birth rates are influenced primarily by economic and educational variables, while institutional and structural factors like childcare or leave policies may require longer time horizons or more granular data to show significance.

## **5. DISCUSSION**

The empirical findings highlight the complexity of fertility determinants in Southeastern Europe. These countries have undergone major socioeconomic transitions, EU integration processes, and institutional reforms. All these events have contributed to changes in reproductive behavior and demographic structures.

Female labor force participation has a negative and statistically significant effect on the birth rates. This finding is consistent with previous studies that confirm that higher female employment – particularly when not accompanied by supportive work–family policies – can suppress birth rates, especially in post – socialist and Southern European contexts where traditional gender roles persist (Matysiak and Vignoli, 2008; Adserà, 2004). All the analyzed countries exhibit a constant rising trend in female employment, reaching 50 – 70% participation, especially with post-EU accession and economic reforms. In the absence of sufficient institutional support, increased employment among women may lead to postponement or reduction of childbearing.

The effect of tertiary education is similarly negative and statistically significant, indicating that as women increasingly pursue higher education, the timing of family formation is delayed. This finding reflects the well-documented postponement of childbirth among highly educated women who delay family formation due to the prolonged education and career prioritization (Kögel, 2004; Billingsley, 2010). In countries like Bulgaria and Romania, education reforms under the Bologna Process (from 2005 onward) and EU structural investments (improved university infrastructure and capacity) have expanded access to higher education for women and the rural population, reinforcing this trend.

Regarding other countries, Albania, Bosnia and Herzegovina, North Macedonia, Montenegro, and Serbia, they all have also adopted the Bologna Process in the period between 2003 to 2005, which caused the restructuring of the Bachelor, Master's and PhD cycles and promoted student mobility and quality assurance. Massification of higher education was evident in the mid-2000s, with a rapid increase in tertiary enrollment, especially among young women. According to the UNESCO Institute for Statistics (2024), female enrollment in higher education began to outpace male enrollment during the late 2000s. A combination of prolonged education and uncertain employment outcomes encouraged postponement of marriage and childbirth, in urban areas and among educated women (Billingsley, 2020; Kreyenfeld *et al.*, 2012). In Serbia and Albania, there was a decline or stagnation in the total fertility rate – around 1.3 to 1.5 children per woman, during the same period, which is in correlation with the increase in the portion of women in tertiary education.

Greece began expanding access to university education in the 1990s, and the reforms followed under the Bologna Process post-1999. By the early 2010s, women in Greece were more likely to hold a tertiary degree than men, contributing to the notion that higher female education is associated with delayed fertility.

These trends are consistent with a substantial body of empirical research linking higher female educational attainment to lower birth rates. Education is associated with delayed childbearing, reduced desired family size, and increased opportunity costs of motherhood (Martin, 1995; Matysiak and Vignoli, 2008). In the context of post-socialist societies, rising female enrolment has also contributed to changing gender roles and reproductive norms (Goldstein *et al.*, 2013; Basten *et al.*, 2014). Furthermore, evidence from low-fertility European countries suggests that increases in tertiary education are among the key drivers of long-term fertility decline (Billari and Kohler, 2004).

GDP per capita emerged as a robust positive determinant of fertility, confirming the procyclical nature of fertility behavior. During periods of economic stability or growth, households may feel more secure in making reproductive decisions. This pattern supports earlier findings from Western Europe and is particularly relevant in post transition economies where economic fluctuations strongly influence life planning (Sobotka *et al.*, 2011; Goldstein *et al.*, 2013).

In contrast, variables such as urbanization, early childhood education, and part-time employment were not statistically significant. While these are often considered fertility-enhancing factors in Western Europe, their limited significance in this sample may reflect differences in policy effectiveness, service coverage, or cultural context. For example, early childhood education systems in Southeastern Europe may not yet be developed enough to significantly reduce childrearing burdens or influence fertility behavior (Scherer *et al.*, 2023). Parental leave policies, both in direct form and as moderators, were also not statistically significant. This may stem from the uniformity of parental leave duration across the countries or from limited policy utilization. Furthermore, the use of first difference transformation may obscure long-term or lagged effects of such policies (Gauthier, 2007). Future research could explore whether different dimensions of leave policy — such as flexibility or wage replacement, or uptake rates— have more substantial demographic effects.

Female part-time employment does not seem to be a relevant determinant of the declining birth rate (Fluchtmann *et al.*, 2023). In Eastern Europe, part-time employment has a negligible effect on the birth rate due to its limited use, low quality, and minimal role in work-family balance.

The birth rate is more strongly influenced by structural factors like emigration or economic insecurity.

The change in significance across models thus reflects increasing model completeness, not loss of robustness; the relationship between female employment and fertility remains negative in direction, but its explanatory power diminishes when policy and structural determinants are included.

These findings reflect broader demographic and institutional transitions in Southeastern Europe. The rise of dual earner households, mass tertiary education, and delayed family formation are shaping fertility patterns. Structural challenges such as emigration, aging populations, and insufficient family policy infrastructure continue to exert downward pressure on birth rates. Consequently, while female labor force participation and education remain key variables, their effects on birth rates depend on broader socioeconomic and institutional environments.

## **6. LIMITATIONS**

The crude birth rate was used as a pragmatic proxy for fertility, given incomplete and non-harmonized annual data on age-standardized measures such as the total fertility rate. While this indicator allows consistent longitudinal modelling, it may reflect structural effects of age composition and migration; thus, results should be interpreted as representing broad fertility trends rather than parity-specific behavior.

A further limitation is the omission of fertility ideals and intentions, which influence childbearing outcomes. European studies show a persistent two-child ideal (Sobotka and Beaujouan, 2014) but a widening gap between intended and achieved births amid economic and partnership instability. Education is positively linked to intentions, though not always to realized fertility (Testa, 2014), and underachievement remains common in low fertility settings (Beaujouan and Berghammer, 2019). These dynamics likely evolved in Southeastern Europe during 2000–2023 but cannot be directly modelled due to missing comparable data, implying potential omitted variable bias.

## **7. CONCLUSION**

This study examined short-term determinants of birth rates in eight Southeastern European countries. The analysis focused on female labor force participation as the primary explanatory variable and included economic, educational, and demographic controls. The findings reveal that increases in female employment and tertiary education are associated with declining birth rates, while economic growth exerts a positive influence on fertility. These relationships reflect both individual-level trade-offs between work and family and macro-level structural conditions.

Notably, parental leave and other family policy variables did not show significant short-term effects in this sample, which may reflect limitations in policy implementation, coverage, or delayed behavioral responses. Similarly, the lack of significance for part-time employment and early childhood education suggests that such measures may not yet be widespread or effective enough in the region to impact fertility decisions.

The results underscore the challenges of reconciling female employment with family formation in post socialist contexts. Policymakers in Southeastern Europe should consider comprehensive work – family policies, including flexible parental leave, affordable childcare, and employment

protections, to mitigate the apparent tradeoff between labor market participation and fertility. Future research should examine long-term effects, policy heterogeneity, and qualitative dimensions of gender roles and family norms to provide a deeper understanding of fertility behavior in transitioning societies.

## REFERENCES

- Adserà, A. (2004), “Changing fertility rates in developed countries: The impact of labor market institutions”, *Journal of Population Economics*, Vol. 17 No. 1, pp. 17-43.
- Aydoğan, E. T. and Uslu, Ç. L. (2019), “The Effects of Female Tertiary Education on Fertility and Development”. *Contemporary Research in Economics and Social Sciences*, 3(1), 153-196.
- Arjan, G., Arnstein, A. and Letizia, M. (2009), “The highest fertility in Europe—for how long? Determinants of fertility change in Albania”, *Demográfia*, Vol. 52 No. 5, pp.76-96.
- Baltagi, B.H. (2008), *Econometric Analysis of Panel Data* 4<sup>th</sup> ed., John Wiley & Sons, Chichester.
- Basten, S., Sobotka, T. and Zeman, K. (2014), “Future fertility in low fertility countries”, in Lutz., W., Butz, W.P. and KC, S.. (Eds.) *World Population and Human Capital in the Twenty-First Century*, Oxford Academic, Oxford, pp. 39-146.
- Beaujouan, E. and Berghammer, C. (2019), “The gap between lifetime fertility intentions and completed fertility in Europe and the United States: A cohort approach”, *Population Research and Policy Review*, Vol. 38 No. 4, pp. 507-535.
- Bharati, T., Chang, S. and Li, Q. (2021), “The effect of tertiary education expansion on fertility: a note on identification”, IZA Discussion Papers, No. 14672, Institute of Labor Economics (IZA), Bonn.
- Billari, F. and Kohler, H.P. (2004), “Patterns of low and lowest-low fertility in Europe”, *Population Studies*, Vol. 58 No. 2, pp. 161-176.
- Billingsley, S. (2010), “The post-communist fertility puzzle”, *Population Research and Policy Review*, Vol. 29 No. 2, pp. 193–231.
- Bloom, D., Canning, D. and Fink, G. (2009), “Fertility, female labor force participation, and the demographic dividend”, *Journal of Economic Growth*, Vol. 14, pp. 79-101.
- Chesnais, J.C. (1992), *The demographic transition: stages, patterns, and economic implications: a longitudinal study of sixty-seven countries covering the period 1720-1984*, Oxford University Press.
- Eurostat (2015), *Enlargement countries – Population and social conditions* (short guide). Luxembourg: Publications Office of the European Union; see note: “2010: break in series” indicating methodological change in 2010.
- Eurostat (2023), *Enrolment in early childhood education by age and sex — % of the population of the corresponding age group (educ\_uoe\_enra10)*.
- Eurostat (2023), *Fertility statistics*. Luxembourg: Publications Office of the European Union.
- Eurostat (2023), *Population having completed tertiary education by age and sex (edat\_lfse\_03)*.
- Fluchtmann, J., van Veen, V. and Adema, W. (2023), *Fertility, employment and family policy: A cross-country panel analysis*, Social, Employment and Migration Working Papers, No. 299, OECD.
- Frejka, T., and Gietel-Basten, S. (2016), “Fertility and Family Policies in Central and Eastern Europe after 1990”. *Comparative Population Studies*, Vol. 41 No. 1, pp. 3-56.
- Gauthier, A.H. (2007), “The impact of family policies on fertility in industrialized countries: A review of the literature”, *Population Research and Policy Review*, Vol. 26 No. 3, pp. 323–346.

- Greene, W. (2012), *Econometric Analysis* 7<sup>th</sup> ed., Prentice Hall, Upper Saddle River, NJ.
- Goldstein, J.R., Kreyenfeld, M., Jasilioniene, A. and Örsal, D.K. (2013), “Fertility reactions to the “Great Recession” in Europe: Recent evidence from order-specific data”, *Demographic Research*, Vol. 29, pp. 85-104.
- Kögel, T. (2004), “Did the association between fertility and female employment within OECD countries really change its sign?”, *Journal of Population Economics*, Vol. 17 No. 1, pp. 45-65.
- Kreyenfeld, M., Andersson, G. and Pailhé, A. (2012), “Economic uncertainty and family dynamics in Europe”, *Demographic Research*, 27, pp. 835-852.
- Luci-Greulich, A. and Thévenon, O. (2013), “The Impact of Family Policies on Fertility Trends in Developed Countries: L’influence des politiques familiales sur les tendances de la fécondité des pays développés”, *European Journal of Population/Revue Européenne de Démographie*, Vol. 29 No. 4, pp. 387-416.
- Luo, Y. and Jarosz, E. (2024), “Job creation, job destruction, and fertility in Germany”, *Demographic Research*, Vol. 52 No. 13, pp. 383-414.
- Magda, I., A. Kiełczewska and N. Brandt (2018), “The “family 500+” child allowance and female labour supply in Poland”, *OECD Economics Department Working Papers*, No. 1481, OECD Publishing, Paris.
- Martin, T.C. (1995), “Women's education and fertility: results from 26 Demographic and Health Surveys” *Studies in Family Planning*, Vol 26 No. 4, pp.187-202.
- Matysiak, A. and Vignoli, D. (2008), “Fertility and women’s employment: a meta-analysis: fécondité et travail des femmes: une méta-analyse”. *European Journal of Population/Revue Européenne de Démographie*, Vol. 24 No. 4, pp. 363-384.
- Mishra, V. and Smyth, R. (2010), “Female labor force participation and total fertility rates in the OECD: New evidence from panel cointegration and Granger causality testing”, *Journal of Economics and Business*, Vol. 62 No. 1, pp. 48-64.
- MONSTAT (Statistical Office of Montenegro) (2010), *Statistical Yearbook 2010*, table 4-10: Birth rates, death rates and marriages. Statistical Office of Montenegro, Podgorica.
- Nikitović, V. and Magdalenić, I. (2024), “Gender differences in cohort fertility patterns in Serbia: the role of educational gradient”, *Stanovništvo*, Vol. 62 No. S1, S5-S35.
- NSI (National Statistical Institute of Bulgaria) (2012), *Demographic Processes 2010*. NSI, Sofia.
- Pesaran, M.H. (2007), “A simple panel unit root test in the presence of cross-section dependence”, *Journal of Applied Econometrics*, Vol. 22 No. 2, pp. 265-312.
- Philipov, D. and Jasilioniene, A. (2008), “Union formation and fertility in Bulgaria and Russia: A life table description of recent trends”, *European Journal of Population*, Vol. 24 No. 1, pp. 91–117.
- Slonimczyk, F. and Yurko, A. (2014), “Assessing the impact of the maternity capital policy in Russia”, *Labour Economics*, Vo. 30, pp. 265-281.
- Spéder, Z., Murinkó, L. and Oláh, L.S. (2020), “Cash support vs. tax incentives: The differential impact of policy interventions on third births in contemporary Hungary”, *Population Studies*, Vol. 74 No. 1, pp. 39-54.
- Scherer, S., Pavolini, E. and Brini, E. (2023), “Formal childcare services and fertility: the case of Italy” *Genus*, Vol. 79 Article 29.
- Shirahase, S. (2000), “Women's increased higher education and the declining fertility rate in Japan”, *Review of Population and Social Policy*, Vol. 9, pp. 47-63.
- Sobotka, T. and Beaujouan, É. (2014), “Two is best? The persistence of a two-child family ideal in Europe”, *Population and Development Review*, Vol. 40 No. 3, pp. 391-419.
- Sobotka, T., Skirbekk, V. and Philipov, D. (2011), “Economic recession and fertility in the developed world”, *Population and Development Review*, Vol. 37 No. 2, pp. 267-306.

- Testa, M.R. (2014), “On the positive correlation between education and fertility intentions in Europe: Individual-and country-level evidence”, *Advances in Life Course Research*, Vol. 21, pp. 28-42.
- UNESCO Institute for Statistics (UIS) (2023), *UIS Data Centre*. Montreal: UNESCO Institute for Statistics.
- UNICEF (2023), *Multiple Indicator Cluster Surveys (MICS)*. New York: United Nations Children’s Fund.
- Vasireddy, S., Berrington, A., Kuang, B. and Kulu, H. (2023), “Education and fertility in Europe in the last decade: A review of literature”, *Comparative Population Studies*, Vol. 48, pp. 553-588.
- World Bank (2025), *World Development Indicators*. Washington, DC: World Bank.
- Wooldridge, J.M. (2008), *Econometric analysis of cross section and panel data*. 2<sup>nd</sup> ed., MIT Press, Cambridge, MA.
- Čipin, I., Zeman, K. and Međimurec, P. (2020), “Cohort fertility, parity progression, and family size in former Yugoslav countries during the twentieth century”, *Comparative Population Studies*, Vol. 45, pp. 229-264.

## ENGINES AND CONSTRAINTS OF GROWTH IN THE WESTERN BALKANS: INSIGHTS FOR EU CONVERGENCE

**Marija Trpkova-Nestorovska**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*Marija.Trpkova-Nestorovska@eccf.ukim.edu.mk*

**Katarina Gacova**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*katarina.gacova@gmail.com*

### **ABSTRACT**

*This paper investigates the determinants of GDP growth in nine Western Balkan countries over the period 2000 – 2023 using an unbalanced panel dataset. The analysis incorporates gross fixed capital formation, exports and imports of goods and services, final and household consumption, labor force growth, and inflation as explanatory variables. To address cross-sectional dependence and unobserved heterogeneity, the study employs a three-model econometric framework, including two-way fixed effects, correlated random effects, and random effects specifications. The results demonstrate that investment, exports, and final consumption are the most robust and statistically significant drivers of GDP growth, while imports consistently exert a negative effect. Labor force growth is positively associated with output but only becomes significant under more robust specifications, whereas inflation shows no systematic impact. The findings highlight the dual role of external competitiveness and domestic demand in sustaining growth, while underscoring structural vulnerabilities linked to import dependence and weak labor market absorption.*

**Keywords:** *GDP growth, Western Balkans, Panel regression, Investment, Trade, Consumption*

**JEL classification:** *C23, E21, E22, F43, O52*

### **1. INTRODUCTION**

Economic growth in the Western Balkans and South-East Europe remains shaped by structural legacies, external dependence, and the ongoing process of EU integration. Despite progress since the early 2000s, the region continues to face challenges related to investment gaps, trade imbalances, and weak labor market absorption, leaving growth trajectories vulnerable to global shocks. Previous research highlights gross fixed capital formation, exports, and consumption as the main engines of growth in small open economies, while labor force dynamics and inflation exert more uncertain effects. In the Western Balkans, capital accumulation is essential for upgrading productive capacity, exports drive convergence through external competitiveness, and domestic consumption sustains short-run growth, whereas high import dependence often offsets these gains. However, empirical studies that systematically evaluate these determinants for the Western Balkans remain relatively scarce compared to EU-wide analyses, leaving important gaps in understanding the drivers of regional growth. This paper investigates the relative importance of these factors for GDP growth in nine countries over 2000 - 2023 using an unbalanced panel and alternative econometric specifications. By addressing cross-sectional dependence and unobserved heterogeneity, the study provides

robust evidence on the dual role of external competitiveness and domestic demand in driving growth, offering insights for both academic debate and regional policy strategies.

## **2. LITERATURE REVIEW**

Economic growth in the Western Balkans and South-East Europe has been shaped by a distinctive combination of structural legacies, external dependencies, and domestic policy frameworks. Unlike advanced European economies, the region continues to grapple with investment gaps, fragile external balances, and labor-market frictions, all of which condition its growth dynamics. Understanding the determinants of GDP in this context is essential, not only for explaining past performance but also for assessing future convergence prospects with the European Union.

The academic and policy literature highlights several key drivers of growth in the region. Gross fixed capital formation emerges as a central determinant, reflecting both the volume and composition of investment. At the same time, trade integration, particularly the role of exports, has supported convergence, though often moderated by the technological content and import dependence of regional economies. Final consumption, driven by rising incomes and remittances, has played an outsized role in sustaining short-run growth, while labor market participation improvements have been offset by persistent productivity challenges. Broader evidence from South-East Europe also underscores the importance of institutional quality and macroeconomic fundamentals in sustaining growth trajectories.

Against this backdrop, the following section reviews the existing literature on GDP determinants in the Western Balkans and South-East Europe, situating investment, trade, consumption, and labor force dynamics within regional perspective. This provides a foundation for the empirical analysis that follows and frames the relevance of the chosen regressors for understanding growth processes in transition and convergence economies.

A substantial strand of South-East Europe work links investment to growth through both the level and composition of gross fixed capital formation. The European Investment Bank documents a persistent investment gap and a low share of infrastructure within gross fixed capital formation across the Western Balkans, implying weaker long-run growth multipliers relative to EU peers (European Investment Bank, 2017). Infrastructure investment in the Western Balkans is constrained by limited financial conditions and rising external debt, despite significant needs in sectors beyond traditional energy and transport (Holzner and Grieveson, 2018). Recent OECD benchmarking confirms that private investment has hovered around EU averages since 2020 but remains uneven across economies (OECD, 2025). IMF analysis adds that imports of capital goods raise productivity beyond their mechanical effect on investment—highlighting the role of external sourcing of machinery and technology for small open Western Balkan economies (IMF, 2019).

Turning to trade, multiple studies identify exports as a core engine of growth. The IMF reports that Western Balkan export growth outpaced world GDP growth over the 2000s, consistent with a trade-led convergence pattern (IMF, 2019). More recent econometric evidence finds both exports and imports positively associated with GDP in the WB6 over 2010 - 2021, interpreting the import effect as technology diffusion and input complementarities in small open economies (Sejfijaj *et al.*, 2025).

On final consumption, the region appears unusually consumption-driven: panel evidence and narrative assessments show private consumption as a primary short-run driver of GDP, with employment and purchasing-power gains reinforcing this channel (Hashani, 2022). World Bank regional reports echo the same composition - robust consumption frequently underpins near-term growth, while investment and exports determine medium-term prospects (World Bank, 2023-2025).

Regarding the labor force, Western Balkans labor markets have improved in participation and employment, but productivity remains low - so employment gains translate into growth, yet convergence hinges on productivity-enhancing investment and skills (WIIW, 2023; OECD, 2025).

Finally, broader South-East Europe panels (not limited to regressors used in this paper) corroborate that macro fundamentals, institutions, and openness matter for growth - a useful context for your specification (Trpkova and Tashevska, 2011).

Growth is strongly tied to the quality and composition of investment (especially capital-imports and infrastructure), an export orientation, and a consumption-heavy demand mix. Labor market improvements help in the short run, but productivity upgrades via investment and skills are decisive for sustained growth.

### **3. DATA, METHODOLOGY, AND EMPIRICAL RESULTS**

The unbalanced panel dataset (due to some missing data for Montenegro) consists of nine countries from the Western Balkans and the region – Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Greece, Montenegro, North Macedonia, Serbia and Slovenia. Data are collected for the period 2000-2023, which makes the sample equal to 216 observations. The model employs the following eight variables presented in Table 1.

*Table 1: Variables of the dataset*

<b>Variable</b>	<b>Description</b>	<b>Source</b>
<i>GDP</i>	Gross domestic product presented as annual %	World Bank
<i>cap_form</i>	Gross fixed capital formation (annual % growth)	World Bank
<i>exp</i>	Exports of goods and services (annual % growth)	World Bank
<i>fin_cons</i>	Final consumption expenditure (annual % growth)	World Bank
<i>hous_cons</i>	Household and non-profit institutions serving households final consumption expenditure (annual % growth)	World Bank
<i>imp</i>	Imports of goods and services (annual % growth)	World Bank
<i>inf</i>	Inflation, GDP deflator (annual %)	World Bank
<i>labor</i>	Labor force, total (annual % growth)	World Bank

*(Source: Authors' presentation)*

The aim of this study is to examine the potential effect that the group of macroeconomic variables have on the gross domestic product. In order to achieve this, the variables were tested for cross sectional dependence and unit roots to determine their stationarity. The cross-sectional dependence tests and their results are presented in Table 2.

The null hypothesis of cross-section dependence test states that there is no cross-sectional dependence or correlation between the countries in the data set. From the presented results it is noticeable that in all variables cross sectional dependence is present, except for the variable

“Labor”. In further testing, different unit root tests will be applied only for this variable. Cross-sectional dependence is confirmed in the Balkans because the countries are economically interlinked, exposed to the same external shocks, and structurally similar transition economies. Their business cycles and macro variables move together, making independence an unrealistic assumption.

Table 2: Cross-section dependence tests

Variable	Test statistic				
	Breusch-Pagan LM	Pesaran scaled LM	Bias-corrected scaled LM	Pesaran CD	CD present
<i>GDP</i>	495.14***	54.11***	53.91***	21.96***	Yes
<i>cap_form</i>	172.10***	16.04***	15.84***	11.67***	Yes
<i>exp</i>	318.84***	33.33***	33.14***	16.79***	Yes
<i>fin_cons</i>	353.62***	37.43***	37.24***	18.57***	Yes
<i>hous_cons</i>	347.84***	36.75***	36.56***	18.35***	Yes
<i>imp</i>	323.20***	33.85***	33.65***	17.38***	Yes
<i>inf</i>	320.05***	33.48***	33.28***	17.13***	Yes
<i>labor</i>	43.36	0.87	0.67	3.07	No

(LM - Lagrange Multiplier; CD - Cross-sectional Dependence)

(\*, \*\*, \*\*\*, statistically significant at 0.1, 0.5 and 0.01)

(Source: Authors' presentation)

For unit root testing of the variables that present cross sectional dependency second-generation unit root tests are used. The specification is previously established according to the development trends of the variables. Results are presented in Table 3.

Table 3: Stationarity tests

Variable	Specification	Test statistic	
		CIPS	Truncated CIPS
<i>GDP</i>	None	-2.60***	-2.60***
<i>cap_form</i>	None	-3.42***	-3.42***
<i>exp</i>	None	-3.28***	-3.28***
<i>fin_cons</i>	Constant	-3.03***	-3.03***
<i>hous_cons</i>	None	-3.01***	-3.01***
<i>imp</i>	None	-3.45***	-3.45***
<i>inf</i>	Constant	-4.92***	-4.59***

Variable	Specification	Test statistic		
		Levin, Lin and Chu	ADF Fisher $\chi^2$	PP Fisher $\chi^2$
<i>labor</i>	None	-5.96***	96.45***	144.09***

(ADF – Augmented Dickey–Fuller; PP - Phillips–Perron)

(\*, \*\*, \*\*\*, statistically significant at 0.1, 0.5 and 0.01)

(Source: Authors' presentation)

The Pesaran CIPS test (Pesaran, 2007) was applied to determine stationarity in variables with confirmed cross-sectional dependency. Maximum lag of 1 with AIC-based selection was used due to the limited number of observations. The results indicate that all variables are stationary. Regarding the variable *labor*, where cross-sectional dependency was not confirmed, classical panel unit root tests were used. Stationarity was also present.

With stationarity confirmed with all variables, the next step is the estimation of panel regression. For the Balkan macro unbalanced panel, the appropriate specification is fixed effects rather than random effects. The main reason is that the countries differ systematically in structural characteristics such as institutions, geography, EU accession status, and governance quality, which are very likely correlated with the explanatory variables. This correlation makes the random effects estimator inconsistent, while fixed effects account for such unobserved heterogeneity. In addition, the region has been subject to major common shocks (2008 global financial crisis, Eurozone debt crisis, COVID-19, energy shocks), which justifies the inclusion of year fixed effects to absorb period-specific influences. Despite the fixed effect being the most likely outcome, Hausman test has  $\chi^2$  statistic 1.73, thus accepting the null hypothesis of estimation of panel model with random effects.

Although the Hausman test points to random effects, a two-way fixed-effects model is the main specification because it is expected that country effects to be correlated with the regressors (e.g., institutions, EU-accession path, geography). Under this realistic correlation — and given cross-sectional dependence and an unbalanced panel — the Hausman result can be unreliable. Country fixed effects purge time-invariant heterogeneity, year fixed effects absorb common shocks, and cluster-by-country standard errors deliver valid inference under heteroskedasticity and within-country serial correlation (and are more resilient to cross-sectional dependency). Practically, fixed effects are also estimable and stable in these data, whereas random effects face degrees-of-freedom constraints. Therefore, random effects are treated as a robustness check, and the two-way fixed effects model is the baseline model.

A three-model approach is used in the analysis to triangulate results under competing assumptions about country effects. Model A (two-way fixed effects, White period (cross-section cluster)) is the baseline because it is expected that country effects to correlate with regressors and because fixed effects are stable/estimable on this unbalanced, cross-section dependent-prone panel. It identifies within-country effects while absorbing common shocks.

Model B (lean correlated random effects specification (Mundlak, 1978; Chamberlain, 1982) is designed to run reliably on small and unbalanced panels (Wooldridge, 2010; 2019). It relaxes pure-fixed effects by allowing correlated random effects via country means, giving the opportunity to check whether results persist when limited between-country information is retained (and when time-invariant factors matter).

Model C (random effects for cross-section, none for period, White cross-section (period cluster)) serves as the formal Hausman counterfactual—efficient if the country effects were exogenous—and provides a transparent appendix comparison.

Agreement in sign and magnitude across the three strengthens credibility. If any divergences appear, they reveal sensitivity to the fixed effects versus random effects assumption rather than the data per se.

Results from the three estimated models are presented in Table 4.

According to the results, credibility can be confirmed, since the size of the estimated coefficients, their signs and significance are similar throughout the models. The only difference is found in `hous_cons` and `labor`. While Models B and C confirm the coefficients are statistically significant, the first model does not.

Table 4: Results from the estimated panel regression models

Variable	Model A	Model B	Model C
constant	0.66**	2.76	0.15
cap_form	0.07**	0.08***	0.08***
exp	0.16***	0.20***	0.19***
fin_cons	0.37**	0.42**	0.42**
hous_cons	0.14	0.22*	0.22*
imp	-0.12***	-0.14**	-0.14**
inf	0.02	0.02	0.01
labor	0.12	0.18***	0.16**
Adjusted R <sup>2</sup>	0.86	0.84	0.85
F-statistic	33.68	155.35	157.02
Akaike info criterion	3.73	n/a	n/a
Schwarz criterion	4.36	n/a	n/a
Durbin-Watson stat.	1.81	1.98	1.95

(\*, \*\*, \*\*\*, statistically significant at 0.1, 0.5 and 0.01)

(Source: Authors' presentation)

Core drivers of the GDP growth are gross fixed capital formation, exports of goods and services, final consumption expenditure, household final consumption expenditure and labor force. The first variable, *cap\_form*, has positive and statistically significant effect across all models (0.07–0.09). This confirms that investment in physical capital (infrastructure, machinery) is a consistent driver of GDP growth in the region. The second variable, *exp* also marks as strongly positive and highly significant (0.16–0.20). Export growth is one of the most robust predictors of GDP performance, reflecting the openness and dependence of these economies on external demand. Final consumption expenditure has a positive and significant effect on GDP (0.37–0.42). Domestic demand, particularly public + private consumption, plays a large role in sustaining GDP. Variable *hous\_cons* is statistically significant in Models B and C (0.22). This indicates that when the household demand is isolated from general consumption, it remains a relevant driver of growth. The effect of labor force growth on GDP is positive across all models but becomes statistically significant only under Models B and C, suggesting that more robust estimators capture the contribution of labor supply more effectively. This indicates that labor force dynamics do play a role in growth, though the strength of this relationship is sensitive to model specification.

Imports of goods and services remain to have negative and statistically significant effects (-0.12 to -0.14). This is expected, as imports represent a leakage from GDP when not matched by productivity gains or export increases.

Inflation has positive but not significant effect, meaning inflationary dynamics (measured by GDP deflator) do not systematically explain GDP growth variation in this sample.

Goodness of fit and diagnostics report good models. The adjusted R<sup>2</sup> is high across all models (0.86–0.88), showing strong explanatory power. The F-statistics are highly significant, confirming the joint validity of regressors. Durbin–Watson statistics are close to 2 (1.8–1.95), suggesting no severe autocorrelation.

Model progression (A → B → C): The three estimated models rely on the same set of explanatory variables but differ in terms of specification and estimation approach. Model A

represents the baseline, while Models B and C introduce alternative treatments of unobserved heterogeneity and covariance structures, which slightly improve model fit and reveal more robust coefficient estimates. Importantly, the core drivers of growth remain stable across specifications. At the same time, labor force growth, while consistently positive, becomes statistically significant in Models B and C, indicating that its contribution to GDP growth is more effectively captured under these more robust frameworks.

#### **4. DISCUSSION OF RESULTS**

The empirical results provide important insights into the drivers of economic growth in the Western Balkans during the 2000–2023 period. Across all model specifications, gross fixed capital formation (Gashi and Sylejmani, 2020; Kaštelan and Konatar, 2022; Nagy and Siljak, 2019), exports of goods and services (Zaimaj, 2024; Trlaković *et al.*, 2018; Krasniqi and Topxhiu, 2017), and final consumption expenditure (Mustafa and Abdullahu, 2024; Bilalli *et al.*, 2023; Hashani *et al.*, 2022) emerge as the most robust and statistically significant determinants of GDP growth. This finding is consistent with the growth literature that highlights the role of investment in physical capital and export performance as key contributors to output expansion in small, open economies. In the Western Balkans, where economic structures are heavily shaped by external trade relations and infrastructure gaps, the strong positive effect of investment and exports confirms their centrality to long-term growth strategies. Similarly, the positive impact of consumption reflects the importance of domestic demand in sustaining growth, especially in economies with relatively small domestic markets and limited high-technology sectors (Halebic, 2016; Hashani *et al.*, 2022; Sulejmani and Thaci, 2021).

Imports of goods and services were found to have a negative and statistically significant effect across all models. This aligns with standard national income accounting, where imports represent a leakage from domestic GDP unless offset by productivity gains or re-exports. In the Balkan context, the persistent negative contribution of imports also reflects structural trade imbalances and dependency on foreign goods, a pattern documented in transition economies with narrow export bases and high import penetration. Similar research can be found in Marjanović *et al.* (2021), Bojat *et al.* (2021), Turan and Karamanaj (2014).

The labor force variable shows a consistently positive effect, becoming statistically significant in Models B and C. This suggests that labor supply expansion does contribute to economic growth, but its effect is only fully captured under more robust estimation frameworks. The result resonates with the broader debate on demographic-economic interactions in post-transition economies: while labor availability is a necessary input for growth, its impact depends on the ability of domestic economies to productively absorb new entrants into the labor market. Given the high levels of emigration and structural unemployment in the Western Balkans, this finding indicates that demographic potential remains underutilized but can become a growth driver when supported by institutional and market conditions.

Inflation, proxied by the GDP deflator, does not show a significant effect on GDP growth. This is not surprising, given that moderate inflation in the region has largely been shaped by external shocks (energy crises, global commodity fluctuations) rather than by endogenous growth dynamics. Similar findings are reported in empirical studies on emerging European economies, where inflation volatility exerts limited explanatory power once structural variables are included.

Importantly, the consistency of signs and magnitudes across all three models strengthens the robustness of the results. The two-way fixed effects model (Model A) remains the preferred specification, as it accounts for unobserved country-specific heterogeneity and common shocks such as the global financial crisis, the Eurozone debt crisis, and the COVID-19 pandemic. The lean correlated random effects model (Model B) and the random effects model (Model C) confirm the core results, while highlighting that labor force effects are sensitive to specification. Such triangulation is in line with best practices in panel econometrics, which stress the importance of robustness checks when dealing with unbalanced panels and cross-sectional dependence (Mundlak, 1978; Chamberlain, 1982; Wooldridge, 2019).

Overall, the findings underscore a dual growth dynamic in the Western Balkans: external competitiveness, captured through export performance, and internal capacity, represented by capital formation and domestic consumption, both act as engines of GDP growth. However, these engines are moderated by structural vulnerabilities, including import dependence and limited labor market absorption capacity. This pattern reflects the broader experience of post-socialist economies integrating into the EU, where openness and domestic demand foster growth but also expose structural weaknesses that require sustained policy attention.

#### **4. CONCLUSION**

GDP growth in the Western Balkans between 2000 and 2023 was primarily driven by capital formation, exports, and final consumption, while imports acted as a drag on growth. Labor force growth showed a consistently positive effect, becoming statistically significant in the more robust model specifications, whereas inflation exhibited no systematic impact. Overall, the findings emphasize the dual role of external competitiveness (exports) and internal demand (consumption and investment) as key engines of growth, tempered by the region's structural dependence on imports.

These results underline both opportunities and vulnerabilities in the region's growth model. On the one hand, sustained investment and export expansion are capable of driving convergence with the European Union, particularly when capital formation is directed toward productivity-enhancing sectors such as infrastructure, manufacturing, and technology. On the other hand, the persistent negative effect of imports reveals structural weaknesses in production and competitiveness, suggesting that without stronger domestic value chains, external openness risks reinforcing dependency rather than fostering balanced growth.

From a policy perspective, three priorities emerge. First, governments should strengthen domestic investment by mobilizing EU funds, improving institutional quality, and reducing barriers to private capital inflows. Second, export competitiveness requires diversification toward higher-value sectors, along with better integration into European supply chains. Third, labor market reforms should focus on enhancing skills, retaining talent, and productively absorbing labor force growth to transform demographic potential into a sustainable growth engine. Addressing these structural gaps would not only mitigate the vulnerabilities revealed in this study but also lay the groundwork for stable, long-term convergence with the EU.

#### **REFERENCES**

Bilalli, A., Beka, A. and Gara, A. (2023), "Financial development and economic growth: Evidence from Western Balkan countries", *InterEULawEast: Journal for International and European Law, Economics and Market Integrations*, Vol. 10 No. 1, pp. 209-222.

- Bojat, M., Kovačević, S. and Kurušić, D. (2021) “Interdependence of foreign trade and the economic growth of Serbia”, *Proceedings of the Faculty of Economics in East Sarajevo/Zbornik Radova Ekonomskog Fakulteta u Istočnom Sarajevu*, Issue 23, pp. 21-42.
- Chamberlain, G. (1982), “Multivariate regression models for panel data”, *Journal of Econometrics*, Vol. 18 No. 1, pp. 5-46.
- European Investment Bank (2017), *Infrastructure Investment in the Western Balkans*. Luxembourg: European Investment Bank.
- Gashi, R. and Sylejmani, L.A. (2020), “Econometric approach of the nexus between remittances and economic growth in Western Balkans”, *International Journal of Economics and Business Administration*, Vol. 8 No. 1, pp. 254-262.
- Halebic, J. (2016), “Consumption function in the Western Balkans countries”, *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, pp. 267-276.
- Hashani, M. (2022), “Private consumption as a driving force of economic growth in Western Balkan countries”, *Journal of Governance and Regulation*, Vol. 11 No. 3, pp. 84-94.
- Hashani, M., Ziberi, B. and Bajrami, R. (2022), “The rule of thumb: Private consumption as a driving force of the economic growth of Western Balkan countries”, *Journal of Governance and Regulation*, Vol. 11 No. 3, pp. 54-60.
- Holzner, M. and Grieveson, R. (2018), *Investment in the Western Balkans: New directions and financial constraints in infrastructure investment*. Policy Notes and Reports No. 27, Vienna Institute for International Economic Studies, Vienna.
- International Monetary Fund (IMF) (2019), *Lifting growth in the Western Balkans*, IMF, Washington, DC.
- Kaštelan, U. and Konatar, M. (2022), “The impact of health capital on economic growth in the Balkan countries”, *ENTRENOVA – ENTERprise REsearch InNOVation*, Vol. 8 No. 1, pp. 99-107.
- Krasniqi, F.X. and Topxhiu, R.M. (2017), “Export and economic growth in the West Balkan countries”, *Romanian Economic Journal*, Vol. 20 No. 65, pp. 88-104.
- Marjanović, D., Beraha, I. and Simović, V. (2021), “The impact of import, export and FDI on the economic growth of the Western Balkans countries”, *Economic Analysis: Journal of Emerging Economics*, Vol. 54 No. 2, pp. 20-29.
- Mundlak, Y. (1978), “On the pooling of time series and cross section data”, *Econometrica*, Vol. 46 No. 1, pp. 69-85.
- Mustafa, B. and Abdullahu, S. (2024), “Public spending and economic growth: Evidence from the Western Balkans”, *International Journal of Applied Economics, Finance and Accounting*, Vol. 18 No. 2, pp. 230-240.
- Nagy, S.G. and Siljak, D. (2019), “Economic convergence of the Western Balkans towards the EU-15”, *Revista Finanzas y Política Económica*, Vol. 11 No. 1, pp. 41-53.
- Organisation for Economic Co-operation and Development (OECD) (2025), *Economic convergence scoreboard for the Western Balkans 2025*, OECD, Paris.
- Pesaran, M.H. (2007), “A simple panel unit root test in the presence of cross-section dependence”, *Journal of Applied Econometrics*, Vol. 22 No. 2, pp. 265-312.
- Sejfičaj, O., Shehu, E., Tahiri, A., Millaku, B., Kovaçi, I. and Selimaj, A. (2025), “The impact of export and import on GDP: The case of Western Balkan countries”, *International Journal of Sustainable Development and Planning*, Vol. 20 No. 2, pp. 807-817.
- Sulejmani, L.A. and Thaci, A. (2021), “The nexus between private consumption and economic growth: Evidence from Western Balkan countries”, *Knowledge – International Journal*, Vol. 46 No. 5, pp. 725-729.

- Trlaković, J., Despotović, D. and Ristić, L. (2018), “Impact of technology-intensive exports on GDP of Western Balkan countries”, *Journal of Policy Modeling*, Vol. 40 No. 5, pp. 1038-1049.
- Trpkova, M. and Tashevska, B. (2011), “Determinants of economic growth in South-East Europe: A panel data approach”, *Perspectives of Innovations, Economics and Business*, Vol. 7 No. 1, pp. 1-4.
- Turan, G. and Karamanaj, B. (2014), “An empirical study on import, export and economic growth in Albania”, *Academic Journal of Interdisciplinary Studies*, Vol. 3 No. 3, p. 428.
- Vienna Institute for International Economic Studies (WIIW) (2023), *Western Balkans Labor Market Brief 2021* (update note). Vienna: WIIW.
- Wooldridge, J.M. (2010), *Econometric Analysis of Cross Section and Panel Data*. 2nd edn. Cambridge, MA: MIT Press.
- Wooldridge, J.M. (2019), “Correlated random effects models with unbalanced panels”, *Journal of Econometrics*, Vol. 211 No. 1, pp. 137–150.
- World Bank (2023–2025), *Western Balkans regular economic report* (various issues), World Bank, Washington, DC.
- World Bank (2025), *World Development Indicators (WDI) database*. Washington, DC: World Bank.
- Zaimaj, E. (2024), “Drivers and performance of exports in WB6 – An empirical investigation of the current dynamics”, *EIRP Proceedings*, Vol. 19 No. 1, pp. 174-180.

*Management Papers*

**A BIBLIOMETRIC INSIGHT TO MACHINE LEARNING APPLICATIONS FOR  
DECISION-MAKING**

**Ivona Serafimovska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
ivona.serafimovska@eccf.ukim.edu.mk*

**Bojan Kitanovikj**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
bojan.kitanovikj@eccf.ukim.edu.mk*

**Filip Peovski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
peovski@eccf.ukim.edu.mk*

**ABSTRACT**

*Using a multi-method bibliometric analysis of published documents from Web of Science and Scopus in the last 34 years, this comprehensive study investigates how machine learning improves advanced decision-making while adhering to the PRISMA guidelines. This study's main goal is to make the methodological patterns, thematic directions, and intellectual structure of research at the nexus of machine learning and decision-making visible. The results show that the U.S., China, India, Germany, and the U.K. are leading a rapidly expanding, cooperative research landscape with a strong emphasis on management, marketing, and finance. Tree-based models, support vector machines, deep learning, reinforcement learning, and explainable artificial intelligence are examples of frequently used algorithms. The field is moving toward applications in big data environments, ethical considerations, and increased interpretability. Digital transformation, competitive intelligence, and strategic planning are highlighted in influential works. This synthesis offers direction for developing more transparent machine learning models and practical frameworks for their use in decision-making, serving both academics and practitioners.*

**Keywords:** *Bibliometric analysis, Decision-making, Machine learning*

**JEL classification:** *B41, C55*

**1. INTRODUCTION**

The adaptability of organizations in the complex and dynamic economic environment is mostly influenced by the scope and quality of the decisions being made. While decisions are mostly intuitive among economic agents, it is preferable to base them on data and information, rather than experience and pure common sense. The decision-making process intertwines multiple factors that span across various goals, environments, policies, and even human behavior. Moreover, the exponential growth of data and information implies the necessity of using

modern and robust methods that can adequately face such challenges. Machine learning (ML), which has been vastly popularized in recent years, adapts well to the aforementioned aspect and rapidly adapts to the continuous evolution of data landscapes, transcending the most common limitations.

Motivated by this, we specify four operational aims for this study: 1) quantify the temporal evolution of citations and publications of ML-based decision-making based on documents indexed in the Web of Science and Scopus databases between 1990 and 2024; 2) recover the intellectual structure (co-citation/bibliographic coupling) and the main thematic clusters (keyword co-occurrence); 3) classify methodological portfolios by decision context; and 4) summarize sectoral adoption and interpretability/ethics signals to surface actionable gaps for future work. Stemming from this, this research endeavor attempts to answer the following research questions: a) *How have ML for decision-making publications and citations changed from 1990 to 2024 across nations and industries?*; b) *What topics and intellectual groups are revealed by bibliographic coupling and co-citation?*; c) *How do decision contexts (strategic, tactical, and operational; forecasting, classification, and optimization) fit with methodological portfolios (algorithms, interpretability orientation)?*; and d) *Which documented limitations and adoption trends serve as the driving forces behind a focused future research agenda?*

Two important contributions are made in this paper. Through a multi-technique bibliometric analysis, it first provides a theoretical framework that synthesizes influential research on machine learning in organizational decision-making, highlighting both current and emerging trends. Second, it offers a thorough analysis of the literature with useful advice for professionals involved in operational and strategic decision-making, including managers and policymakers.

The paper is structured in the following manner. The bibliometric theory and earlier research are reviewed in Section 2. Following the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines, Section 3 describes the data and methodology. The results of the bibliometric network are shown in Section 4. Leading machine learning applications in decision-making are methodically covered in Section 5. Section 6 concludes the study.

## **2. LITERATURE REVIEW**

Although their focus deviates from the decision-analytic lens, early surveys present machine learning (ML) as a general-purpose technology for inductive inference (Dietterich, 1996; Jordan and Mitchell, 2015). While Athey (2017) emphasizes integrating predictive models within explicit decision objectives, econometrics-focused contributions (Mullainathan and Spiess, 2017; Athey and Imbens, 2019) contend that prediction enhances causal reasoning in policy settings rather than replacing it. On the other hand, value creation is repositioned toward judgment, supervision, and complementary skills by management-of-work perspectives (Brynjolfsson and Mitchell, 2017; Agrawal *et al.*, 2019). Our review adds value by contrasting these camps. For instance, while decision-first approaches ensure relevance, they run the risk of under-exploiting complex patterns. Approaches where prediction is the primary goal, on the other hand, face interpretability and external validity constraints in high-stakes decisions. Our coding of “decision context” and “interpretability orientation” in the empirical corpus is driven by this nexus.

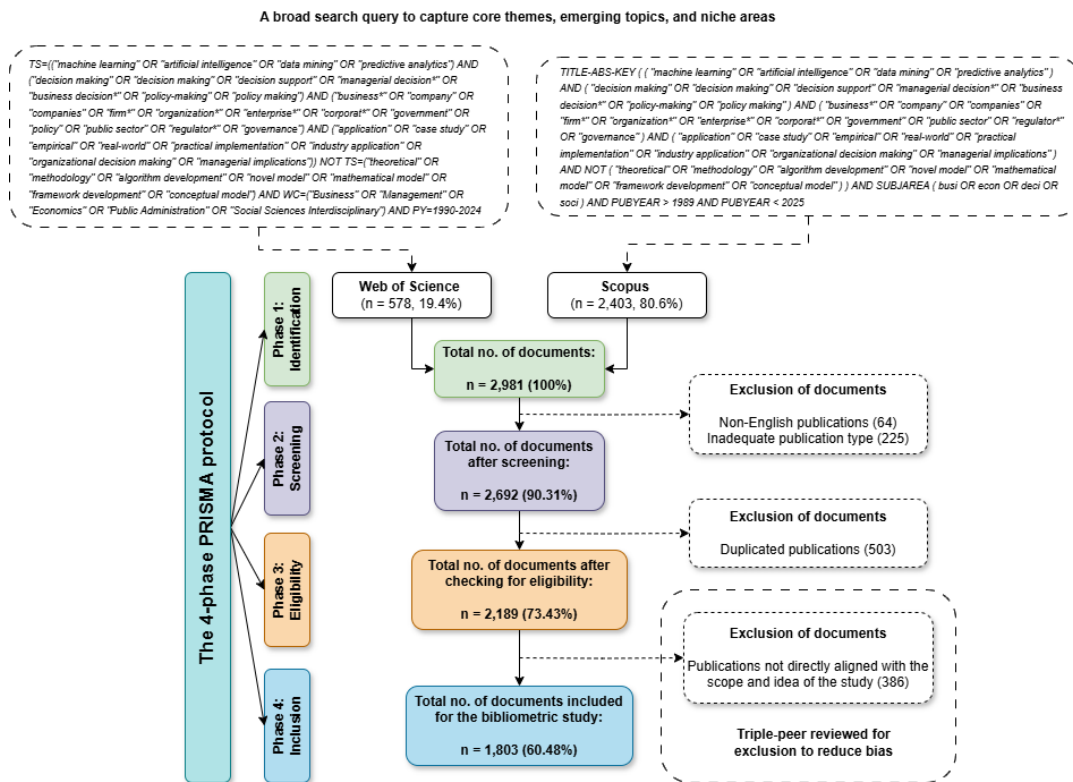
Empirical research is dominated by financial markets. For Brazilian day trading, Paiva *et al.* (2019) combine portfolio theory and support-vector machines, producing better risk-adjusted returns. In his review of reinforcement-learning advances in dynamic allocation and option pricing, Hambly (2023) points out that value-based and policy-gradient approaches perform better under regime changes than traditional stochastic control. By focusing on the most instructive observations, decision-centric active learning (Saar-Tsechansky and Provost, 2007) further lowers data-acquisition costs in credit scoring. Supply-chain and industrial environments both gain in this case, as Chen and Zhou (2020) combine real-time parameter estimation with model-predictive control to stabilize time-varying production systems, while Bertolini *et al.* (2021) document how predictive maintenance and defect detection drive Industry 4.0. Park and Yang (2022), whose interpretable LSTM predicts economic crises with explainable-AI overlays, and Guo *et al.* (2021), who combine neural networks with multi-criteria decision aiding to elicit stakeholder preferences, both address macro-policy and labor perspectives. De Laat (2018) contends that partial explainability, as opposed to complete code disclosure, strikes a balance between accountability and proprietary incentives, but transparency is still a limitation.

Despite the widespread adoption, disparities still exist between sectors. Empirical research is dominated by financial markets (Paiva *et al.*, 2019; Hambly, 2023), while public-sector and emerging-market deployments lag, especially in areas where interpretability and governance are most important (De Laat, 2018; Bucker *et al.*, 2022; Monken *et al.*, 2023). Our comparative synthesis between methods and decision fit and our agenda prioritizing contexts where human-AI complementarity is consequential are motivated by this asymmetry as well as the well-documented trade-off between accuracy and transparency (Kratsch *et al.*, 2021; Makridakis *et al.*, 2023).

### 3. METHODS AND DATA

Bibliometric analysis tracks conceptual evolution, identifies influential works, and quantifies scholarly output. In quickly expanding fields, it provides reproducible insights and improves literature reviews (Zupic and Čater, 2015) by offering a macro-level perspective (Donthu *et al.*, 2021). Performance analysis evaluates publishing and citation metrics (Lamovšek and Černe, 2023), while the five techniques, such as citation analysis, co-citation, bibliographic coupling, co-author, and co-occurrence analyses, provide insights into the topic (Marzi *et al.*, 2025). We supplemented these methods with a manual classification of each included document by decision context (strategic, tactical, operational), task type (e.g., forecasting, optimization, classification), and human-AI interaction mode in order to go beyond basic clusters and better capture the decision-making dimension. By using a hybrid approach, we were able to understand the networks both as abstract knowledge structures and in terms of the application of machine learning to real world decision-making. To ensure transparency through identification, screening, eligibility, and inclusion, this study uses the PRISMA protocol (Moher *et al.*, 2009; Page *et al.*, 2021). For identification, we use Web of Science Core Collection and Scopus (Baas *et al.*, 2020), while Google Scholar is not used due to noted problems with access and quality (Lim *et al.*, 2024). With an emphasis on practical applications, we review literature combining machine learning and decision-making in public administration, economics, and business and management, with a focus on papers published between 1990 and 2024. Figure 1 visually presents the flow of the process in acquiring and pre-processing data alongside the extensive search queries for both databases.

Figure 1: The conducted PRISMA protocol



(Source: Authors' work)

Combining these large databases through intensive data wrangling (see Koehler *et al.*, 2017), which included standardization and duplicate removal, was a significant challenge. Although Web of Science entries generally have more detailed metadata, Scopus was incorporated into its structure. There are several methods for merging, including open-source tools (Nikolić *et al.*, 2024), a three-step protocol (Caputo and Kargina, 2022), and bibliometrix/biblioshiny in R (Lim *et al.*, 2024). Despite warnings in favor of automation (Kasaraneni and Rosaline, 2024), we chose to manually merge using both full and abbreviated field names (Kumpulainen and Seppänen, 2022). The deduplication of documents under the screening phase was based on repetitive DOIs.

To avoid bias, eligibility was manually evaluated by three separate reviewers in the third phase. Documents (a total of 386 or 12.95% of the original sample) that were judged inappropriate by two or more reviewers were eliminated. Items that mentioned machine learning or decision-making in passing or that didn't have any real-world applications for ML decision-making were eliminated. These fields included healthcare, education, maritime, urban governance, agriculture, meteorology, psychology, and cybersecurity. A total of 1,803 (60.48%) documents were included for analysis beyond phase four.

## 4. MAIN INSIGHTS

### 4.1. Cluster analysis of authors' keywords

Using a LinLog/modularity normalization and a minimum occurrence threshold of 10, the author's keyword co-occurrence network offers an empirically supported perspective on the theoretical underpinnings of machine learning applications in decision-making. Three distinct theme clusters were produced when 50 of the 4,826 keywords were judged suitable for inclusion.

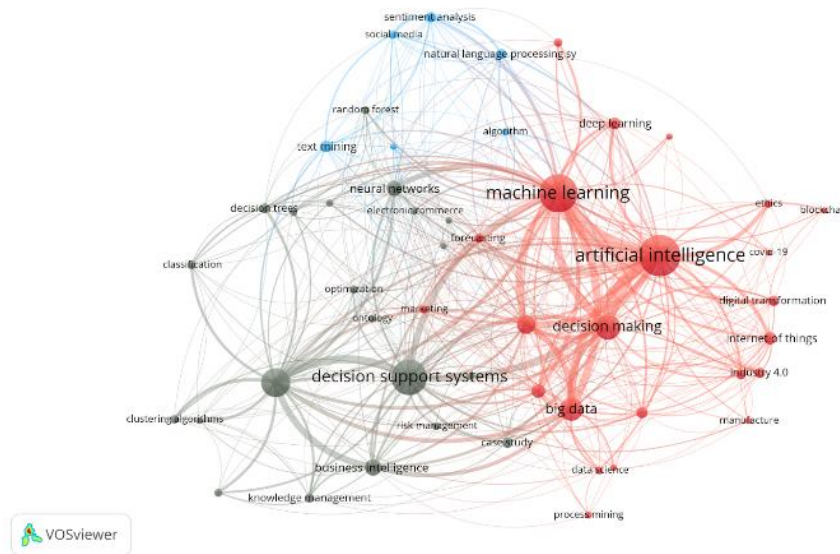
*Cluster 1: Core analytics, strategic decision processes, and emerging technologies.* The conceptual center of the field is represented by Cluster 1, which consists of 23 keywords and is anchored by three highly recurring terms: “artificial intelligence” (352 occurrences, total link strength: 359), “machine learning” (300 occurrences, total link strength: 354), and “decision making” (135 occurrences, total link strength: 196). To address complex business challenges, such as supply-chain optimization and Industry 4.0 applications, research encompasses deep learning, big-data analytics, and predictive tools. Growing concerns about interpretability, ethics, and long-term socioeconomic impact are reflected in emerging themes like explainable AI, digital transformation, and sustainability. All things considered, this cluster demonstrates how sophisticated analytical techniques meet organizational strategy and governance to tackle modern decision-making issues.

*Cluster 2: Foundational decision support, knowledge integration, and established algorithms.* Cluster 2 is dominated by methodological and infrastructural constructs that have long supported data-driven decision processes. As conceptual anchors, three keywords stand out: “business intelligence” (58 occurrences, total link strength: 97), “data mining” (179 occurrences, total link strength: 199), and “decision support systems” (257 occurrences, total link strength: 188). By using analytical techniques such as support vector machines, neural networks, decision trees, and classification algorithms integrated into knowledge management and data warehouse platforms, this cluster organizes and leverages structured information to support decision-making. For well-informed decision-making, it makes optimization, predictive modeling, and historical data analysis possible. The enduring significance of these frameworks for effective resource allocation and strategic insights across organizations is demonstrated by domain-specific applications in e-commerce and CRM.

*Cluster 3: Linguistic dimensions, content-driven analytics, and cognitive insight.* The third cluster, which is composed of six keywords, focuses on using unstructured textual data to extract meaning and structure. The terms “text mining” (31 occurrences, total link strength: 39), “sentiment analysis” (26 occurrences, total link strength: 47), and “natural language processing systems” (21 occurrences, total link strength: 34) are the most prominent examples of this cluster. This highlights a move toward language-based analytics of digital communications, which adds qualitative information to decision-making. Multiple insights into stakeholder preferences, consumer behavior, and public concerns are subsequently obtained by incorporating text analysis into quantitative models. This improves interpretability and directs the creation of policies, user experience design, and reflects in brand management.

We find two empirically supported bridges in the network beyond the three-way structure. For instance, explainable/transparent terms linking “decision support/assurance/governance” have grown since 2019 (Bücker *et al.*, 2022; Monken *et al.*, 2023), and second, there are weak but growing ties between reinforcement learning and optimization, indicating an early convergence of sequential choice and stakeholder-preference modeling (Guo *et al.*, 2021; Hambly *et al.*, 2023). These time-stamped connections support our assertion that the field is moving from purely algorithmic performance to auditable decision benefit, with a focus on interpretability and governance.

Figure 2: Network of authors' keywords.

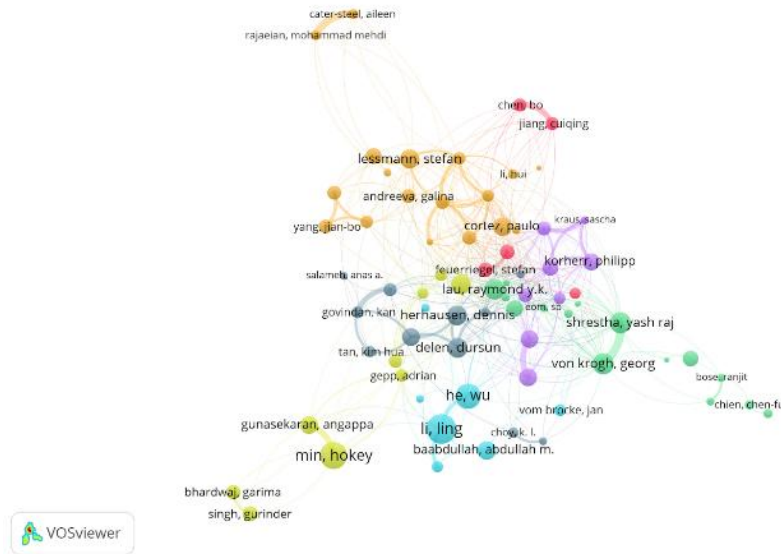


(Source: Authors' work)

#### 4.2. The most prominent authors

The following bibliographic coupling technique focuses on authors, employing fractional counting to measure connections based on shared references, which resulted in the visualization map in Figure 3. The analysis considers authors with a minimum of 2 publications and at least 10 citations per author, narrowing the dataset from 5,597 authors to 121 who meet the threshold for influence and productivity. Among these, the largest connected set consists of 73 authors, further divided into 7 distinct clusters, each with a minimum size of 5 authors. The analysis uses the LinLog/modularity method to optimize clustering, emphasizing relationships within and between groups. Weights are based on normalized citations, ensuring that the scientific impact of authors is equitably represented. The authors highlight topics like operations research, decision-support systems, and machine learning applications by grouping them into seven clusters based on common references. Well-known writers Lessmann (10.52), Cortez (9.36), and Stahlbock (6.26) act as intellectual linkages; their work covers supply-chain management, marketing, and the choice of AI methods (Hoffmann *et al.*, 2020; Lessmann *et al.*, 2021).

Figure 3: Bibliographic coupling of authors

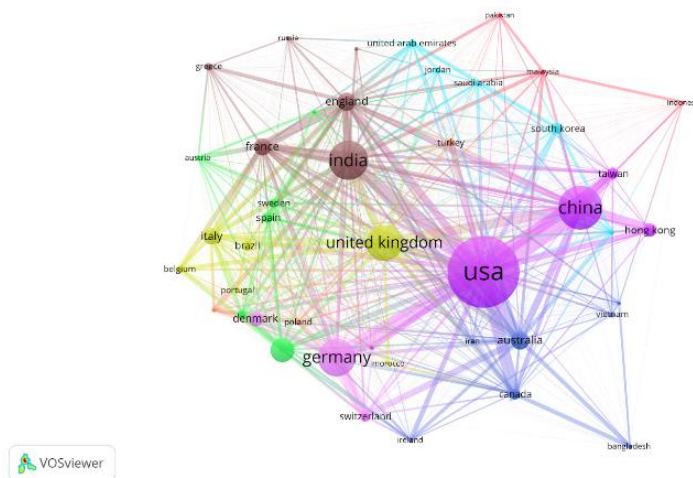


(Source: Authors' work)

### 4.3. Countries driving the global research

An additional bibliographic coupling technique on involved countries in the research field was conducted, based on shared references in their publications. The analysis includes countries with a minimum of 10 publications and at least 50 citations, reducing the dataset from 114 to 44 countries that meet the threshold for scientific impact and productivity. Using normalized citations as weights, the analysis emphasizes the relative influence of each country's contributions, ensuring fair comparison. The LinLog/modularity normalization method was applied to optimize clustering, resulting in 3 distinct clusters, each with a minimum size of 5 countries. Also, the minimal cluster size is now reduced to 3 countries. This produced 9 clusters presented in Figure 4.

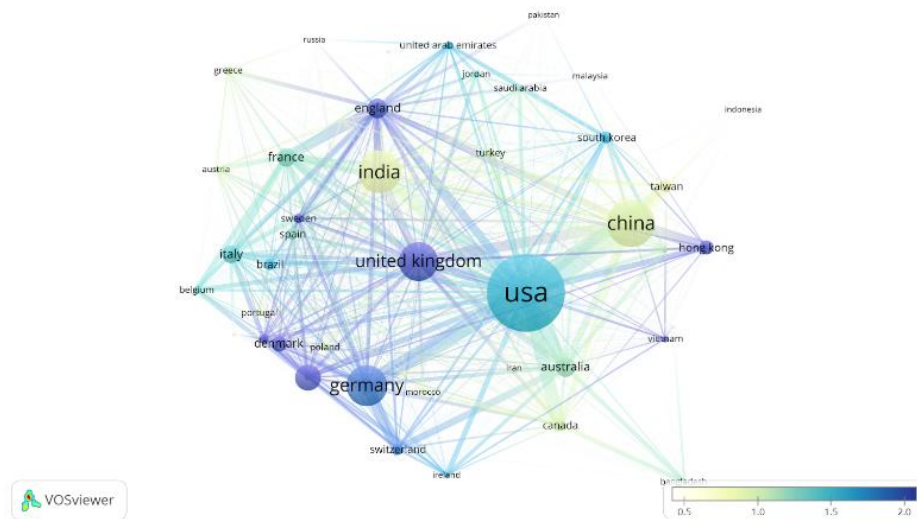
Figure 4: Bibliographic coupling of countries with scaled by normalized citations.



(Source: Authors' work)

Global ML research is led by the US, China, India, the UK, and Germany, and these countries frequently collaborate and have a high citation impact (see Figure 5). In addition to cross-regional ties between China, India, and the US, regional clusters include European alliances (Germany, Italy, France) and Asian-Middle Eastern ties (South Korea, Saudi Arabia, India). Emerging niches are reflected in peripheral contributors such as Greece and Bangladesh. While Taiwan, Germany, the US, and the UK dominated earlier work, bibliographic coupling reveals recent (post-2020) growth from the UAE, Saudi Arabia, Jordan, and India. Southeastern Europe, Latin America, and Africa are underrepresented regions that should be the focus of future studies.

Figure 5: Bibliographic coupling of countries by average normalized citations.



(Source: Authors' work)

## 5. SYSTEMATIC REVIEW AND DISCUSSION

The maps suggest a straightforward narrative. Uneven adoption across sectors can be explained by the fact that work in information systems and operations research is still only tangentially related to text-centric analytics. The rapid growth of studies that focus on interpretability after 2019 and their increased connections to decision support and assurance topics are indicators of growing governance concerns (De Laat, 2018; Bücken *et al.*, 2022; Monken *et al.*, 2023). While auditable and naturally interpretable approaches are more important in public sector contexts, application-driven strands in finance and logistics tend toward deep models (Makridakis *et al.*, 2023; Pugliese *et al.*, 2021). These trends can be inferred from the timing of links that connect clusters as well as from the composition of bibliographic coupling and co-citation communities. Using Table 1, we expand on the previous analysis by demonstrating how machine learning techniques are linked to decision-making contexts in the literature. The table shows how these approaches are frequently combined with neighboring approaches and how these combinations occur across various domains, rather than treating them as a strict taxonomy. This facilitates a clearer understanding of the types of techniques that are most frequently used in particular decision-making contexts and where their effects have been most noticeable.

This has two-fold implications relating to the scientific communities interested in exploring the applications of machine learning and those investigating how decision-making in organizations can be enhanced, particularly in a business context, and practitioners' communities, whose

members can embed the findings in organizational and unit-based strategies, policies, and activities on an operational level.

A fundamental application of machine learning is the use of data-driven predictive analytics to make intelligent decisions by forecasting by utilizing the relationships between predictors and outcomes (Mahdavejad *et al.*, 2018). Financiers, marketers, and policymakers use algorithms such as deep learning, sentiment analysis, explainable AI, and natural language processing (NLP) to forecast trends and comprehend behavior (Lessmann *et al.*, 2021; Monken *et al.*, 2023). Decision-making across sectors, including government and nonprofits, is improved by techniques like support vector machines (SVMs), decision trees, decision support systems (DSS), and neural networks, which optimize inventory, logistics, risk management, resource allocation, and strategy (see Sebastião *et al.*, 2020). Beyond traditional domains, ML supports context-aware computing, image/speech recognition, cybersecurity, digital attack prevention, IoT-based smart city traffic and energy management, and sustainable agriculture (Saba *et al.*, 2023), indicating future ML trajectories.

Table 1: Decision-making areas of global research and machine learning applications

Machine Learning Concepts	Decision-Making Areas	Key Domains
<b>Core AI/Analytics</b> Deep learning, neural networks, SVM, decision trees, random forests, explainable AI, generative AI, NLP, text mining, sentiment analysis	Forecasting and classification; sentiment/content analysis; strategic decision support; innovation management	Finance, marketing, and brand strategy, policy and public opinion, social media, and consumer behavior
<b>Optimization and planning</b> Genetic algorithms, fuzzy logic, multi-criteria decision making, time-series and predictive analytics, big data, and IoT analytics	Supply-chain and resource optimization; sustainability planning; digital-transformation strategy; real-time automation	Supply-chain management and Industry 4/5.0; smart cities; sustainable development; healthcare; project management
<b>Decision support and risk</b> Decision-support systems, data mining, clustering, expert systems, simulation, reinforcement, and unsupervised learning	Organizational decision support, CRM and e-government, risk assessment and crisis management, ERP and resilience	Cybersecurity and financial markets, e-commerce, agriculture, information systems, and engineering education
<b>Human-Centric and contextual ML</b> Adversarial ML, supervised learning, multi-objective optimization, automated decision frameworks	Human-in-the-loop support; UX and stakeholder preference elicitation; scenario simulation; robustness testing	Virtual/augmented reality; digital education; human-machine collaboration; HR management; emerging digital platforms

(Source: Authors' work)

The bibliometric approach employed yielded a multi-level insight into how decisions are formed based on applications of machine learning algorithms. For instance, we can highlight several key points.

While classical models sacrifice raw accuracy for auditability and small-sample stability, deep models are sensitive to drift and can leak target information in the absence of robust pipelines (Kratsch *et al.*, 2021; Makridakis *et al.*, 2023). Compared to models that are naturally interpretable, post-hoc explainers may not always meet accountability requirements in regulated settings (De Laat, 2018; Bücken *et al.*, 2022). Furthermore, decision quality depends not only on predictive fit, which is a problem that has been identified in the corpus but is rarely

operationalized, but also on matching learning objectives with cost-sensitive losses and constraints. ML techniques correspond with different types of decisions. For instance, explainable AI and fuzzy logic are used for policy ambiguities in strategic and policy decisions, while big-data analytics and deep learning are used for complex forecasting (Makridakis *et al.*, 2023). For routine tasks, operational decisions employ decision trees and random forests; for dynamic scheduling, they employ genetic algorithms and reinforcement learning (Pallathadka *et al.*, 2023). In DSS and business intelligence (BI), organizational support employs data mining, regression, and clustering. Moreover, the industries that drive innovation through significant budgets and data availability include healthcare, government, ICT, finance, marketing, management, and transportation (Kratsch *et al.*, 2021). Through improved decision-making, they optimize social media, customer relationship management (CRM), e-commerce, human resources, and resource management, proving that benefits outweigh costs. Although the field is dominated by tree-based models, explainable artificial intelligence, deep learning, reinforcement learning, and support vector machines, each has unique trade-offs. Deep learning is less appropriate for controlled or high-stakes situations because it frequently lacks transparency and is susceptible to data drift, despite its high accuracy. Support vector machines and tree-based ensembles, on the other hand, are simpler to validate but have the potential to oversimplify intricate decision environments. Although reinforcement learning promises flexibility, its actual application is constrained by the volume of data and the challenge of matching learning signals to actual decision-making goals. Even useful explainable AI tools occasionally provide post-hoc explanations instead of true interpretability.

The application of ML algorithms in decision-making processes in organizations is uneven across organizations and industries due to facilitating conditions and affecting barriers. Data quality, computing power, expertise, and digital readiness facilitate the adoption of machine learning in decision-making, but privacy, ethics, resistance, skill gaps, and costs impede its use in small and medium enterprises and under-resourced industries (Burggräf *et al.*, 2024). The goal of new accessible, humanized machine learning trends is to make their use more accessible to all. The post-2018 corpus clearly shows the rise of interpretability and ethical considerations, as keywords like “explainable AI,” “transparency,” and “accountability” become more common and occur alongside terms related to decision support. This temporal pattern suggests that, concurrent with the technical drive toward explainable models, conversations about ethical and governance issues are becoming more popular.

It is important to recognize several limitations, even though this bibliometric analysis offers valuable insights into the domain of machine learning and decision-making. First, by using a limited set of keywords, the study might have missed new or specialized fields. Second, references that are cited to criticize rather than support specific findings can distort the metrics used to measure citation impact, making them less representative of true scholarly influence. Third, there is potential for different interpretations of the data due to the subjective nature of the process of grouping important terms and assigning themes. Last but not least, the study mostly used keywords that the authors themselves supplied, which may have combined different ideas or missed minute differences that could have revealed other important patterns in ML-driven decision-making.

## **6. CONCLUSION**

Following PRISMA guidelines, we conducted a multi-technique bibliometric review of Scopus and Web of Science documents to investigate how machine learning aids in decision-making, pinpoint research hotspots, and suggest future directions. The United States, the United

Kingdom, India, Germany, and China are the top contributors to the growing publication trends. Although machine learning has shown great promise in automating and redefining decision-making processes, applications are still dispersed, which could lead to a gap between research and practice.

Classification algorithms help with segmentation, fraud detection, and prioritization, while predictive models dominate forecasting and diagnostics, providing strategic and operational insights. Techniques based on optimization improve efficiency and resource allocation (Sarker *et al.*, 2019). These approaches address both particular problems and cross-functional goals, and they cover a wide range of industries, including healthcare, government, ICT, finance, marketing, management, and transportation. We identified two main research concepts: a domain-driven focus on contextual constraints like ethical and regulatory considerations, and a technology-driven focus on algorithmic features (complexity, scalability, and interpretability). It is still difficult to integrate these viewpoints; research frequently uses a single technique in a limited domain (such as genetic algorithms for supply chains or deep learning for clinical imaging) without synthesizing the entire machine learning ecosystem.

Limits of the study focus on four frictions, such as alignment of decision losses with predictive targets, accountability transparency, robustness to drift and scarce labels, and external validity under domain shift and governance. The following are the methodological priorities: preference-aware reinforcement learning, prediction under uncertainty, cost-sensitive and constrained learning as well as causal machine learning for counterfactual decision support (Athey and Imbens, 2019; Makridakis *et al.*, 2023; Hambly *et al.*, 2023). We provide a framework for choosing the best approaches and identifying uncharted territory by mapping popular approaches, deep learning, neural networks, decision trees, text mining, sentiment analysis, natural language processing, and explainable AI, to their respective application domains. Both researchers and practitioners can benefit from this synthesis, since researchers can look into why particular algorithms perform well in particular situations, and practitioners can use machine learning more strategically, for example. Consequently, additional research questions can be derived from here. Compared to black boxes that are explained after the fact, do models that are inherently interpretable lead to better regulatory outcomes? When labels are limited, can human-in-the-loop active learning provide better decisions than end-to-end deep models? In what circumstances do hybrid ML and MCDM systems perform better on stakeholder satisfaction than pure ML optimizers? By doing this, the field can transcend disjointed applications and create a cohesive body of knowledge that connects advancements in algorithms with real-world enhancements in decision outcomes.

## REFERENCES

- Agrawal, A., Gans, J.S. and Goldfarb, A. (2019) "Artificial intelligence: The ambiguous labor market impact of automating prediction", *Journal of Economic Perspectives*, Vol. 33 No. 2, pp. 31-50. <https://doi.org/10.1257/jep.33.2.31>
- Athey, S. (2017) "Beyond prediction: Using big data for policy problems", *Science*, Vol. 355 No. 6324, pp. 483-485. <https://doi.org/10.1126/science.aal4321>
- Athey, S. and Imbens, G.W. (2019), "Machine learning methods that economists should know about", *Annual Review of Economics*, Vol. 11 No. 1, pp. 685-725. <https://doi.org/10.1146/annurev-economics-080217-053433>
- Baas, J., Schotten, M., Plume, A., Côté, G. and Karimi, R. (2020), "Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies",

- Quantitative Science Studies*, Vol. 1 No. 1, pp. 377–386.  
[https://doi.org/10.1162/qss\\_a\\_00019](https://doi.org/10.1162/qss_a_00019)
- Bertolini, M., Mezzogori, D., Neroni, M. and Zammori, F. (2021), “Machine Learning for industrial applications: A comprehensive literature review”, *Expert Systems with Applications*, Vol. 175, p. 114820. <https://doi.org/10.1016/j.eswa.2021.114820>
- Brynjolfsson, E. and Mitchell, T. (2017), “What can machine learning do? Workforce implications”, *Science*, Vol. 358 No. 6370, pp. 1530-1534. <https://doi.org/10.1126/science.aap8062>
- Bücker, M., Szepannek, G., Gosiewska, A., and Biecek, P. (2022), “Transparency, auditability, and explainability of machine learning models in credit scoring” *Journal of the Operational Research Society*, Vol. 73 No. 1, pp. 70-90. <https://doi.org/10.1080/01605682.2021.1922098>
- Burggräf, P., Steinberg, F., Sauer, C. R. and Nettesheim, P. (2024), “Machine learning implementation in small and medium-sized enterprises: Insights and recommendations from a quantitative study”, *Production Engineering*, Vol. 18, pp. 751-764. <https://doi.org/10.1007/s11740-024-01274-2>
- Caputo, A. and Kargina, M. (2022), “A user-friendly method to merge Scopus and Web of Science data during bibliometric analysis”, *Journal of Marketing Analytics*, Vol. 10 No. 1, pp. 82-88. <https://doi.org/10.1057/s41270-021-00142-7>
- Chen, Y. and Zhou, Y. (2020), “Machine learning based decision making for time varying systems: Parameter estimation and performance optimization”, *Knowledge-Based Systems*, Vol. 190, p. 105479. <https://doi.org/10.1016/j.knosys.2020.105479>
- De Laat, P.B. (2018), “Algorithmic decision-making based on machine learning from big data: can transparency restore accountability?”, *Philosophy and Technology*, Vol. 31 No. 4, pp. 525-541. <https://doi.org/10.1007/s13347-017-0293-z>
- Dietterich, T. (1996), “Machine learning”, *ACM Computing Surveys*, Vol. 28 No. 4es, p. 3. <https://doi.org/10.1145/242224.242229>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. and Lim, W. M. (2021), “How to conduct a bibliometric analysis: An overview and guidelines”, *Journal of Business Research*, Vol. 133, pp. 285-296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Guo, M., Zhang, Q., Liao, X., Chen, F.Y. and Zeng, D.D. (2021), “A hybrid machine learning framework for analyzing human decision-making through learning preferences”, *Omega*, Vol. 101, p. 102263. <https://doi.org/10.1016/j.omega.2020.102263>
- Hambly, B., Xu, R. and Yang, H. (2023), “Recent advances in reinforcement learning in finance”, *Mathematical Finance*, Vol. 33 No. 3, pp. 437-503. <https://doi.org/10.1111/mafi.12382>
- Hoffmann, N., Stahlbock, R. and Voß, S. (2020), “A decision model on the repair and maintenance of shipping containers”, *Journal of Shipping and Trade*, Vol. 5 No. 1, p. 22. <https://doi.org/10.1186/s41072-020-00070-2>
- Jordan, M. I. and Mitchell, T. M. (2015), “Machine learning: Trends, perspectives, and prospects”, *Science*, Vol. 349 No. 6245, pp. 255-260. <https://doi.org/10.1126/science.aaa8415>
- Kasaraneni, H. and Rosaline, S. (2024), “Automatic merging of Scopus and Web of Science data for simplified and effective bibliometric analysis”, *Annals of Data Science*, Vol. 11 No. 3, pp. 785-802. <https://doi.org/10.1007/s40745-022-00438-0>
- Koehler, M., Bogatu, A., Civili, C., Konstantinou, N., Abel, E., Fernandes, A. A. A., Keane, J., Libkin, L. and Paton, N. W. (2017), “Data context informed data wrangling”, *2017 IEEE International Conference on Big Data (Big Data)*, pp. 956-963. <https://doi.org/10.1109/BigData.2017.8258015>

- Kratsch, W., Manderscheid, J., Röglinger, M. and Seyfried, J. (2021), "Machine learning in business process monitoring: A comparison of deep learning and classical approaches used for outcome prediction", *Business and Information Systems Engineering*, Vol. 63, pp. 261-276. <https://doi.org/10.1007/s12599-020-00645-0>
- Kumpulainen, M. and Seppänen, M. (2022), "Combining Web of Science and Scopus datasets in citation-based literature study", *Scientometrics*, Vol. 127 No. 10, pp. 5613-5631. <https://doi.org/10.1007/s11192-022-04475-7>
- Lamovšek, A. and Černe, M. (2023), "Past, present and future: A systematic multitechnique bibliometric review of the field of distributed work", *Information and Organization*, Vol. 33 No. 2, 100446. <https://doi.org/10.1016/j.infoandorg.2022.100446>
- Lessmann, S., Haupt, J., Coussement, K. and De Bock, K. W. (2021), "Targeting customers for profit: An ensemble learning framework to support marketing decision making", *Information Sciences*, Vol. 557, pp. 286-301. <https://doi.org/10.1016/j.ins.2019.05.027>
- Lim, W. M., Kumar, S. and Donthu, N. (2024), "How to combine and clean bibliometric data and use bibliometric tools synergistically: Guidelines using metaverse research", *Journal of Business Research*, Vol. 182, 114760. <https://doi.org/10.1016/j.jbusres.2024.114760>
- Mahdavinejad, M. S., Rezvan, M., Barekatin, M., Adibi, P., Barnaghi, P. and Sheth, A. P. (2018), "Machine learning for Internet of Things data analysis: A survey", *Digital Communications and Networks*, Vol. 4 No. 3, pp. 161-175. <https://doi.org/10.1016/j.dcan.2017.10.002>
- Makridakis, S., Spiliotis, E., Assimakopoulos, V., Semenovoglou, A. A., Mulder, G. and Nikolopoulos, K. (2023), "Statistical, machine learning and deep learning forecasting methods: Comparisons and ways forward", *Journal of the Operational Research Society*, Vol. 74 No. 3, pp. 840-859. <https://doi.org/10.1080/01605682.2022.2118629>
- Marzi, G., Balzano, M., Caputo, A. and Pellegrini, M. M. (2025), "Guidelines for Bibliometric-Systematic Literature Reviews: 10 steps to combine analysis, synthesis and theory development", *International Journal of Management Reviews*, Vol. 27 No. 1, pp. 81-103. <https://doi.org/10.1111/ijmr.12381>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and The PRISMA Group (2009), "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement", *PLoS Medicine*, Vol. 6 No. 7, e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Monken, A., Ampeh, W., Haberkorn, F., Krishnaswamy, U. and Batarseh, F. A. (2023), "Assuring AI methods for economic policymaking", In *AI Assurance*, Academic Press, pp. 371-427. <https://doi.org/10.1016/B978-0-32-391919-7.00025-1>
- Mullainathan, S. and Spiess, J. (2017), "Machine learning: an applied econometric approach", *Journal of Economic Perspectives*, Vol. 31 No. 2, pp. 87-106. <https://doi.org/10.1257/jep.31.2.87>
- Nikolić, D., Ivanović, D. and Ivanović, L. (2024), "An open-source tool for merging data from multiple citation databases", *Scientometrics*, Vol. 129 No. 7, pp. 4573-4595. <https://doi.org/10.1007/s11192-024-05076-2>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021), "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews", *Systematic Reviews*, Vol. 10 No. 1, p. 89. <https://doi.org/10.1186/s13643-021-01626-4>
- Paiva, F.D., Cardoso, R.T.N., Hanaoka, G.P. and Duarte, W.M. (2019), "Decision-making for financial trading: A fusion approach of machine learning and portfolio selection", *Expert Systems with Applications*, Vol. 115, pp. 635-655. <https://doi.org/10.1016/j.eswa.2018.08.003>

- Pallathadka, H., Ramirez-Asis, E. H., Loli-Poma, T. P., Kaliyaperumal, K., Ventayen, R. J. M. and Naved, M. (2023), “Applications of artificial intelligence in business management, e-commerce and finance”, *Materials Today: Proceedings*, Vol. 80, pp. 2610-2613. <https://doi.org/10.1016/j.matpr.2021.06.419>
- Park, S. and Yang, J.-S. (2022), “Interpretable deep learning LSTM model for intelligent economic decision-making”, *Knowledge-Based Systems*, Vol. 248, p. 108907. <https://doi.org/10.1016/j.knosys.2022.108907>
- Pugliese, R., Regondi, S. and Marini, R. (2021), “Machine learning-based approach: Global trends, research directions, and regulatory standpoints”, *Data Science and Management*, Vol. 4, pp. 19-29. <https://doi.org/10.1016/j.dsm.2021.12.002>
- Saar-Tsechansky, M. and Provost, F. (2007), “Decision-centric active learning of binary-outcome models”, *Information Systems Research*, Vol. 18 No. 1, pp. 4-22. <https://doi.org/10.1287/isre.1070.0111>
- Saba, T., Rehman, A., Haseeb, K., Bahaj, S. A. and Lloret, J. (2023), “Trust-based decentralized blockchain system with machine learning using Internet of agriculture things”, *Computers and Electrical Engineering*, Vol. 108, 108674. <https://doi.org/10.1016/j.compeleceng.2023.108674>
- Sarker, I. H., Kayes, A. S. M. and Watters, P. (2019), “Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage”, *Journal of Big Data*, Vol. 6 No. 1, pp. 1-28. <https://doi.org/10.1186/s40537-019-0219-y>
- Sebastião, H. M. C. V., Godinho, P. M. C. and Westgaard, S. (2020), “Using Machine Learning to Profit on the Risk Premium of the Nordic Electricity Futures”, *Scientific Annals of Economics and Business*, Vol. 67 (SI), pp. 1-17. <https://doi.org/10.47743/saeb-2020-0024>
- Zupic, I. and Čater, T. (2015), “Bibliometric methods in management and organization”, *Organizational Research Methods*, Vol. 18 No. 3, pp. 429-472. <https://doi.org/10.1177/1094428114562629>

## STAKEHOLDERS AND SUSTAINABILITY IN THE BALKAN WINE INDUSTRY: A MAPPING APPROACH TO GOVERNANCE

**Tamara Kaftandjieva**

*School of Economics and Business,  
University of Ljubljana, Slovenia  
tamara.kaftandzieva@gmail.com*

**Metka Tekavčič**

*School of Economics and Business,  
University of Ljubljana, Slovenia  
metka.tekavcic@ef.uni-lj.si*

### **ABSTRACT**

*The Balkan wine industry, dominated by small and medium-sized enterprises (SMEs), faces increasing pressure to adopt sustainability practices, yet efforts remain fragmented due to misaligned stakeholder interests and weak governance coordination. This paper applies stakeholder mapping to identify key actors, relationships, and governance levers influencing sustainability across Slovenia, Serbia, and North Macedonia. Using eight semi-structured interviews, document analysis, and triangulation through policy review, the study constructs a stakeholder map based on power, interest, and feasibility dimensions. Stakeholders were classified as internal or external and positioned within a power–interest matrix using qualitative coding and document triangulation. Findings reveal four clusters: Key Players, Context Setters, Subjects, and Minimal Stakeholders, whose interactions shape governance outcomes. Misalignments emerge between policy ambition and SME capacity, between market power and sustainability incentives, and between fragmented institutional levels. The analysis highlights how high interest often fails to produce action when feasibility is constrained. The study contributes a region-specific governance roadmap that links stakeholder alignment with actionable policy levers. It provides insights into how SMEs, associations and policymakers can strengthen coordination, integrate demand-side incentives and accelerate sustainability transitions in the Balkan wine value chain.*

**Keywords:** *Stakeholder mapping; Power–interest matrix; Wine value chain; Sustainability governance; Balkan region; Small and medium-sized enterprises (SMEs); Stakeholder clusters; Value chain governance*

**Jel Classification:** *Q56*

### **1. INTRODUCTION**

Sustainability integrates environmental, economic, and social dimensions of development, requiring a balance between ecological responsibility, competitiveness, and social equity (Elkington and Rowlands, 1999; Lozano, 2008). In the agri-food sector, sustainability transitions depend on governance mechanisms that reconcile diverse stakeholder expectations and align private incentives with public goals (Reed *et al.*, 2009; Alonso, 2016; Gilinsky *et al.*, 2016; Probst *et al.*, 2021). Within the European Green Deal and the Common Agricultural Policy (CAP), sustainability now underpins rural development priorities, yet implementation remains uneven across small and medium-sized enterprises (SMEs), particularly in emerging wine regions such as the Balkans (Erjavec *et al.*, 2021; Martins *et al.*, 2022). The wine industry

offers a valuable lens for studying these governance challenges because it combines local production, cultural heritage, and export orientation (Baiano, 2021).

Western European regions, such as Bordeaux, Tuscany, and Rioja, have progressed through coordinated certification schemes, collective branding, and strong policy–industry interfaces. In contrast, Balkan producers operate in fragmented institutional settings with limited cross-border coordination and uneven policy enforcement (Mariani and Vastola, 2015; Merli *et al.*, 2018). Although SMEs preserve local identity, maintain rural employment, and sustain cultural landscapes, they face persistent financial, informational, and administrative barriers that hinder the adoption of sustainability practices (Flores and Medeiros, 2019; Licastro and Sergi, 2021; Lozano, 2008; Reed *et al.*, 2009). Research on the Balkan wine sector remains scarce and primarily descriptive, focusing on competitiveness and exports rather than stakeholder alignment or governance (Santini *et al.*, 2013). There is little comparative evidence on how sustainability is interpreted along the value chain or how governance mechanisms can strengthen coordination across multiple actors.

This paper addresses that gap by applying a stakeholder-mapping approach to identify how power, interest, and feasibility interact across the Balkan wine value chain. It examines Slovenia, Serbia, and North Macedonia, countries with shared viticultural traditions but differing institutional capacities. The study links institutional and market dynamics to policy design and practice, identifying where misalignments occur and which governance levers could realign stakeholder incentives.

Accordingly, this paper addresses three research questions: (1) Which stakeholders influence or are affected by sustainability, and how do their power, interest, and feasibility differ?; (2) Where do misalignments between power and interest create barriers to implementation?; (3) Which governance mechanisms could support feasible, SME-sensitive sustainability transitions? The paper contributes in two ways. First, it extends stakeholder mapping to a post-transition, SME-based wine sector where informal networks and limited institutional capacity shape sustainability differently from Western European regions. Second, it develops an actionable governance roadmap that identifies priority interventions, responsible actors, and short-term feasibility for SMEs.

## **2. LITERATURE REVIEW**

### **2.1. Stakeholder Theory and Sustainability**

Stakeholder theory shifts the focus from shareholders to all actors who affect or are affected by organisational outcomes (Freeman, 2010). Stakeholder salience depends on power, legitimacy, and urgency (Mitchell *et al.*, 1997). In sustainability transitions, outcomes depend not only on technology or policy but also on negotiation among actors with conflicting interests and asymmetric power (Bansal, 2005; Dyllick and Hockerts, 2002).

In food and agricultural sectors, sustainability is shaped by supply chain complexity, territorial embeddedness, and cultural identity (De Steur *et al.*, 2016). However, most studies focus on large-scale Western European contexts and overlook SME-based, post-transition economies (Dodds *et al.*, 2013; Gabzdylova *et al.*, 2009). Scholars argue that stakeholder theory must account for place-based and cultural dimensions when applied to sectors such as wine (Mariani and Vastola, 2015; Flores and Medeiros, 2019).

## 2.2. Stakeholder Mapping Approaches

Stakeholder mapping operationalises stakeholder theory by classifying actors according to their influence and interest. The power–interest matrix is widely used, dividing stakeholders into high power/high interest, high power/low interest, low power/high interest, and low power/low interest groups (Ackermann and Eden, 2011). This framework helps identify engagement priorities and governance gaps.

Recent literature emphasises that mapping is not only diagnostic but also collaborative. Sandroni *et al.* (2022) describe stakeholder mapping as a transdisciplinary practice fostering shared understanding among scientists, policymakers, and communities. Monticone *et al.* (2024) show how network structure and actor centrality shape innovation diffusion across food supply chains. Similarly, Van den Akker *et al.* (2024) find that powerful institutions in precision agriculture can delay sustainable practices. These studies illustrate that mapping can expose governance imbalances where actors with the highest stakes lack decision-making power, while those with structural power may have low interest in sustainability.

## 2.3. Sustainability in the Wine Industry

Research on wine sustainability spans environmental, social, and economic dimensions. Environmental studies quantify carbon and water footprints (Schäufele and Hamm, 2017), agrochemical impacts (Corbo *et al.*, 2014), and climate-change risks (Mozell and Thach, 2014). Economic analyses explore the relationship between sustainability capabilities and performance (Annunziata *et al.*, 2018; Broccardo and Zicari, 2020), while social research examines community engagement and labor conditions (Santini *et al.*, 2013). Because wine production is embedded in territory and culture, sustainability also involves maintaining identity and heritage (Flores, 2018; Mariani and Vastola, 2015). Consumers increasingly value authenticity and ethical production (Schäufele and Hamm, 2017), creating potential competitive advantages.

Western European wine regions have developed structured sustainability frameworks through certification systems, cooperative governance, and public support (Annunziata *et al.*, 2018; Merli *et al.*, 2018). In contrast, Balkan regions face fragmented institutions, limited incentives, and inconsistent regulation (Licastro and Sergi, 2021). Yet, they possess strong potential in wine tourism, indigenous varieties, and rural development (Vukojević *et al.*, 2022; Oltean and Gabor, 2022). These conditions justify using stakeholder mapping to analyse where sustainability in the Balkan wine value chain is enabled, resisted, or constrained.

## 3. METHODOLOGY

### 3.1. Research Design

This study adopts a qualitative, concept-driven approach grounded in stakeholder theory (Freeman, 2010; Mitchell *et al.*, 1997). It combines document analysis with exploratory semi-structured interviews to understand how sustainability is interpreted, enabled, or constrained along the Balkan wine value chain. The purpose was not statistical representativeness but to capture diverse perspectives across value-chain stages and governance levels, consistent with stakeholder mapping in sustainability research (Sandroni *et al.*, 2022; Monticone *et al.*, 2024).

### 3.2. Data Collection

Data were gathered in two phases. First, academic literature, EU sustainability frameworks (Green Deal; CAP Strategic Plans 2023–2027), and national wine and rural development strategies from North Macedonia, Serbia, and Slovenia were analysed to identify stakeholder

groups and policy contexts. Second, eight semi-structured interviews were conducted between April and July 2024 with winery owners/managers (n = 3), policymakers (n = 2), a wine association representative (n = 1), an NGO representative (n = 1), and a distributor (n = 1). All interviews were recorded and transcribed with consent.

Several interviewees held multiple roles (e.g., winery owner and association board member), explaining why more stakeholder perspectives and quotations are presented than the number of individuals interviewed. Stakeholder views were triangulated with policy reports, media interviews, and association documents to strengthen reliability rather than inflate the sample. The focus was on diversity of perspectives, appropriate for exploratory stakeholder mapping (Flores and Medeiros, 2019; Sandroni *et al.*, 2022).

### 3.3. Stakeholder Identification and Classification

Stakeholders were categorised using a two-layered approach aligned with literature and policy frameworks (Santini *et al.*, 2013; Sandroni *et al.*, 2022). Internal stakeholders include actors directly involved in the production, processing, or distribution of wine who create economic value, such as vineyard owners, winery managers, employees, oenologists, cooperatives, and distributors. External stakeholders are those who do not produce wine themselves but influence sustainability through regulation, finance, certification, knowledge, or community engagement. These include ministries, local governments, EU bodies, financial institutions, certification authorities, NGOs, local communities, research institutes, and tourism operators.

This classification was initially derived from policy documents (EU Green Deal, CAP, and national wine strategies) and refined using interview insights. It was further expanded to incorporate both vertical governance relations (EU → national → SME winery) and horizontal dynamics (winery–association, NGO–local authority), reflecting the multi-level nature of sustainability governance in agri-food systems (Monticone *et al.*, 2024).

### 3.4. Operationalisation of Power, Interest and Feasibility

Stakeholders were positioned in the power–interest matrix through qualitative coding of interview transcripts and document sources. Table 1 presents the operational definitions and indicators of the three constructs.

Table 1: The operational definitions and indicators

<i>Construct</i>	<i>Definition</i>	<i>Indicators / Evidence Sources</i>
<b>Power</b>	Ability to influence sustainability decisions, control resources, or shape rules	Formal authority (ministries, EU bodies), financial control (banks, CAP/IPARD funds), standard-setting power (certification bodies), market leverage (distributors, associations)
<b>Interest</b>	The degree to which sustainability is relevant to stakeholder goals, values, or strategies	Interview statements, participation in sustainability projects, inclusion of sustainability in strategic plans, policies, or public discourse
<b>Feasibility</b>	Practical capacity, especially for SMEs, to act on sustainability expectations	Access to finance, human resources, technical knowledge, administrative burden, support programmes (e.g., advisory services, subsidies)

(Source: Authors' elaboration based on Freeman (2010); Mitchell *et al.* (1997); Ackermann and Eden (2011); and Moore and Manring (2009)).

The criteria are built on stakeholder theory (Mitchell *et al.*, 1997), governance literature (Sandroni *et al.*, 2022; Monticone *et al.*, 2024), and SME sustainability studies (Moore and Manring, 2009; Martins *et al.*, 2022). Power and interest were coded as low, medium or high (Ackermann and Eden, 2011). Feasibility, used to interpret motivation, especially for high-interest/low-capacity actors (SMEs, NGOs), was not a separate axis. Coding followed three steps: (1) extracting stakeholder statements; (2) assigning relative power and interest; and (3) cross-validating with policy and funding documents and association reports. A traceability table (Table 2) links each stakeholder group to its matrix position and supporting evidence.

#### 4. FINDINGS: STAKEHOLDER MAPPING AND ANALYSIS

##### 4.1. Overview of Stakeholder Distribution

Twelve stakeholder groups active in sustainability across the Balkan wine value chain were identified. They differ in power (ability to influence decisions), interest (motivation to engage) and feasibility (capacity to act). Based on interviews and documents, they were placed in four quadrants of a power–interest matrix (Figure 1) supported by a traceability table (Table 2).

##### 4.2. Stakeholder Traceability and Placement

Table 2 summarizes each stakeholder cluster, its position in the power–interest grid, and supporting evidence.

Table 2: Stakeholder Clusters, Power–Interest Position and Evidence Source

Cluster	Stakeholder Groups Included	Power	Interest	Evidence (Interview/Docum
Key Players	Ministries of Agriculture, EU institutions (CAP/DG AGRI), Winery owners/managers, Wine associations	High to Medium–High	High	INT1, INT2, INT6, INT7, INT8; CAP Strategic Plans 2023–2027; National Rural Development Strategy 2023
Context Setters	Distributors/intermediaries, Certification bodies, Financial institutions	High	Low to Medium	INT1, INT5; EU Organic Regulation 2023; IPARD II Guidelines 2024
Subjects	Local communities, NGOs, Tourism operators, Winery employees	Low to Medium	High	INT3, INT4, INT6, INT8; Tourism Strategy Slovenia 2023
Minimal Stakeholders	Bulk wine buyers, External investors, Low-engagement suppliers	Low	Low	INT5, INT6; Trade Data Report 2023

(Source: Author’s elaboration using interview and document data)

Note: The same interviewees may appear across clusters due to overlapping roles within the wine value chain. For instance, INT1 is both a winery owner and association board member; INT5 manages a distribution company influencing certification; INT6 combines winery management with cross-border sustainability work; and INT8 represents an NGO collaborating with wineries. Coding was role-based rather than individual-based, reflecting the multifunctional nature of actors in SME-dominated sectors.

The stakeholder matrix is illustrated in Figure 1.

Figure 1: Stakeholder Mapping of the Balkan Wine Industry (Power – Interest Matrix)



(Source: Author’s elaboration using interview and document data)

#### 4.2.1 Key Players: High Power, High Interest

Key Players include national Ministries of Agriculture, EU institutions, wine associations, and proactive winery owners/managers. They shape policies, funding mechanisms, and sector-wide coordination. Five of eight interviewees (INT1, INT2, INT6, INT7, INT8) emphasised that sustainability is now integrated into agricultural policy, eco-schemes, and national wine strategies. As one policymaker stated, “Sustainability targets exist in all three analysed countries, though implementation depends on administrative capacity and available funding” (INT2). Similarly, wineries link sustainability with export competitiveness, tourism value, and vineyard resilience. One winery owner noted: “We are ready to go greener, but we can’t do it alone; collaboration is the missing ingredient” (INT1). However, despite high interest, many face feasibility constraints: administrative burdens, lack of staff, and limited access to CAP/IPARD funds, particularly for small wineries. Wine associations help bridge gaps by disseminating knowledge, organising workshops, and mediating between policymakers and producers, though cross-border collaboration remains weak (INT7).

#### 4.2.2 Context Setters: High Power, Low Interest

Context Setters (distributors, certification bodies, financial institutions) hold considerable economic or regulatory influence but show limited or inconsistent commitment to sustainability. Distributors determine which wines reach domestic and regional markets, yet purchasing criteria remain dominated by price, reliability, and volume rather than environmental or social performance. As one distributor explained: “Our buyers want low prices, not sustainability labels. If a wine is certified but costs more, it simply doesn’t move” (INT5). Certification bodies control access to sustainability labels but are often perceived as bureaucratic, expensive, and poorly adapted to SMEs. Four interviewees described certification as “paperwork rather than improvement” (INT1, INT3, INT6, INT8). Financial institutions and

CAP/IPARD mechanisms officially promote green transitions, yet access depends on collateral, compliance, and administrative capacity, thresholds often too high for small wineries (INT2, INT7). Consequently, sustainability often remains a procedural requirement rather than a strategic priority for this group.

#### 4.2.3 Subjects: Low Power, High Interest

Subjects: local communities, NGOs, winery employees, and tourism operators demonstrate strong interest in sustainability but lack institutional power. Interviewees (INT3, INT4, INT6, INT8) highlighted local concerns about heritage, landscapes, and rural livelihoods, yet noted that decisions are made without meaningful community involvement. One tourism actor stated: *“Wine tourism affects our villages and landscapes, yet decisions are made without asking us”* (INT4). NGOs promote awareness and biodiversity projects but remain consultative rather than decision-making actors. Tourism operators view sustainability as vital for wine tourism competitiveness, but they lack authority in agricultural or certification policy. Winery employees were mentioned in relation to labour conditions and pesticide exposure (INT3, INT6), yet they rarely participate in governance discussions.

#### 4.2.4 Minimal Stakeholders: Low Power, Low Interest

Minimal Stakeholders include bulk wine buyers, external investors, and low-engagement suppliers. Their influence and interest in sustainability are both limited, except when it directly affects profitability or export requirements. Interview evidence (INT5) shows that bulk buyers primarily evaluate wine on price, supply volume, and consistency. External investors often prioritise short-term gains over long-term sustainability commitments. Input suppliers (for packaging, logistics, or chemicals) remain operational rather than strategic actors and are seldom involved in sustainability planning.

### 4.3. Misalignments and Tensions

The stakeholder mapping reveals several structural misalignments that explain why sustainability in the Balkan wine sector remains fragmented despite being widely endorsed in policy and discourse. A key tension lies between ambitious policy frameworks and the limited feasibility for SMEs to implement them. Although sustainability is embedded in EU and national strategies, many wineries lack the financial, human, and administrative capacity to meet certification requirements or access CAP/IPARD funding. As one winery owner noted, *“EU rules seem made for bigger producers; we don’t have the people or time to handle all the paperwork.”* (INT6). This gap shows that high interest does not always translate into action when feasibility is constrained. A second misalignment concerns the role of market intermediaries, particularly distributors and financial institutions, whose decisions largely remain driven by price and volume rather than sustainability criteria. As one distributor admitted, *“Our buyers want low prices, not sustainability labels.”* (INT5). This weakens economic incentives for wineries to adopt sustainability practices unless exporting to markets where certification is required. Certification systems constitute another source of tension. While intended to ensure quality and traceability, they are widely perceived by wineries as bureaucratic, expensive, and insufficiently adapted to small producers. Several interviewees described certification as *“paperwork more than improvement,”* which discourages engagement rather than enabling it. Stakeholders with high interest but low power, such as local communities, NGOs, tourism actors, and winery employees, also experience exclusion from formal decision-making processes despite being directly affected by vineyard expansion, landscape changes, or tourism flows. A tourism representative explained, *“Wine tourism affects our villages, yet decisions are made without asking us.”* (INT4). Their participation remains informal and reactive rather than institutionalised. Finally, coordination across the

value chain and governance levels is weak. Ministries, associations, and producers often act in isolation, resulting in disconnected initiatives rather than an integrated governance framework. Cross-border collaboration among Serbian, Slovenian, and Macedonian institutions is sporadic, despite shared challenges and opportunities.

Together, these tensions indicate that the main barriers to sustainability are structural misalignments between power, interest, and capacity rather than a lack of awareness. These findings provide the analytical basis for the Discussion section.

## 5. DISCUSSION

This section interprets the stakeholder-mapping results through the lenses of governance, power relations, and feasibility. It explains why sustainability progresses unevenly in the Balkan wine sector and how stakeholder dynamics shape implementation. Four main themes emerge: **power asymmetries**, **interest without influence**, **feasibility constraints**, and **fragmented coordination**. A final part reflects on the study's contribution and proposes demand-side governance instruments, as suggested by reviewers.

### 5.1. Power and Governance Asymmetries

The results show that actors with the strongest influence over sustainability in the Balkan wine sector are national ministries, EU institutions, producer associations, and proactive wineries. These stakeholders shape policy, funding allocation, and sector-wide coordination, reflecting governance structures within the Common Agricultural Policy (CAP), where sustainability priorities are defined at the EU level and implemented nationally (European Commission, 2023; Merli *et al.*, 2018). This supports findings that sustainability in viticulture is largely policy-driven rather than market-driven (Flores and Medeiros, 2019).

Conversely, distributors, certification bodies, and financial institutions hold high power but demonstrate low or inconsistent interest in sustainability. Their decisions shape which wines reach consumers and under what conditions, yet price, volume, and reliability remain the main purchasing criteria (Van Den Akker *et al.*, 2024). This alignment with market-based logic over sustainability performance mirrors patterns in global agri-food chains (Santini *et al.*, 2013; Gilinsky *et al.*, 2016). Certification bodies, despite promoting sustainability in principle, are often perceived as administrative gatekeepers rather than facilitators, reinforcing bureaucratic rather than participatory governance (Mitchell *et al.*, 1997; Moore and Manring, 2009).

### 5.2. Interest Without Influence: Stakeholders at the Margins

Stakeholders such as local communities, NGOs, tourism actors, and winery employees show strong interest in sustainability but limited institutional power. Their concerns relate to landscape preservation, cultural identity, and rural development, yet they remain largely excluded from decision-making processes. This reflects broader evidence that stakeholders with moral or cultural legitimacy often lack structural power in agricultural governance (Reed *et al.*, 2009; Mariani and Vastola, 2015).

Tourism actors, particularly in Slovenia and Serbia, connect sustainability with economic potential via wine tourism, but remain outside formal agricultural policymaking. Winery employees face direct implications from sustainability decisions, such as labour standards and pesticide exposure, yet have almost no representation in governance structures. These dynamics indicate that interest alone does not translate into influence without institutional mechanisms.

### 5.3. Feasibility Constraints in SME-Dominated Wine Sectors

Even when stakeholders demonstrate high interest, especially SMEs, associations, and communities, feasibility often limits implementation. The Balkan wine sector is dominated by SMEs that face financial constraints, limited administrative capacity, and high compliance costs. Access to CAP and IPARD funds requires paperwork, co-financing, and collateral that many SMEs lack (Erjavec *et al.*, 2021; Martins *et al.*, 2022). Certification systems are perceived as costly, rigid, and designed for larger producers, discouraging participation from small wineries (INT6). These findings align with research demonstrating that sustainability frameworks often overlook SME constraints (Moore and Manning, 2009).

### 5.4. Fragmented Coordination Across the Value Chain and Territory

Beyond individual stakeholder constraints, sustainability in the Balkan wine sector is also limited by fragmented coordination across institutional, territorial, and value chain levels. Stakeholders often operate in isolation rather than as part of an integrated governance system.

**Horizontal coordination gaps** appear between stakeholders at the same level of the value chain. Producers, distributors, certification bodies, and tourism actors frequently lack structured platforms for dialogue or joint action. For example, associations can facilitate knowledge-sharing and regional branding but have limited influence on market incentives or certification policies. Distributors and tourism actors rarely engage with producers on sustainability planning, despite influencing consumer demand and value perception.

**Vertical coordination gaps** exist between governance levels: EU, national, regional and local. Sustainability objectives are formulated at the EU level through CAP and Green Deal policies, interpreted by national ministries, and expected to be implemented by local administrations and SMEs. However, this top-down path often overlooks local variations in infrastructure, knowledge systems, and cultural practices. Interviewees noted that while policies exist on paper, practical support mechanisms, such as advisory services, local partnerships, or simplified application procedures, are often missing.

**Cross-border fragmentation** further restricts collective action. The Balkan wine sector is regionally diverse but institutionally disconnected. Associations from Slovenia, Serbia, and North Macedonia develop sustainability initiatives separately rather than collaboratively, limiting opportunities for shared certification schemes, joint tourism routes, or common branding. Unlike Western European regions such as Tuscany or Bordeaux, regional identity in the Balkans is not institutionally organised around wine sustainability. This pattern reflects what Alonso (2016) observed in other wine regions, where wineries act as local anchors yet rarely form cohesive governance structures.

This lack of coordination reinforces the power of individual actors, ministries, large wineries, or distributors, while weakening collective capacity to innovate or negotiate sustainability standards. As argued in value chain governance literature, sustainability transitions require not only individual commitment but networked alignment across economic and territorial actors (Trienekens, 2011; Probst *et al.*, 2021).

### 5.5. Contribution of the Study and Demand-Side Governance

This study contributes analytically, empirically, and practically. Analytically, it integrates feasibility, often overlooked in stakeholder-mapping, into the traditional power-interest framework, better reflecting the realities of an SME-dominated sector. Empirically, it provides

one of the first systematic stakeholder maps of the Balkan wine value chain, covering three countries and twelve stakeholder groups. Practically, it clarifies that sustainability is not primarily hindered by opposition but by misalignments between power, motivation, and capacity.

While most sustainability policies in the Balkan wine sector rely on supply-side instruments (regulation, subsidies, certification), the findings suggest potential in demand-side governance. Such instruments could incentivise high-power but low-interest actors like distributors, retailers, and financial institutions by making sustainability economically relevant. Examples include sustainable procurement standards in retail and hospitality, tourism-linked labelling schemes, and integration of sustainability in public catering or wine tourism branding. Evidence from Italy and Spain shows that such approaches can re-engage passive intermediaries and shift market expectations (Flores, 2018; Santini *et al.*, 2013).

## **6. POLICY AND GOVERNANCE IMPLICATIONS**

Sustainability in the Balkan wine sector will not advance through individual initiatives alone, but through coordinated governance that aligns policy frameworks, market incentives, and local realities. While sustainability goals are embedded in EU and national strategies, their translation into practice remains inconsistent because responsibilities are spread across actors who hold unequal power, resources, and motivation. Thus, sustainability is not only a technical challenge but a governance issue involving institutional capacity, power dynamics, and territorial identity.

National ministries and EU institutions define standards, funding mechanisms, and regulatory frameworks, yet these often assume administrative and financial capacities that small wineries do not have. To increase feasibility, support schemes must be simplified, application procedures reduced, and advisory services strengthened. Certification and funding requirements should be adapted to SME realities rather than scaled down from large-producer models. Funding mechanisms such as CAP and IPARD should encourage not only compliance, but also capacity-building, shared infrastructure, and collaborative projects, which are crucial in a fragmented production landscape.

Market intermediaries, distributors, certification bodies, and financial institutions hold significant influence but show limited motivation to promote sustainability unless required by export markets or specific buyers. Governance should therefore focus not on creating new bodies but on reorienting existing ones towards enabling behaviour: introducing sustainability into loan evaluations, certification procedures, and purchasing contracts. Wine associations are well-positioned to coordinate this shift by linking wineries, policymakers, and market actors.

High-interest but low-power actors, local communities, NGOs, tourism operators, and employees, remain largely excluded from formal governance. Involving them through regional wine councils or participatory planning platforms would strengthen local legitimacy and ensure that sustainability reflects territorial identity, not only regulatory compliance. A key policy opportunity lies in demand-side instruments. If retailers, tourism operators, or public institutions reward sustainable wine through procurement standards, labelling schemes, or tourism branding, this could activate currently disengaged actors and make sustainability economically attractive. Rather than relying solely on subsidies and regulation, this approach turns sustainability from a cost into a market advantage.

In summary, effective sustainability governance in the Balkan wine sector depends on three shifts: (1) reducing feasibility barriers for SMEs, (2) aligning powerful but passive market actors with sustainability goals, and (3) integrating local voices into decision-making. These elements form the foundation for the concluding policy roadmap.

## 7. CONCLUSION

Sustainability in the Balkan wine sector is not absent but unevenly applied, constrained by imbalances of power, interest, and feasibility. While embedded in EU and national policies and supported by many wineries and communities, implementation remains hindered by limited administrative capacity, weak market incentives, and fragmented coordination. The stakeholder mapping shows that policy institutions and associations guide sustainability, yet their efforts are not reinforced by distributors and certifiers, while local communities and tourism actors remain marginalised.

The findings underline that progress depends not only on regulations or individual initiatives but on governance mechanisms capable of aligning incentives, reducing feasibility barriers for SMEs, and involving actors who possess local knowledge and long-term commitment to the landscape. This includes simplifying certification and funding procedures, enabling cross-sector collaboration, and recognising sustainability as both an economic strategy and a territorial responsibility. What is learned here for policymakers is clear: sustainability will not be achieved solely through policy documents or funding calls but through adaptive systems that connect institutions, producers, intermediaries, and place-based actors.

For policymakers and associations, three practical lessons emerge. *First*, administrative complexity and certification procedures must be simplified and adapted to SME capacities. *Second*, regional governance platforms, such as cross-sector wine councils involving tourism actors, associations, and local communities, can reduce coordination gaps. *Third*, demand-side tools such as sustainable wine labelling, tourism-linked branding, and green public procurement can motivate currently disengaged intermediaries. Sustainability in the Balkan wine sector will advance not only through regulations but through governance arrangements that align authority, incentives, and feasibility.

## 8. LIMITATIONS AND FUTURE RESEARCH

This study has several limitations that should be acknowledged and addressed in future work. It draws on eight interviews across three countries, providing depth but not representativeness; perspectives from retailers, seasonal workers, or local authorities may be underrepresented. The analysis is primarily qualitative and interpretive, based on interviews and document analysis, and while triangulated with policy evidence, it remains perception-driven. Future research could employ mixed-method approaches, combining stakeholder mapping with surveys, social network analysis, or quantitative indicators, to strengthen robustness. The findings also reflect a specific period during ongoing CAP and Green Deal policy reforms; as funding instruments and regulations evolve, stakeholder priorities and influence may shift, calling for longitudinal studies to trace governance changes over time. Finally, this research focuses on governance dynamics rather than measuring environmental or economic performance; future studies could link stakeholder configurations to tangible sustainability outcomes such as water use, certification uptake, or export results.

## REFERENCES

- Ackermann, F. and Eden, C. (2011), "Strategic management of stakeholders: Theory and practice", *Long Range Planning*, Vol. 44 No. 3, pp. 179-196.
- Alonso, A.D. (2016), "Wineries' contribution to the local community: a stakeholder view", *International Journal of Economics and Business Research*, Vol. 12 No. 4, pp. 295-312.
- Annunziata, E., Pucci, T., Frey, M. and Zanni, L. (2018), "The role of organizational capabilities in attaining corporate sustainability practices and economic performance: Evidence from the Italian wine industry", *Journal of Cleaner Production*, Vol. 171, pp. 1300-1311.
- Baiano, A. (2021), "An overview on sustainability in the wine production chain", *Beverages*, Vol. 7 No. 1, pp. 1-15.
- Bansal, P. (2005), "Evolving sustainably: A longitudinal study of corporate sustainable development", *Strategic Management Journal*, Vol. 26 No. 3, pp. 197-218.
- Broccardo, L. and Zicari, A. (2020), "Sustainability as a driver for value creation: A business model analysis of small and medium enterprises in the Italian wine sector", *Journal of Cleaner Production*, Vol. 259, 120852.
- Corbo, C., Lamastra, L. and Capri, E. (2014), "From environmental to sustainability programs: A review of sustainability initiatives in the Italian wine sector", *Sustainability*, Vol. 6 No. 4, pp. 2133-2159.
- De Steur, H., Wesana, J., Dora, M.K., Pearce, D. and Gellynck, X. (2016), "Applying value stream mapping to reduce food losses and wastes in supply chains: A systematic review", *Waste Management*, Vol. 58, pp. 359-368.
- Dodds, R., Graci, S., Ko, S. and Walker, L. (2013), "What drives environmental sustainability in the New Zealand wine industry? An examination of driving factors and practices", *International Journal of Wine Business Research*, Vol. 25 No. 3, pp. 164-184.
- Dyllick, T. and Hockerts, K. (2002), "Beyond the business case for corporate sustainability", *Business Strategy and the Environment*, Vol. 11 No. 2, pp. 130-141.
- Elkington, J. and Rowlands, I.H. (1999), *Cannibals with Forks: The Triple Bottom Line of 21st Century Business*, Capstone, Oxford.
- Erjavec, E., Volk, T., Rednak, M., Ciaian, P. and Lazdinis, M. (2021), "Agricultural policies and European Union accession processes in the Western Balkans: aspirations versus reality", *Eurasian Geography and Economics*, Vol. 62 No. 1, pp. 46-75.
- European Commission (2023), *CAP Strategic Plans 2023–2027*, European Commission, Brussels.
- Flores, S.S. (2018), "What is sustainability in the wine world? A cross-country analysis of wine sustainability frameworks", *Journal of Cleaner Production*, Vol. 172, pp. 2301-2312.
- Flores, S.S. and Medeiros, M. (2019), "Sustainability challenges in emerging wine regions: A comparative analysis", *Wine Economics and Policy*, Vol. 8 No. 2, pp. 157-168.
- Freeman, R.E. (2010), *Strategic Management: A Stakeholder Approach*, Cambridge University Press, Cambridge.
- Gabzdylova, B., Raffensperger, J.F. and Castka, P. (2009), "Sustainability in the New Zealand wine industry: drivers, stakeholders and practices", *Journal of Cleaner Production*, Vol. 17 No. 11, pp. 992-998.
- Gilinsky, A., Newton, S.K. and Vega, R.F. (2016), "Sustainability in the global wine industry: Concepts and cases", *Agriculture and Agricultural Science Procedia*, Vol. 8, pp. 37-49.
- Licastro, A. and Sergi, B.S. (2021), "Drivers and barriers to a green economy: A review of selected Balkan countries", *Cleaner Engineering and Technology*, Vol. 4, 100228.
- Lozano, R. (2008), "Envisioning sustainability three-dimensionally", *Journal of Cleaner Production*, Vol. 16 No. 17, pp. 1838-1846.
- Mariani, A. and Vastola, A. (2015), "Sustainable winegrowing: current perspectives", *International Journal of Wine Research*, Vol. 7, pp. 37-48.

- Martins, A., Branco, M.C., Melo, P.N. and Machado, C. (2022), “Sustainability in small and medium-sized enterprises: A systematic literature review and future research agenda”, *Sustainability*, Vol. 14 No. 11, 6493.
- Merli, R., Preziosi, M. and Acampora, A. (2018), “Sustainability experiences in the wine sector: toward the development of an international indicators system”, *Journal of Cleaner Production*, Vol. 172, pp. 3791–3805.
- Mitchell, R.K., Agle, B.R. and Wood, D.J. (1997), “Toward a theory of stakeholder identification and salience: defining the principle of who and what really counts”, *Academy of Management Review*, Vol. 22 No. 4, pp. 853-886.
- Monticone, F., Samoggia, A., Specht, K., Schröter, B., Rossi, G., Wissman, A. and Bertazzoli, A. (2024), “Harvesting connections: the role of stakeholders’ network structure, dynamics and actors’ influence in shaping farmers’ markets”, *Agriculture and Human Values*, Vol. 41 No. 4, pp. 1503-1520.
- Moore, S.B. and Manring, S.L. (2009), “Strategy development in small and medium sized enterprises for sustainability and increased value creation”, *Journal of Cleaner Production*, Vol. 17 No. 2, pp. 276-282.
- Mozell, M.R. and Thach, L. (2014), “The impact of climate change on the global wine industry: challenges & solutions”, *Wine Economics and Policy*, Vol. 3 No. 2, pp. 81-89.
- Oltean, F.D. and Gabor, M.R. (2022), “Wine tourism — A sustainable management tool for rural development and vineyards: Cross-cultural analysis”, *Agriculture*, Vol. 12 No. 10, 1614.
- Probst, L., Boyer, M., Baggio, G. and Bonanno, A. (2021), “Participatory governance in agri-food systems: a systematic review”, *Agriculture and Human Values*, Vol. 38, pp. 1025-1043.
- Reed, M.S., Graves, A., Dandy, N., Posthumus, H., Hubacek, K., Morris, J., Prell, C., Quinn, C.H. and Stringer, L.C. (2009), “Who’s in and why? A typology of stakeholder analysis methods for natural resource management”, *Journal of Environmental Management*, Vol. 90 No. 5, pp. 1933-1949.
- Sandroni, L.T., de Barros Ferraz, K.M.P.M., Marchini, S., Percequillo, A.R., Coates, R., Paolino, R.M. et al. (2022), “Stakeholder mapping as a transdisciplinary exercise for jaguar conservation in the Brazilian Atlantic Forest”, *Conservation Science and Practice*, Vol. 4 No. 5, e12651.
- Santini, C., Cavicchi, A. and Casini, L. (2013), “Sustainability in the wine industry: key questions and research trends”, *Agricultural and Food Economics*, Vol. 1 No. 1, pp. 9.
- Schäufele, I. and Hamm, U. (2017), “Consumers’ perceptions and willingness-to-pay for wine with sustainability characteristics: a review”, *Journal of Cleaner Production*, Vol. 147, pp. 379-394.
- Trienekens, J.H. (2011), “Agricultural value chains in developing countries: a framework for analysis”, *International Food and Agribusiness Management Review*, Vol. 14 No. 2, pp. 51-82.
- Van den Akker, A., Fabbri, A., Slater, S., Gilmore, A.B., Knai, C. and Rutter, H. (2024), “Mapping actor networks in global multi-stakeholder initiatives for food system transformation”, *Food Security*, Vol. 16 No. 5, pp. 1223-1234.
- Vukojević, D., Tomić, N., Marković, N., Mašić, B., Banjanin, T., Bodiroga, R. et al. (2022), “Exploring wineries and wine tourism potential in the Republic of Srpska”, *Sustainability*, Vol. 14 No. 5, 2485.

## THE LONG-RUN AVERAGE COST CURVE: EVIDENCE FROM THE BOTTLED WATER INDUSTRY

**Bojan Mladenović**

*Faculty of Economics and Business, University of Belgrade, Serbia  
Business Development Manager, Magroni DOO Skopje, North Macedonia  
bojan.mladenovic@aurora.ekof.bg.ac.rs; mladenovic@magroni.com.mk*

### **ABSTRACT**

*This paper examines the shape of the long-run average cost (LRAC) curve, a central concept in production economics and strategic management. While traditional neoclassical theory suggests a U-shaped curve, with costs declining and then rising due to diseconomies of scale, a significant body of empirical research points toward an L-shaped curve, where costs decline to the minimum efficient scale (MES) and then stabilise. Drawing on both theoretical and empirical perspectives, this paper applies a case study of a bottled water manufacturer in North Macedonia. Using production and cost data collected over a ten-year period, the study tests whether the LRAC follows the U-shaped or L-shaped pattern. The results show that after a phase of declining costs, the firm reached a zone of constant returns to scale, supporting the L-shaped hypothesis. The findings contribute to the debate on cost curve theory and provide managerial implications for capacity planning and efficiency.*

**Keywords:** *Long-run average cost curve, Economies of scale, Bottled water industry, Cost efficiency, Case study*

**JEL classification:** *M41, D24, L67*

### **1. INTRODUCTION**

In economic and business practice, firms continually face the challenge of balancing output expansion and cost efficiency. As a result, the behaviour of average costs has become a central issue in both theory and management. During the early decades of the twentieth century, theoretical constructions established the average cost curve as one of the most frequently used instruments in microeconomic analysis and an indispensable element of the theory of the firm. However, precisely because of its ubiquity, the cost curve has often been applied inconsistently. As a result, interpretations of cost behaviour have often diverged between theoretical models and real-world business practice.

The average total cost (ATC)<sup>1</sup> represents the total cost per unit of output, calculated by dividing total production costs by the quantity produced, and serves as a key indicator for optimal scale and pricing. It provides a measure of efficiency at different levels of output and serves as a key indicator in determining optimal production scale and pricing strategies. In economic theory, the short run refers to a period when some production factors remain fixed, while in the long run, all inputs can be adjusted. This distinction depends not on time itself but on a firm's ability to modify resources and expand capacity as conditions change (Paunović, 2021). In practical terms, this distinction reminds us that short-run decisions are constrained by existing resources, whereas long-run efficiency depends on strategic investment choices.

---

<sup>1</sup> In this paper, the term "average total cost (ATC)" refers to short-run conditions where both fixed and variable costs are included. In the long run, all inputs are variable; hence the term "average cost" is used instead, as the "total" distinction becomes redundant. Both terms denote the average cost per unit of output.

In the short run, the ATC curve in manufacturing companies is almost always presented as U-shaped. This shape suggests that short-run ATC first declines due to the fall in average fixed costs; then, as output increases further, costs stabilise and eventually rise (Larson, 1991). The “U” shape is thus explained by the notion that fixed costs within the production process cause ATC to be high at low output levels, while, on the other hand, there exist certain inputs that cannot be increased indefinitely, at least in the short run. When output is high, the shortage of these inputs constrains the efficiency with which they can contribute to increased production. This naturally raises a question that has shaped decades of debate: What is the true form of the cost curve in the long run, once all inputs become variable? The literature provides two broad perspectives:

- **Traditional perspective.** Rooted in neoclassical economics, this view holds that the long-run average cost (LRAC) curve is also U-shaped, much like the short-run curve. This position is typically associated with Alfred Marshall, who explained increasing and diminishing returns using the concept of the representative firm (Marshall, 1920; Keppler & Lallement, 2006). As production expands, unit costs fall due to economies of scale; eventually, economies are exhausted, diseconomies arise, and unit costs increase.
- **Empirical perspective.** Most empirical studies indicate that the LRAC curve is L-shaped rather than U-shaped. Average costs decline until the minimum efficient scale (MES) is reached and then remain constant, or decline slightly further. According to this perspective, decreasing returns to scale are rarely observed in practice.

This paper addresses the question of LRAC shape through a combination of theory and empirical evidence. After reviewing the key literature, the paper applies a case study of a bottled water producer in North Macedonia, using company data to test whether the LRAC follows the U-shape or L-shape.

## **2. LITERATURE REVIEW**

The precise shape of the LRAC curve has been the subject of economic debate for more than a century. This is understandable given its significance for strategic managerial decisions. While the LRAC has little operational role in day-to-day production management, it is highly relevant in determining the optimal size of production capacity. The distinction lies in the right-hand side of the curve: in the U-shape, costs eventually rise, while in the L-shape, they remain flat or even decline slightly. Given their distinct characteristics, the two shapes lead to fundamentally different managerial insights regarding capacity planning and efficiency improvement. While theory has often assumed a U-shape, empirical evidence increasingly suggests an L-shape. This section reviews the key contributions. Together, these contrasting perspectives underscore how the long-run cost concept has evolved from a theoretical construct into an empirically testable relationship.

### **2.1 The U-shaped LRAC**

The U-shaped LRAC curve is strongly associated with neoclassical economics. It has been used to illustrate economies and diseconomies of scale <sup>2</sup> (Marshall, 1920; Scherer and Ross,

---

<sup>2</sup> The term “economies (and diseconomies) of scale” is used throughout this paper in accordance with standard microeconomic theory, where “scale” refers to changes in output level and associated cost behaviour. In contrast, some authors such as Chambers (1988), Debertin (1986), and Paunović (2021) distinguish between the two concepts, associating “economies of scale” with the technical properties of the production function and “economies of size” with long-run cost behaviour. Others, including Gravel and Rees (2004), Mauris and Thomas

1990). As output expands, average costs decline due to economies of scale; however, beyond a certain point, diseconomies emerge, primarily linked to managerial inefficiencies, leading to rising costs (Coase, 1937; Canback, 2002).

Marshall introduced the concept of the representative firm, arguing that firms initially benefit from internal and external economies of scale but eventually face diminishing returns (Marshall, 1920; Keppeler and Lallemand, 2006). Later, Pigou (1927) elaborated the idea of the equilibrium firm, which reaches an optimal size where further expansion no longer reduces costs. Together, these contributions shaped the traditional U-shaped view of LRAC.

According to select modern literature, the traditional U-shaped LRAC curve remains a valid theoretical framework under specific conditions, though its empirical relevance is limited. Keat, Young, and Erfle describe the LRAC curve as typically U-shaped, reflecting economies of scale at lower output levels and diseconomies of scale at higher levels, and emphasise that the LRAC reaches its lowest point at the MES, beyond which average costs begin to rise (Keat *et al.*, 2017). Pindyck and Rubinfeld (2018) present the U-shaped LRAC in their foundational economic models, highlighting that average costs may eventually rise due to managerial diseconomies and coordination inefficiencies at higher output levels. Similarly, Hubbard,

O'Brien, and Rafferty (2022) acknowledge that the LRAC can be U-shaped in industries where operational complexity and logistical constraints increase with scale. Though empirical studies increasingly suggest flat or L-shaped cost structures, these authors defend the pedagogical and analytical utility of the U-shaped curve in understanding firm cost dynamics, particularly in sectors prone to scale inefficiencies.

## **2.2 The L-shaped LRAC**

Empirical studies increasingly challenge the U-shape. Viner (1931) argued that in the absence of an “absolutely fixed factor” such as land, LRAC tends to flatten after the MES. Implicitly, this reasoning positioned constant average costs as the most common case in industries that did not depend on land or other absolutely fixed factors (Aslanbeigui and Naples, 1997). Scherer and Ross (1990) highlighted managerial and technological innovations that allow firms to sustain constant returns, while Williamson (1970) emphasised the importance of decentralised corporate structures in mitigating diseconomies of scale.

Wiles (1956) and Johnston (1960) found strong evidence that most industries exhibit an L-shaped LRAC, where costs fall and then stabilise. As Johnston observed, production may sometimes generate a U-shaped ATC curve in the short run, since firms attempting to expand output encounter capacity constraints. In the long run, however, enterprises can expand their capacities by constructing new facilities. If each facility operates efficiently, firms can grow as much as they wish without increasing average costs (Besanko *et al.*, 2017). Similarly, Jeganraj (2015) concluded that technological progress and sustainable practices enable firms to maintain costs at minimal levels, consistent with the L-shape hypothesis.

According to modern research, the traditional view of the LRAC curve as strictly U-shaped is increasingly challenged by empirical and theoretical findings. Ghani (2019) argues that the U-shaped LRAC is valid only under restrictive conditions and that actual cost behaviour

---

(2002), Keat *et al.* (2017), Besanko *et al.* (2017), as well as classical and neoclassical economists such as Marshall (1920) and Viner (1931), generally use the term “economies of scale” to describe both phenomena. For the sake of consistency, this paper adopts only the term “of scale” when referring to cost behaviour relative to output volume.

frequently deviates from textbook models. Yinger (2023) supports this re-evaluation, highlighting a shift in economic thought toward recognizing the limitations of earlier assumptions and embracing models that better account for sustained economies of scale.

Lynham (2018) further illustrates that over a significant range of output, average production costs tend to remain stable—indicating a flatter LRAC curve. Taken together, these contributions suggest that in many process industries the LRAC remains flat over wide output ranges, reframing capacity planning as a problem of maintaining utilisation rather than avoiding inevitable diseconomies.

### **3. METHODOLOGY AND CASE STUDY**

This section outlines the methodological approach and presents the case study evidence. The aim is to connect theoretical assumptions about the LRAC curve with the realities of business practice, illustrating how cost dynamics unfold in real production settings. By focusing on one company in the bottled water industry, it becomes possible to observe how production scale, cost dynamics, and efficiency interact in real operating conditions. The methodology combines descriptive cost analysis with a graphical interpretation of both short-run and LRAC curves, allowing a visual insight into how efficiency evolves with scale.

#### **3.1 Research Design**

The research paper applies the case study method to test whether the LRAC curve of a specific manufacturing firm follows a U-shape or an L-shape. Primary data were collected from company records through its Enterprise Resource Planning (ERP) system, covering the period from December 2011 to December 2021. The analysis is based on a real production cost data for a single product. For confidentiality reasons, the product name is omitted, while cost data are presented per unit in aggregated form.

#### **3.2 Company Background**

The case company, located in Skopje, North Macedonia, operates within the bottled water and soft drinks sector, a segment characterised by relatively stable demand and high capital intensity. Founded in 2000, it employs about 130 workers, with an annual turnover of approximately €10 million and total assets of €12 million.

#### **3.3 Data and Variables**

Data were collected from randomly selected production work orders over a ten-year period. The dataset covers production in 3 different scales. For each scale, two variables were measured:

- Q = quantity of output (in number of units)
- ATC = average total cost (€ per single unit)

Table 1 summarises the dataset used in the analysis.

*Table 1: Summary of production volumes and ATC*

Scale 1		Scale 2		Scale 3	
Quantity	ATC	Quantity	ATC	Quantity	ATC
102	0,294 €	9.864	0,159 €	13.430	0,191 €
738	0,194 €	10.080	0,142 €	15.430	0,180 €
1.194	0,175 €	10.326	0,137 €	16.124	0,170 €
2.520	0,172 €	10.350	0,130 €	18.996	0,135 €

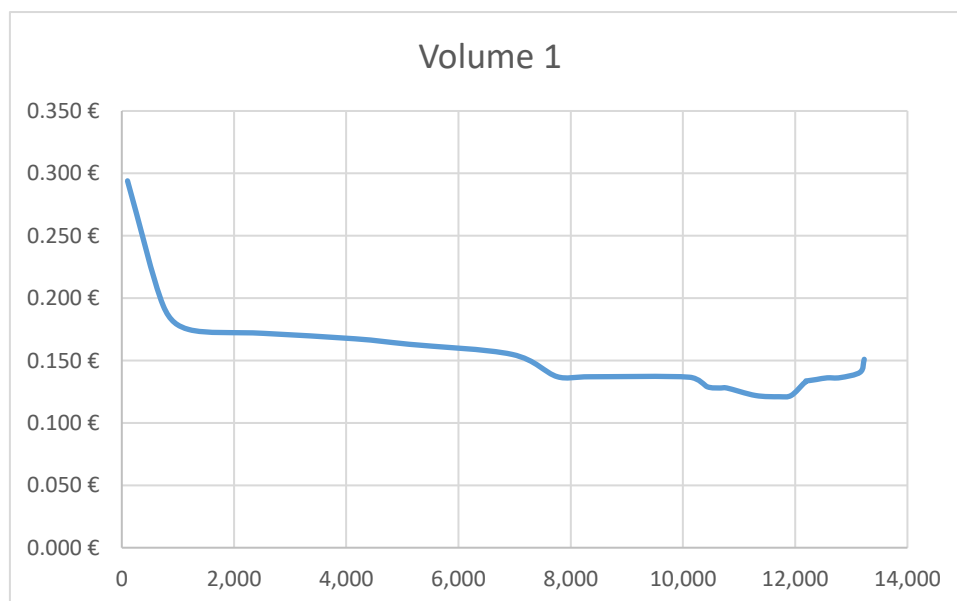
4.158	0,167 €	10.446	0,129 €	18.996	0,135 €
5.118	0,163 €	10.512	0,128 €	21.032	0,132 €
6.954	0,155 €	10.644	0,128 €	26.388	0,130 €
7.764	0,137 €	10.758	0,127 €	26.724	0,126 €
8.340	0,137 €	11.148	0,125 €	29.178	0,121 €
10.116	0,137 €	12.084	0,124 €	29.220	0,118 €
10.432	0,129 €	12.318	0,124 €	29.628	0,110 €
10.656	0,128 €	13.140	0,123 €	33.486	0,109 €
10.782	0,128 €	13.578	0,122 €	34.308	0,104 €
11.292	0,122 €	13.728	0,122 €	35.010	0,103 €
11.712	0,121 €	14.412	0,119 €	35.088	0,102 €
11.928	0,122 €	15.084	0,119 €	40.062	0,101 €
12.150	0,132 €	15.504	0,115 €	40.194	0,100 €
12.180	0,133 €	15.594	0,113 €	43.008	0,100 €
12.192	0,134 €	16.386	0,112 €	44.206	0,097 €
12.210	0,134 €	18.048	0,111 €	44.256	0,097 €
12.384	0,135 €	19.542	0,111 €	44.808	0,097 €
12.570	0,136 €	19.632	0,113 €	50.376	0,096 €
12.804	0,136 €	20.064	0,120 €	52.944	0,093 €
13.158	0,141 €	20.418	0,130 €	53.112	0,092 €
13.230	0,151 €	20.790	0,133 €	59.822	0,091 €

(Source: Author's calculations based on the company's ERP data)

### 3.4 Case Study Findings

- Scale 1. The short-run ATC is U-shaped. Costs fall steeply at low output due to declining fixed costs, then stabilise, and finally rise when output exceeds ~12.000 units.

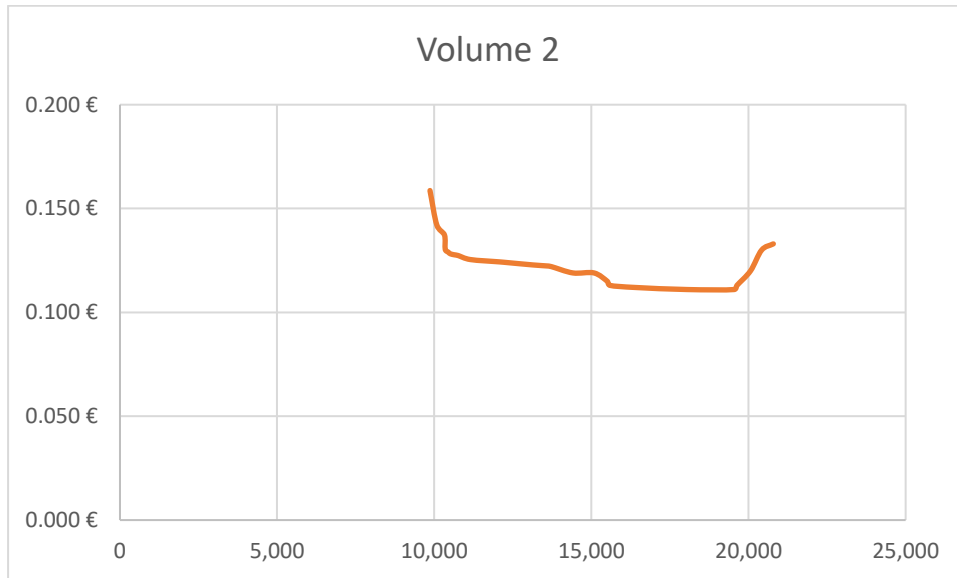
Figure 1: Short-run ATC curve, Scale 1



(Source: Author's own calculations based on collected data)

- Scale 2. The ATC again follows a U-shape. MES is reached at an output of ~18.000 units, while average costs rise beyond ~20.000 units, manifesting diseconomies of scale.

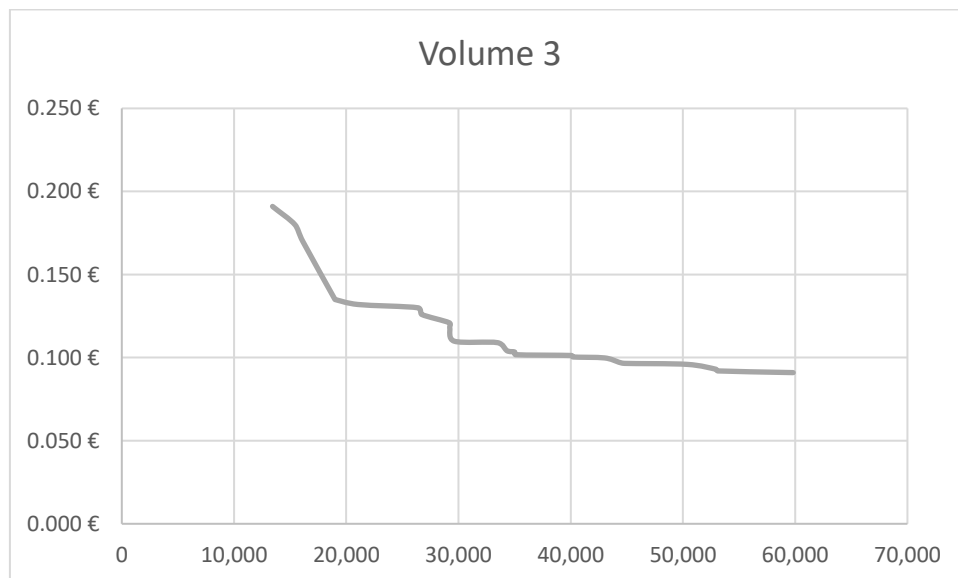
Figure 2: Short-run ATC curve, Scale 2



(Source: Author's own calculations based on collected data)

- Scale 3. The ATC is closer to L-shaped, with no evidence of rising costs at higher volumes. This demonstrates operations under constant returns to scale.

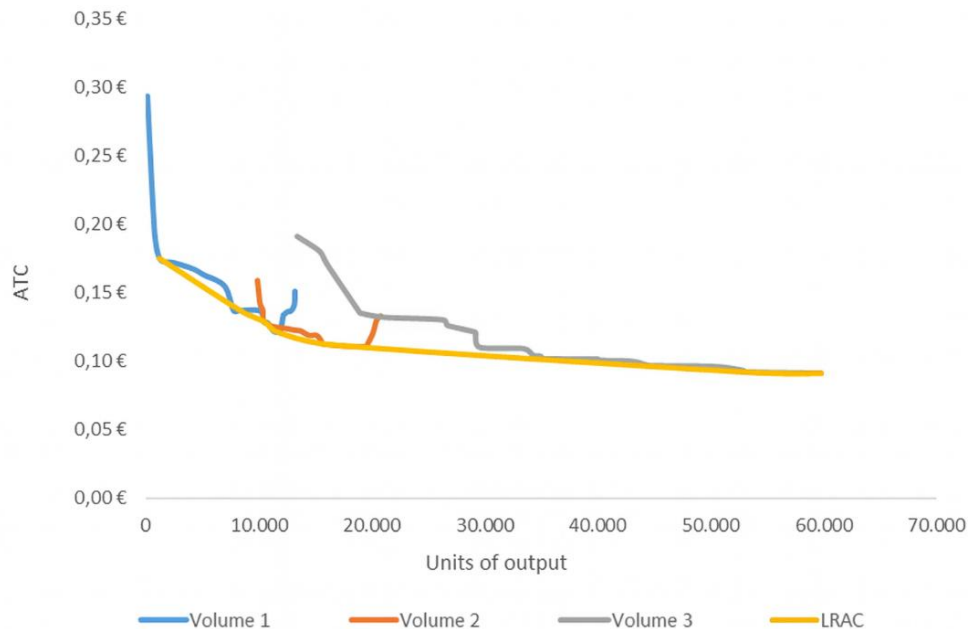
Figure 3: Short-run ATC curve, Scale 3



(Source: Author's own calculations based on collected data)

By integrating the three short-run ATCs, we can obtain the LRAC curve. It shows an initial phase of declining costs, followed by a long zone of constant returns to scale, with no evidence of rising costs at high output. As observed from the combined data, this integration highlights the transition from plant-level efficiency to long-run scale behaviour.

*Figure 4: LRAC curve*



*(Source: Author's own calculations based on collected data)*

### **3.5 Statistical Considerations**

Although this study primarily relies on graphical analysis to examine the relationship between output and ATC, future research could strengthen the empirical validity of these findings through statistical testing. Regression analysis could be applied to measure the significance and stability of cost-output relationships across different production scales. However, given the limited sample size and focus on a single firm, this paper emphasises a qualitative interpretation of the data, which aligns with the exploratory character of the research.

## **4. RESULTS AND DISCUSSION**

The case study supports the L-shaped LRAC. While small-scale capacities showed U-shaped short-run ATC, the LRAC curve stabilised at constant returns. This aligns with empirical findings by Wiles (1956), Johnston (1960), and Viner (1931), who emphasised that in most industries costs do not rise once MES is reached. Managerially, this means scale expansions can proceed without automatic cost penalties provided utilisation discipline is sustained.

The results challenge the neoclassical U-shaped view (Marshall, 1920; Pigou, 1927), but support the modern perspective that technology, organisation, and managerial innovations (Scherer and Ross, 1990; Williamson, 1970) mitigate diseconomies of scale.

From a managerial standpoint, the findings suggest that firms can expand operations without inevitable cost increases—so long as efficiency discipline and proactive capacity planning are maintained.

#### **4.1 Generalizability and Industry Context**

While the case study provides solid, company-level evidence on LRAC behaviour, its generalizability is necessarily limited. The bottled water industry exhibits several structural features that may shape cost dynamics in ways not fully representative of other manufacturing settings: (i) The industry relies heavily on fixed assets such as PET bottling lines, blow-moulding equipment, labelling, clean-in-place (CIP) systems and automatic packaging, which allow fixed costs to be spread over a larger output; (ii) Production is standardised and process-oriented, with relatively few product variations, so set-up and changeover costs although evident, are still lower than in plants producing many different products; (iii) Demand in regional markets is generally stable and predictable, which helps maintain high capacity utilisation; and (iv) The main costs come from packaging materials and energy, offering little flexibility to substitute inputs in the short run. In addition, company-specific factors such as access to technological know-how, maintenance practices, and the maturity of costing systems may further tilt observed average costs toward an L-shaped profile by constraining managerial diseconomies and supporting efficient scale. Consequently, the L-shaped LRAC identified here should be interpreted as context-dependent: comparable patterns are most likely in process industries with high fixed costs, standardised outputs, and sustained utilisation, whereas sectors with greater product heterogeneity, volatile demand, or complex coordination burdens may display different LRAC trajectories. These contextual differences reinforce the importance of analysing cost structures within their specific industrial settings rather than assuming uniform cost behaviour.

#### **4.2 Limitations**

This study has several limitations that should be acknowledged. First, the analysis is based on a single case study within the bottled water industry, and consequently, the cost behaviour observed may reflect the specific technological, organisational, and market characteristics of this firm rather than the entire sector. Second, the study relies primarily on historical production and cost data, which limits the ability to control for external factors such as changes in input prices or demand fluctuations. Third, graphical methods were used to interpret cost trends, without applying more advanced econometric techniques that could statistically confirm the shape of the LRAC curve.

### **5. CONCLUSION**

This paper analysed the LRAC curve through theoretical and empirical perspectives. The study combined a literature-based discussion of cost curve behaviour with a case study from the bottled water industry, where production data from a single firm were used to construct short-run ATC curves and derive the LRAC curve.

Based on the evidence presented, three key conclusions emerge from this analysis, each reflecting a different dimension of the cost–scale relationship. First, while the U-shaped LRAC remains an important theoretical construct, it does not fully align with the empirical evidence obtained in this study. The observed cost pattern better corresponds to the L-shaped curve, characterised by initially declining and subsequently constant average costs. Second, the results support the view that firms can sustain efficiency at optimal scale without triggering diseconomies of scale, especially in industries with standardised production processes and stable demand conditions. Third, the case study demonstrates that industry-specific characteristics—such as capital intensity, process standardisation, and energy-dependent cost

structures—can significantly shape the cost-output relationship, limiting the generalisability of results.

Future research should expand the analysis to a broader sample of firms and employ advanced econometric or artificial intelligence (AI) techniques to assess statistical significance and predict LRAC dynamics. Such approaches could deepen the empirical understanding of cost behaviour and, in practice, strengthen managerial decision-making in capacity planning and cost optimisation. Ultimately, the evidence supports treating LRAC in process industries as flat beyond MES, with managerial focus on sustaining utilisation and reliability rather than anticipating diseconomies by default.

## REFERENCES

- Aslanbeigui, N. and Naples, M.I. (1997), “Scissors or horizon: Neoclassical debates about returns to scale, costs, and long-run supply, 1926-1942”, *Southern Economic Journal*, Vol. 64 No. 2, pp. 517-537. doi:10.2307/1060864
- Besanko, D., Dranove, D., Shanley, M. and Schaefer, S. (2017), *Economics of Strategy*, 7th ed., Hoboken: Wiley.
- Canback, S. (2002), *Bureaucratic Limits of Firm Size: Empirical Analysis Using Transaction Cost Economics*, Doctoral dissertation, Brunel University.
- Chambers, R.G. (1988), *Applied Production Analysis: A Dual Approach*, Cambridge: Cambridge University Press.
- Coase, R.H. (1937), “The nature of the firm”, *Economica*, Vol. 4 No. 16, pp. 386-405. doi:10.1111/j.1468-0335.1937.tb00002.x
- Debertin, D.L. (1986), *Agricultural Production Economics*, New York: Macmillan Publishing Company.
- Ghani, H. (2019), “Different variations of cost curve in economics”, *International Journal of Tax, Economy and Management*, Vol. 2 No. 2.
- Gravell, J. and Rees, R. (2004), *Microeconomics*, 3rd ed., Harlow: Pearson Education.
- Hubbard, R. G., O'Brien, A. P., & Rafferty, M. (2022), *Economics for Managers*, 5th ed., Pearson.
- Jeganraj, M. (2015), “Are long-run average cost curves ever U-shaped? A critical review of literature”, *International Journal of Management and Social Science Research Review*, Vol. 1 No. 8, pp. 138-142.
- Johnston, J. (1960), *Statistical Cost Analysis*, New York: McGraw-Hill.
- Keat, P.G., Young, P.K. and Erfle, S. (2017), *Managerial Economics: Economic Tools for Today's Decision Makers*, 7th ed., Harlow: Pearson.
- Keppler, J.H. and Lallement, J. (2006), “The origins of the U-shaped average cost curve: Understanding the complexities of the modern theory of the firm”, *History of Political Economy*, Vol. 38 No. 4, pp. 733-774. doi:10.1215/00182702-2006-018
- Larson, B. (1991), “A dilemma in the theory of short-run production and cost”, *Southern Economic Journal*, Vol. 58 No. 2, pp. 465-474. doi:10.2307/1060188
- Lynham, J. (2018), 7.3 *The structure of costs in the long run*, In *Principles of microeconomics – Hawaii edition*, University of Hawai'i at Mānoa.
- Marshall, A. (1920), *Principles of Economics*, 8th ed., London: Palgrave Macmillan.
- Mauris, T. and Thomas, C. (2002), *Managerial Economics*, New York: McGraw-Hill
- Paunović, B. (2021), *Ekonomika preduzeća - Preduzeće, okruženje i ulaganja*, CID Ekonomski fakultet, Univerzitet u Beogradu.
- Pigou, A.C. (1927), “The laws of diminishing and increasing cost”, *The Economic Journal*, Vol. 37 No. 146, pp. 188-207. doi:10.2307/2224353

- Pindyck, R. S. and Rubinfeld, D. L. (2018), *Microeconomics*, 9th ed., Pearson.
- Scherer, F.M. and Ross, D. (1990), *Industrial Market Structure and Economic Performance*, Boston: Houghton Mifflin.
- Viner, J. (1931), “Cost curves and supply curves”, *Zeitschrift für Nationalökonomie*, 3, pp. 23-46. doi:10.1007/978-3-662-39842-5\_1
- Wiles, P.J.D. (1956), *Price, Cost, and Output*, New York: Frederick A. Praeger.
- Williamson, O.E. (1970), *Corporate Control and Business Behavior: An Inquiry into the Effects of Organisation Form on Enterprise Behavior*, Englewood Cliffs: Prentice Hall.
- Yinger, J. (2023), “Envelopes for economists: An intellectual history”, *The European Journal of the History of Economic Thought*, Vol. 30 No. 1, pp. 1-27.  
<https://doi.org/10.1080/09672567.2022.2123543>

## SHAPING THE NEW BUSINESS CHANGEMAKERS: TOWARD A NOVEL TAXONOMY OF SOCIAL ENTREPRENEURSHIP EDUCATION

**Bojan Kitanovikj**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
bojan.kitanovikj@eccf.ukim.edu.mk*

**Stojan Debarliev**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
stojan.debarliev@eccf.ukim.edu.mk*

**Aleksandra Janeska-Iliev**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
aleksandra.janeska-iliev@eccf.ukim.edu.mk*

### **ABSTRACT**

*Social entrepreneurship appears as a newer business model that, in addition to profit, aims to solve a social problem with the help of social innovation. As the number of social enterprises increases, so does the interest in acquiring knowledge, skills, and abilities to create a new social entrepreneur who is ready to respond to modern challenges. That is why the eyes are immediately focused on education and educational institutions that are tasked with laying the foundations of learning for social entrepreneurship. To date, the literature on social entrepreneurship education (SEE) is fragmented and without a clear systematic framework. To fill this research gap, the purpose of this paper is to synthesize and map the landscape of SEE and create a new taxonomy of the ecosystem. Namely, this is one of the first research endeavors to map the educational opportunities for learning social entrepreneurship. The research is aimed at examining key trends in SEE, key themes and concepts, their interrelationships, and impact on the social entrepreneurship ecosystem based on abstract mining of scholarly articles published in journals, which are indexed in the Scopus database. The results have implications for a multitude of stakeholders such as academia, managers and entrepreneurs, practitioners, and decision-makers.*

**Keywords:** *Social entrepreneurship, Education, Social entrepreneur, Training, Taxonomy*

**JEL classification:** *L31, I21*

### **1. INTRODUCTION**

The importance of education and training for the development of entrepreneurship is widely recognized (Ndofirepi, 2020). Consequently, entrepreneurship education has attracted increasing interest from private sector stakeholders, higher education institutions, practitioners and policymakers around the world who see this type of educational opportunity as a key driver and enabler of innovative business ideas and economic growth (Nabi *et al.*, 2017; Janeska Iliev *et al.*, 2023). The scholarly discussion on the topic is also developing to examine cross-country and inter-institutional differences in entrepreneurship education models in terms of objectives and pedagogical approaches (Jones and Matlay, 2011). Although there is a lack of an integrated

pedagogical approach to entrepreneurship education for students at all levels (Blankesteyn *et al.*, 2021), there is an increasing body of empirical evidence to support the relevance of entrepreneurship education across disciplines, levels, and types of education (Pollard and Wilson, 2014; Debarliev *et al.*, 2022).

Hence, entrepreneurship education has gradually emerged and developed into a mature field (Jones and Matlay, 2011). On the other hand, in response to the increasing importance of sustainable development, there has recently been an expansion of entrepreneurship education into the domains of sustainable and social entrepreneurship as separate and unique types of entrepreneurial education.

Educational opportunities for social entrepreneurship are generally aimed at building capacity and acquiring knowledge, skills, and experiences related to creating business models that, in addition to profit goals, try to solve a social problem (Shahid and Alarifi, 2021). Additionally, social entrepreneurship education (SEE) is focused on building awareness, shaping attitudes, and expectations about social problems versus sustainable development (Satar and John, 2016), developing relevant mindsets and competencies for starting and managing social enterprises and other sustainable business ventures (Kummitha and Kummitha, 2021).

Consequently, higher education is beginning to contribute to the field of training new social entrepreneurs by offering special programs on social entrepreneurship aimed at business professionals as well as professionals from other educational backgrounds. These programs are often offered by various other institutions around the world, such as the Skoll Center for Social Entrepreneurship, the Ashoka Foundation, the Schwab Foundation, GIZ, and the like (Kremel and Wetter Edman, 2019; Sen, 2007). The increasing number of programs from non-formal education providers and universities shows the strong interest among students and educational institutions in this type of educational opportunity (Kirby and Ibrahim, 2011).

Recently, there has been an intense emergence of academic programs, training and courses on social entrepreneurship in most developed and many developing countries (Brock and Kim, 2011; Nakao and Nishide, 2020). However, there is still room for improvement. Existing curricula and teaching pedagogies for SEE vary widely, resulting in a lack of consensus on what constitutes best practice for this type of education (Pittaway and Cope, 2007).

Researchers note that the design and maintenance process of the programs are primarily driven by teacher preferences (García-González and Ramírez-Montoya, 2021). Previous research efforts have incompletely presented and analyzed the evolutionary development of SEE, offering few relevant findings regarding optimal forms of pedagogy, curriculum design and development, the inclusion of teaching approaches, systemic components, and training needs for social entrepreneurship (Capella-Peris *et al.*, 2020). Consequently, there are limited systematic frameworks for teaching and learning social entrepreneurship around the world (Hoppe *et al.*, 2017).

To address this gap in research for a comprehensive overview of the current situation, opportunities, strategies, and outcomes of SEE from a more systematic and rigorous approach, this research is positioned to provide an answer to these topics. Additionally, after searching the Scopus database, where the highest quality and most cited papers are indexed worldwide, it was noted that no study has used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol for a systematic review and framework building. Hence,

the aim of the research is to synthesize and map the landscape of SEE and create a new taxonomy of the ecosystem.

After defining the purpose, subject and scope of the research, the paper asks the following research questions: 1) What are the dominant themes, clusters, and intellectual structures in the global literature on social entrepreneurship education?; 2) Which recurring dimensions of social entrepreneurship education emerge from the literature and how do they relate with the wider SEE ecosystem?; 3) In what ways do current initiatives interact with formal curricula to expand access, deepen engagement, and support venture creation in SEE?

To answer these questions, the paper presents a systematic literature review of data obtained from the Scopus database, which aims to examine educational opportunities and programs for social entrepreneurship, grounded in the PRISMA protocol for systematic literature reviews and meta-analyses.

Namely, the paper is conceptualized in a way that in the next section lays out the foundations of the theory of social entrepreneurship and educational challenges, followed by a presentation of the methodological approach. Finally, the results and conclusions of the research are presented.

## **2. THEORETICAL BACKGROUND AND RELEVANT RESEARCH**

### **2.1. Achievements in social entrepreneurship education programs**

In practice to date, significant efforts are evident from universities of business, management, and economics, as well as from other academic institutions and organizations for offering training and short-term courses, while undergraduate and postgraduate programs in the field of social entrepreneurship and other related disciplines are less common (Roslan *et al.*, 2022). Even outside the traditional academic world, education for future social entrepreneurs is further strengthened through the establishment of special social entrepreneurship centers around the world. These centers primarily support a wide range of teaching, learning, mentoring, and research initiatives for community development (Rahman and Tekula, 2014). In other words, the focus of social entrepreneurship centers lies in stimulating, incubating, mentoring, and providing financial support to enthusiastic and active social entrepreneurs (Rahman and Tekula, 2014).

One of the most comprehensive analyses of social entrepreneurship curricula was conducted by Brock and Steiner (2009), who examined nearly 110 social entrepreneurship curricula taught at over 70 universities across the United States, Asia, and Europe. In the context of the curriculum, the authors identified seven salient themes in curriculum design, such as addressing social needs or problems, innovating, advancing social enterprises, securing resources, identifying opportunities, developing a sustainable business model, and measuring results (Brock and Steiner, 2009). Furthermore, the study highlighted that universities tend to implement numerous teaching methods, including traditional lectures, class discussions, case study elaborations, guest lectures and interviews, business plan development, and the like (Brock and Steiner, 2009).

However, despite attempts to conceptualize social entrepreneurship education (Mitra *et al.*, 2019), the literature on the topic is fragmented and without clear systematic frameworks of how it functions and, more importantly, how it relates to the broader social entrepreneurship ecosystem. Nevertheless, it is clear that the number of academic and other institutions planning

to initiate various social entrepreneurship education programs is continuously growing (Thomsen *et al.*, 2018).

## **2.2. Social entrepreneurial intention and mindset in the educational context**

Social entrepreneurship represents a paradigm shift by encouraging a bottom-up development approach or a socially innovative business model to address socioeconomic problems (Gupta *et al.*, 2020). As a type of business model, it is the focus of various stakeholders, and therefore understanding and influencing the intention of the social entrepreneur, especially through educational intervention, can greatly help in the process of developing new social entrepreneurs (García-Morales *et al.*, 2020). Researchers believe that the inclusion of social entrepreneurship as part of the curriculum is very likely to encourage the development and replication of social innovations (Thomsen *et al.*, 2018).

There is significant empirical evidence that entrepreneurial intentions and mindsets are fundamental factors influencing entrepreneurial behavior (Farrukh *et al.*, 2018). Social entrepreneurship intention is understood as the desire to create a social enterprise to bring about social change through innovation (Tran and Von Korfflesch, 2016). In other words, intention can be understood as the individual's goal to start an entrepreneurial venture to create social change in society. The positive correlation between entrepreneurial education and entrepreneurial intention has been confirmed in different contexts, including research that targets undergraduate or postgraduate students, given that these respondents tend to think more about their career after education (Chipeta *et al.*, 2020).

Furthermore, research highlights that students should be equipped with a general entrepreneurial mindset as a prerequisite for sustainable development (Pollard and Wilson, 2014). This mindset is growth-oriented and has obvious benefits in terms of improving tolerance of risk, proactivity, innovation, passion for solving social problems, and other elements of entrepreneurial orientation (Satar and Natasha, 2019). Consequently, developing an entrepreneurial mindset can help strengthen the entrepreneurial competencies of young entrepreneurs and students at all levels of education (Schaltegger and Wagner, 2011). Thus, it is necessary for students to be exposed to conditions that stimulate their entrepreneurial behavior for the development of social enterprises (Satar, 2024).

As a result, researchers have often relied on theoretical foundations such as the social capital theory, human capital theory, theory of planned behavior, social cognitive theory, social cognitive career theory, and social learning theory as some of the crucial postulates that describe SEE efforts (Chen and Shabbir, 2025).

## **2.3 Key dimensions of social entrepreneurship education programs**

Diving into the theoretical background of the topic, one can identify several key aspects of social entrepreneurship education programs. Entrepreneurship education generally aims to foster and enhance entrepreneurial awareness, motivation, knowledge, skills, and other competencies necessary to successfully launch and manage a venture (Naveed *et al.*, 2021). By extension, social entrepreneurship education is expected to pursue similar objectives. Ensuring the sustainability of social enterprises demands key entrepreneurial abilities, including the generation and exploitation of ideas, opportunity recognition, securing financial resources, achieving commercial returns for reinvestment, effective management, resource conservation, and an understanding of business models and market dynamics (Shahid and Alarifi, 2021).

Social entrepreneurship education should be designed around broad objectives such as raising awareness, fostering entrepreneurial intentions, building capacity, supporting ongoing ventures, embedding a societal mission within the enterprise vision, and facilitating transitions to employment within social enterprises (Shahid and Alarifi, 2021). Because social enterprises differ in their missions, strategies, structures, and processes (Satar and Natasha, 2019), the purposes and learning goals of these programs must distinguish between narrow and broad conceptualizations of social entrepreneurship. As a result, the learning objectives may vary considerably depending on the form and scope of the educational approach (Kwong *et al.*, 2022).

There is broad agreement that students should be exposed to experiences that stimulate their inclination and intention to act in ways aligned with the values of social entrepreneurship (Mitra *et al.*, 2019). As a multidisciplinary domain, social entrepreneurship encompasses diverse missions, strategies, structures, and processes (Satar and Natasha, 2019). Consequently, it can be integrated into the curricula of various disciplines, including the humanities, arts, sciences, and engineering. This diversity widens the potential audience for such programs to include not only students but also managers, entrepreneurs, activists, and academic professionals (Kwong *et al.*, 2022).

Within the context of SEE, learning centers can play a pivotal role by offering curricula and specialized modules on entrepreneurship and social entrepreneurship. Scholars argue that such centers should be embedded within a broader higher education strategy aimed at systematically developing social entrepreneurship programs, much like traditional entrepreneurship centers (Maas and Jones, 2017). Their organizational positioning, however, is not uniform: some operate within universities or faculties, while others function independently (Shahid and Alarifi, 2021). Ideally, these centers adopt an interdisciplinary orientation and remain accessible across faculties, engaging stakeholders from the wider ecosystem, such as regulators, policymakers, researchers, entrepreneurs, donors, and civil society organizations, to foster a holistic approach to social entrepreneurship education (Maas and Jones, 2017).

### **3. METHODOLOGY**

To fulfill the research objective, we develop the SEE taxonomy based on a systematic literature review and bibliometric analysis techniques, which are instruments for evaluating and analyzing scientific literature and are considered a form of scientific mapping (Zupic and Čater, 2015). Thus, the main goal of bibliometric methods is to discover the relationships and connections between published scientific papers and relevant publications. In addition, there is a descriptive component of this type of research that is used to collect new data about the authors and papers that receive the most citations in the field, their interconnectedness, and the occurrence of keywords and comparable topics (Donthu *et al.*, 2021).

Having said that, bibliometric analysis is not necessarily a new method (Kessler, 1963), but it has recently seen a significant increase in interest from the scientific community due to the easily accessible online databases, which contain almost all documents ever published. This development is supported by new and improved bibliometric software solutions such as VOSviewer and BibExcel, which significantly facilitate data structuring and the analysis process (Bunjak *et al.*, 2022).

For the purposes of this research, a search was conducted in the Scopus database, one of the most influential databases of quality research from around the world. On the other hand, the

fact that the review and the underlying studies it draws upon, rely heavily on literature indexed in Scopus may lead to introducing some bias and omissions due to restricting the source of data to a single database and language. Hence, high-quality research published in other languages or regional journals may not appear in prestige databases and thus may be overlooked, skewing our understanding of SEE. As a result, this is framed as a potential methodological limitation and a call for future research. Still, the data necessary for the bibliometric analysis was obtained by searching for several keywords divided into four pillars:

- social AND entrepren\* AND education\* – 1.597 identified publications,
- “social entrepren\*” AND education\* – 322 identified publications,
- social AND entrepren\* AND training – 548 identified publications,
- “social entrepren\*” AND training – 93 identified publications.

In the search, the year 2025 was excluded because at the time of writing this year has not yet ended. In addition, the intention was to analyze only papers in peer-reviewed scientific journals, which are indexed in the Scopus database. In addition, a pre-selection criterion was set in terms of language, so that the papers had to be written exclusively in English. Furthermore, the search was further refined and limited to the category of business, management and accounting. A more detailed overview of the pre-selection criteria is presented in Table 1.

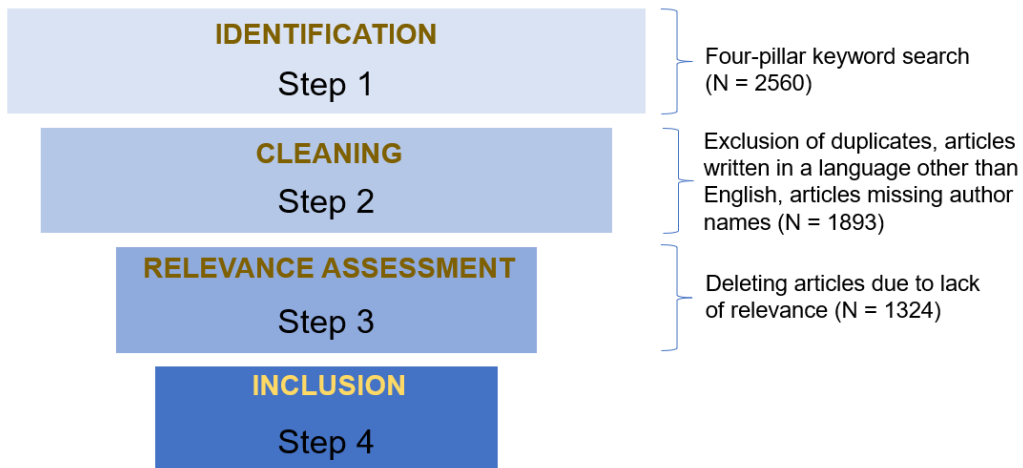
*Table 1: Criteria for inclusion and exclusion of papers from the analysis*

<b>Criteria for excluding papers</b>	<b>Criteria for including papers</b>
1. Comments on other papers, government reports, abstracts and papers presented at conferences, editorials, poster papers, research protocols, gray literature, graduation theses	1. Original scientific papers published in peer-reviewed, international, scientific journals
2. Papers whose primary purpose is not related to the relationship between social entrepreneurship and education	2. Papers explaining the relationship between social entrepreneurship and formal or informal education
3. Papers not written in English	3. Papers written in English
4. Papers published in 2025	4. Papers written up to and including 2024, with no restrictions on initial publication.

*(Source: Authors' own analysis)*

While excluding certain types of publications enhances quality control, it means potentially valuable practical knowledge, like detailed program evaluations or innovations reported in industry reports or doctoral theses, is not incorporated. Publication bias is also a factor, considering that studies showing positive effects of SEE are more likely to be published than null or negative findings, potentially overstating the benefits in the literature. We mitigated bias by following systematic review protocols, employing the PRISMA framework to rigorously identify, screen, and include articles as proposed by Moher *et al.* (2015). Taking this into account, 2,560 articles or papers were identified in the database. A detailed overview of the phases of the PRISMA protocol is shown in Figure 1.

Figure 1: Database cleaning stages according to the PRISMA protocol



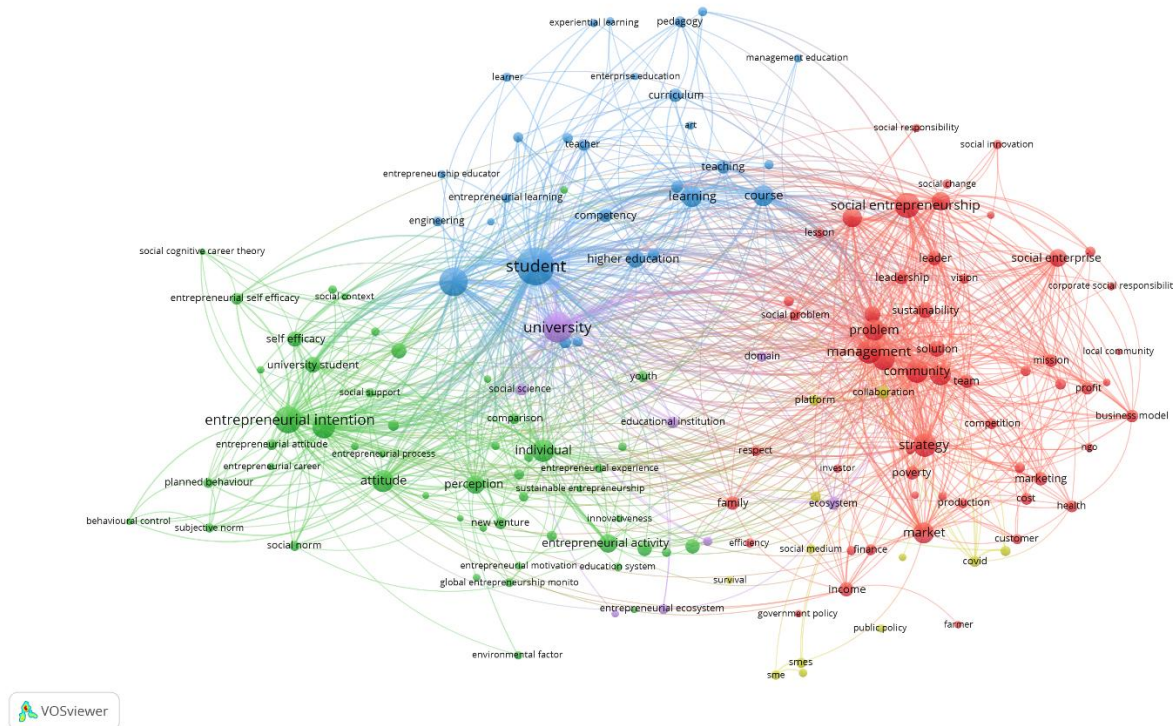
(Source: Adapted from Moher et al. (2015))

For the purposes of bibliometric analysis, the VOSviewer software (version 1.6.19) was used to measure and display the frequency of occurrence of keywords (Van Eck and Waltman, 2014).

#### 4. RESEARCH FINDINGS AND DISCUSSION

To develop a taxonomy of SEE, a keyword co-occurrence analysis was performed to gauge the main topics, areas and themes in the research field (Figure 2).

Figure 2: Network visualization map of SEE keyword co-occurrence



(Source: Authors' own analysis)

Namely, the map shows a dense network of keywords related to social entrepreneurship and education, where the size of each keyword or bubble reflects how often the keyword appears in the literature and the colors indicate groups of terms that frequently co-occur (Van Eck and Waltman, 2014). The thick links between the keywords represent strong relationships in the published work.

One prominent cluster appears in *blue* and centers on terms such as “student,” “entrepreneurship education,” “higher education,” “learning,” “curriculum,” “teaching,” and “experiential learning”. This area captures the pedagogical and institutional side of entrepreneurship education, pointing out research on how universities design curricula, how teachers deliver courses, and how students experience entrepreneurship and social innovation in higher-education settings. Moreover, it is focused more on the individual aspects of gaining competencies and delivering knowledge. The cluster’s main theoretical move is adapting experiential or constructivist learning and “authentic” entrepreneurship education to the social-impact context, arguing that real-problem engagement and reflection are necessary to build changemaker competence (Aadland and Aaboen, 2020). Practically, studies converge on two levers of effectiveness – institutional commitment through mission alignment and dedicated centers or incubators, on one hand, and educator capacity through faculty development and practitioner co-teaching on the other (Vázquez-Parra *et al.*, 2022). Yet evidence comparing instructional designs remains thin, and assessment standards are uneven (Chen and Shabbir, 2025).

A second major cluster is shown in *red* and focuses on the essence, context, and content of SEE. Some of the most notable keywords here include “social entrepreneurship,” “management,” “community,” “leadership,” “sustainability,” “strategy”, and “problem solving”, representing the management and practice dimension of the field. The keywords here emphasize how social enterprises create value, the strategies they use to address social problems, and the importance of collaboration, leadership, and community engagement in achieving social change. The contribution is a competence view that integrates business intelligence with mission-driven and systems-thinking skills. Curricula increasingly include design thinking, stakeholder engagement, and impact logic, but comparative studies of syllabi and learning outcomes are still scarce (Azqueta *et al.*, 2023).

The *green* cluster brings together terms such as “entrepreneurial intention,” “attitude,” “self-efficacy,” “perception,” and “social norm”. This points to a psychological and behavioral research stream, where scholars examine how education and social context shape students’ entrepreneurial attitudes and their intention to start ventures as outcomes of the process. It links theories of planned behavior and social cognitive career development to entrepreneurship education. Empirical work generally shows positive effects, especially for experiential formats, yet effects vary across contexts and program quality, joining the call for linking intentions to action and track post-program behaviors (Nabi *et al.*, 2017). This means that in the SEE context, intentions are still seen as outcomes of educational activities and antecedents for becoming a social entrepreneur (Michel and Förster, 2025).

A smaller *yellow* cluster appears around keywords such as “policy,” “SMEs,” “market,” “finance,” “income,” and “COVID”. As such, the cluster reflects work on the external environment in the form of public policy, market forces, funding mechanisms and crises, which influence social entrepreneurship and entrepreneurship education. In comparison with the other clusters, it shows an emerging but still less developed body of research together with the *purple* cluster, which contains terms like “ecosystem,” “domain,” “university,” “social science,” and

similar. In other words, these emergent themes situate SEE within enabling or constraining environments in terms of policy, finance, partnerships, and shocks such as COVID, arguing that program outcomes depend on ecosystem scaffolding or the abundance of incubators, funding, or multi-stakeholder projects (Pache and Chowdhury, 2012). While universities with social-innovation hubs and wide and well-established social networks report stronger venture outputs (Hassan *et al.*, 2022), sustainable financing and long-term models remain under-researched, especially in fragile contexts (Shahid and Alarifi, 2021; Albatran and Atikbay, 2025).

In this vein, despite rapid growth, SEE literature remains fragmented and marked by several gaps and tensions. First, scholars note an impact-intention gap where research predominantly measures shifts in student intentions or attitudes rather than long-term social venture creation or demonstrable societal impact (Azqueta *et al.*, 2023; Chen and Shabbir, 2025). While SEE is shown to boost social entrepreneurial intentions and self-efficacy, empirical evidence on the translation of these intentions into sustained entrepreneurial action and social value creation remains limited (Alzate *et al.*, 2024). This calls for longitudinal designs and outcome tracking beyond course completion.

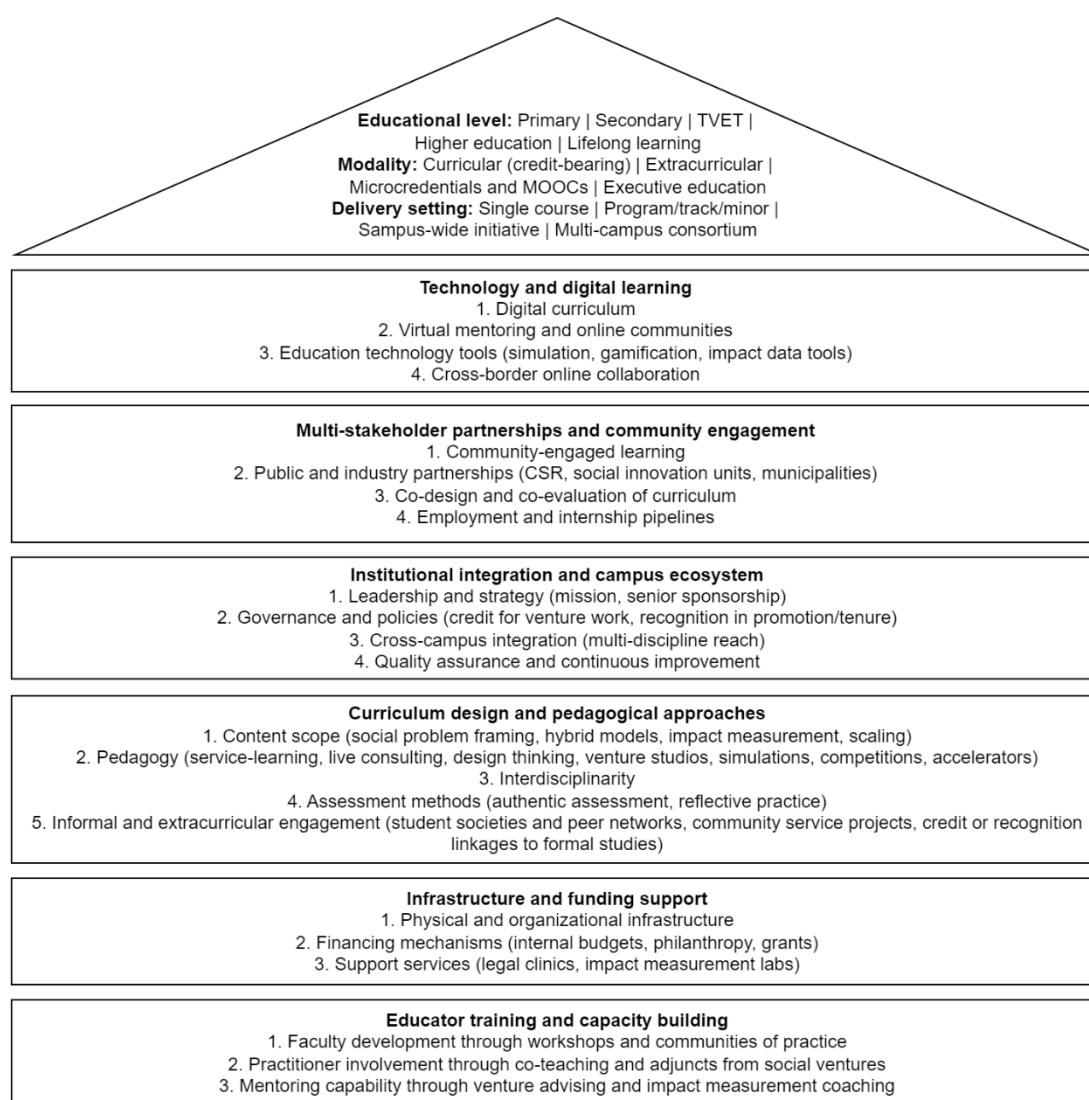
Second, SEE exhibits pedagogical inconsistency. Programs vary widely in methods – from lectures to service learning and incubator-based models (Roslan *et al.*, 2022). Yet systematic comparisons and pedagogical evidence remain scarce (Chen and Shabbir, 2025). Although experiential and authentic learning approaches show promise, the field lacks consensus on core competences, instructional design, and assessment standards (Hockerts, 2018; Capella-Peris *et al.*, 2020). This variability risks impeding cumulative knowledge development and complicates efforts to benchmark program effectiveness.

Third, SEE faces conceptual and interdisciplinary tensions. Social entrepreneurship combines business, social innovation, and public-sector logics, yet educational models often privilege either managerial or social-mission perspectives (Hoppe *et al.*, 2017). Scholars highlight the need for clearer conceptual boundaries and integrated competence frameworks that reconcile economic sustainability with social impact and stakeholder value creation (Satar, 2024; Gupta *et al.*, 2020).

Finally, the literature increasingly recognizes the underexplored role of informal learning ecosystems, including accelerators, community-based programs, and online platforms (Debarliev *et al.*, 2022; Hockerts, 2018). Such initiatives often foster deeper engagement and real-world practice, however, they remain peripheral in research and loosely integrated into formal programs (Vázquez-Parra *et al.*, 2022). Thus, we believe that bridging formal and informal learning pathways and examining their combined effects represents a promising avenue for advancing SEE theory and practice.

With that being said, much of the literature measures student intentions or competencies, while giving less attention to the structural and programmatic mechanisms that make SEE effective (Shahid and Alarifi, 2021; Chen and Shabbir, 2025). Across levels of education, especially higher education, the field now calls for integrative frameworks that explain how curricula, institutions, and ecosystems interact to nurture social entrepreneurs (Alzate *et al.*, 2024). We propose one such taxonomy in Figure 3.

Figure 3: Taxonomy of SEE ecosystem



(Source: Authors' own construction)

Multiple reviews identify a shortage of trained educators capable of delivering interdisciplinary, impact-oriented instruction (Roslan *et al.*, 2022). Faculty development programs, practitioner co-teaching, and communities of practice are emerging solutions, but empirical evidence on how these capacity-building efforts influence program quality remains limited (Alzate *et al.*, 2024).

Multi-stakeholder partnerships and infrastructure support emerge as powerful enablers in SEE initiatives, too. Collaborations with nonprofits, social enterprises, government agencies, and industry provide real problem contexts, mentorship, and funding (Pache and Chowdhury, 2012; Roslan *et al.*, 2022). Universities with incubators, social innovation hubs, and dedicated funding lines demonstrate stronger outcomes, yet financing models and long-term sustainability of such infrastructure are under-researched (Shahid and Alarifi, 2021).

At the program level, curriculum design and pedagogy remain diverse and inconsistent. Effective practices combine social venture content with experiential methods in the form of service learning, live consulting projects, and design thinking (Hockerts, 2018). Scholars agree on the importance of learner-centered, action-oriented approaches, but systematic comparisons



actors in controlled settings), and finally to real contexts where students operate genuine social enterprises with external stakeholders and tangible risk.

As in the original model, these categories are hierarchical, as higher levels of self-direction presuppose elements of the preceding stages. For example, a self-directed, real-context program still contains educator support and participatory elements to ensure students acquire the necessary theoretical grounding before assuming full responsibility.

## **5. CONCLUSION**

This study provides one of the first systematic attempts to map and classify the rapidly expanding field of SEE. By applying the PRISMA protocol and bibliometric techniques to a large Scopus dataset, we identify four dominant research clusters – pedagogical and institutional design, social entrepreneurship practice, entrepreneurial intention and psychology, and policy/market context. We synthesized them into a new taxonomy of the SEE ecosystem. Our findings show that while SEE is gaining momentum across higher education and non-formal settings, the literature remains fragmented and heavily outcome-oriented, focusing primarily on student intentions and competences rather than on the programmatic and structural mechanisms that make SEE effective (Shahid and Alarifi, 2021; Chen and Shabbir, 2025). The proposed taxonomy highlights seven interrelated domains that together capture the complexity of how education and social entrepreneurship intersect.

Some key takeaways can be found useful for both theory and practice. First, universities that embed social entrepreneurship within their strategic missions and create dedicated centers or hubs are better positioned to sustain SEE at scale (Vázquez-Parra *et al.*, 2022). Second, effective curricula blend social venture content with experiential pedagogies such as service learning and live consulting projects, yet systematic comparisons of pedagogical approaches remain rare (Hockerts, 2018). Third, a shortage of trained educators continues to constrain program quality, underlining the importance of faculty development, practitioner co-teaching, and communities of practice (Roslan *et al.*, 2022). Finally, partnerships with external stakeholders and the emergence of digital platforms offer powerful but under-studied avenues for extending SEE beyond traditional classroom boundaries.

Despite these contributions, the study does not come free of some limitations. Bibliometric methods rely on publication data indexed in Scopus and therefore may omit relevant work in other languages, databases or practitioner outlets (Donthu *et al.*, 2021). The taxonomy is conceptual and derived from secondary data, so future empirical testing across different national and institutional contexts is needed to validate and refine its categories. These limitations open several avenues for future research. Comparative case studies and longitudinal designs could examine how combinations of curriculum, partnerships, and infrastructure shape student outcomes and venture creation. Cross-country analyses would help uncover contextual differences in policy support, funding models, and cultural factors influencing SEE. Additionally, the aforementioned digital learning environments and hybrid delivery models merit particular attention, given their growing role in scaling SEE and fostering global networks.

Theoretically, this taxonomy extends work on authentic entrepreneurship education and foundational theories like social learning theory, human capital theory, social cognitive theory, and similar by reconceptualizing the hierarchical dimensions for a social impact context, thus linking levels of educational conception with degrees of real-world engagement. Practically, it

offers educators and policymakers a structured framework for designing and evaluating SEE programs, guiding decisions about institutional investment, partnership development, and the integration of experiential and digital elements. By moving the conversation beyond outcomes to the mechanisms of partnership and implementation, this research contributes to the maturation of SEE as a distinct field that climbs high on global agendas when it comes to research and education, in general.

## REFERENCES

- Aadland, T. and Aaboen, L. (2020), "An entrepreneurship education taxonomy based on authenticity", *European Journal of Engineering Education*, Vol. 45 No. 5, pp. 711-728.
- Albatran, A.A. and Atikbay, T. (2025), "Entrepreneurship education in fragile contexts: bridging the intention–action gap through psychological and contextual pathways", *Sustainability*, Vol. 17 No. 16, 7447.
- Alzate, P., Mejía-Giraldo, J.F., Jurado, I., Hernandez, S. and Novozhenina, A. (2024), "Research perspectives on youth social entrepreneurship: strategies, economy, and innovation", *Journal of Innovation and Entrepreneurship*, Vol. 13 No. 1, p. 49.
- Azqueta, A., Sanz-Ponce, R. and Núñez-Canal, M. (2023), "Trends and opportunities in social entrepreneurship education research", *Administrative Sciences*, Vol. 13 No. 11, p. 232.
- Blankesteyn, M., Bossink, B. and van der Sijde, P. (2021), "Science-based entrepreneurship education as a means for university-industry technology transfer", *International Entrepreneurship and Management Journal*, Vol. 17 No. 2, pp. 779-808.
- Brock, D.D. and Kim, M. (2011), *Social entrepreneurship education resource handbook*, Available at SSRN 1872088.
- Brock, D.D. and Steiner, S. (2009), *Social entrepreneurship education: is it achieving the desired aims?*, Available at SSRN 1344419.
- Bunjak, A., Černe, M. and Schölly, E.L. (2022), "Exploring the past, present, and future of the mindfulness field: a multitechnique bibliometric review", *Frontiers in Psychology*, Vol. 13, 792599.
- Capella-Peris, C., Gil-Gómez, J., Martí-Puig, M. and Ruíz-Bernardo, P. (2020), "Development and validation of a scale to assess social entrepreneurship competency in higher education", *Journal of Social Entrepreneurship*, Vol. 11 No. 1, pp. 23-39.
- Chen, N. and Shabbir, M.S. (2025), "Social entrepreneurship education landscape mapping: a bibliometric analysis and analysis of empirical research", *The International Journal of Management Education*, Vol. 23 No. 2, 101122.
- Chipeta, E.M., Kruse, P. and Surujlal, J. (2020), "Effects of gender on antecedents to social entrepreneurship among university students in South Africa", *International Journal of Business and Management Studies*, Vol. 12 No. 1, pp. 18-33.
- Debarliev, S., Janeska-Iliev, A., Stripeikis, O. and Zupan, B. (2022), "What can education bring to entrepreneurship? Formal versus non-formal education", *Journal of Small Business Management*, Vol. 60 No. 1, pp. 219-252.
- Dixon, S.E. and Clifford, A. (2007), "Ecopreneurship – a new approach to managing the triple bottom line", *Journal of Organizational Change Management*, Vol. 20 No. 3, pp. 326-345.
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. and Lim, W.M. (2021), "How to conduct a bibliometric analysis: an overview and guidelines", *Journal of Business Research*, Vol. 133, pp. 285-296.
- Estrin, S., Mickiewicz, T. and Stephan, U. (2016), "Human capital in social and commercial entrepreneurship", *Journal of Business Venturing*, Vol. 31 No. 4, pp. 449-467.

- Farrukh, M., Alzubi, Y., Shahzad, I.A., Waheed, A. and Kanwal, N. (2018), "Entrepreneurial intentions: the role of personality traits in perspective of theory of planned behaviour", *Asia Pacific Journal of Innovation and Entrepreneurship*, Vol. 12 No. 3, pp. 399-414.
- García-González, A. and Ramírez-Montoya, M.S. (2021), "Social entrepreneurship education: changemaker training at the university", *Higher Education, Skills and Work-Based Learning*, Vol. 11 No. 5, pp. 1236-1251.
- García-Morales, V.J., Martín-Rojas, R. and Garde-Sánchez, R. (2020), "How to encourage social entrepreneurship action? Using Web 2.0 technologies in higher education institutions", *Journal of Business Ethics*, Vol. 161, pp. 329-350.
- Gupta, P., Chauhan, S., Paul, J. and Jaiswal, M.P. (2020), "Social entrepreneurship research: a review and future research agenda", *Journal of Business Research*, Vol. 113, pp. 209-229.
- Hassan, H.K., Igel, B. and Shamsuddoha, M. (2022), "Entrepreneurship education and social entrepreneurial intentions: the mediating effects of entrepreneurial social network", *Frontiers in Psychology*, Vol. 13, 860273.
- Hockerts, K. (2018), "The effect of experiential social entrepreneurship education on intention formation in students", *Journal of Social Entrepreneurship*, Vol. 9 No. 3, pp. 234-256.
- Hoppe, M., Westerberg, M. and Leffler, E. (2017), "Educational approaches to entrepreneurship in higher education: A view from the Swedish horizon", *Education + Training*, Vol. 59 No. 7/8, pp. 751-767.
- Janeska Iliev, A., Debarliev, S. and Cvetkoska, V. (2023), "What is next in entrepreneurial education: bibliometric analysis on building competencies for the entrepreneurial world", in *Proceedings of the 4th International Conference Economic and Business Trends Shaping the Future in Skopje, North Macedonia, 2023*, pp. 177-184.
- Jones, C. and Matlay, H. (2011), "Understanding the heterogeneity of entrepreneurship education: going beyond Gartner", *Education + Training*, Vol. 53 No. 8/9, pp. 692-703.
- Kessler, M.M. (1963), "Bibliographic coupling between scientific papers", *American Documentation*, Vol. 14 No. 1, pp. 10-25.
- Kirby, D.A. and Ibrahim, N. (2011), "The case for (social) entrepreneurship education in Egyptian universities", *Education + Training*, Vol. 53 No. 5, pp. 403-415.
- Koe Hwee Nga, J. and Shamuganathan, G. (2010), "The influence of personality traits and demographic factors on social entrepreneurship start up intentions", *Journal of Business Ethics*, Vol. 95, pp. 259-282.
- Kremel, A. and Wetter Edman, K. (2019), "Implementing design thinking as didactic method in entrepreneurship education. The importance of through", *The Design Journal*, Vol. 22 Suppl. 1, pp. 163-175.
- Kummitha, H.R. and Kummitha, R.K.R. (2021), "Sustainable entrepreneurship training: a study of motivational factors", *The International Journal of Management Education*, Vol. 19 No. 1, 100449.
- Kwong, C., Cheung, C., Bhattarai, C. and Fieldhouse, S. (2022), "How entrepreneurial are social entrepreneurship education providers? The role of universities' entrepreneurial ecosystems in the provision of elective social entrepreneurship courses to business students", *Studies in Higher Education*, Vol. 47 No. 5, pp. 1046-1056.
- Maas, G. and Jones, P. (Eds.). (2017), *Entrepreneurship centres: Global perspectives on their contributions to higher education institutions*. Springer.
- Michel, S. and Förster, M. (2025), "Challenge-based learning for social entrepreneurship education in vocational schools—a promising learning environment to foster entrepreneurial intentions?!", *Empirical Research in Vocational Education and Training*, Vol. 17 No. 1, p. 21.
- Mitra, P., Kickul, J., Gundry, L. and Orr, J. (2019), "The rise of hybrids: a note for social entrepreneurship educators", *International Review of Entrepreneurship*, Vol. 17 No. 2.

- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M. and Prisma-P Group (2015), "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement", *Systematic Reviews*, Vol. 4, pp. 1-9.
- Nabi, G., Liñán, F., Fayolle, A., Krueger, N. and Walmsley, A. (2017), "The impact of entrepreneurship education in higher education: a systematic review and research agenda", *Academy of Management Learning and Education*, Vol. 16 No. 2, pp. 277-299.
- Nakao, K. and Nishide, Y. (2020), "The development of social entrepreneurship education in Japan", *Entrepreneurship Education*, Vol. 3, pp. 95-117.
- Naveed, M., Zia, M.Q., Younis, S. and Shah, Z.A. (2021), "Relationship of individual social entrepreneurial orientations and intentions: role of social entrepreneurship education", *Asia Pacific Journal of Innovation and Entrepreneurship*, Vol. 15 No. 1, pp. 39-50.
- Ndofirepi, T.M. (2020), "Relationship between entrepreneurship education and entrepreneurial goal intentions: psychological traits as mediators", *Journal of Innovation and Entrepreneurship*, Vol. 9 No. 1, pp. 1-20.
- Pache, A.C. and Chowdhury, I. (2012), "Social entrepreneurs as institutionally embedded entrepreneurs: toward a new model of social entrepreneurship education", *Academy of Management Learning and Education*, Vol. 11 No. 3, pp. 494-510.
- Pittaway, L. and Cope, J. (2007), "Simulating entrepreneurial learning: integrating experiential and collaborative approaches to learning", *Management Learning*, Vol. 38 No. 2, pp. 211-233.
- Pless, N.M., Maak, T. and Stahl, G.K. (2011), "Developing responsible global leaders through international service-learning programs: the Ulysses experience", *Academy of Management Learning and Education*, Vol. 10 No. 2, pp. 237-260.
- Pollard, V. and Wilson, E. (2014), "The 'entrepreneurial mindset' in creative and performing arts higher education in Australia", *Artivate*, Vol. 3 No. 1, pp. 3-22.
- Rahman, N. and Tekula, R. (2014), "A replicable evaluation method of social entrepreneurship centers and programs", *Emerging Research Directions in Social Entrepreneurship*, pp. 111-123.
- Roslan, M.H.H., Hamid, S., Ijab, M.T., Yusop, F.D. and Norman, A.A. (2022), "Social entrepreneurship in higher education: challenges and opportunities", *Asia Pacific Journal of Education*, Vol. 42 No. 3, pp. 588-604.
- Satar, M.S. (2024), "Inspiring role model and compassion in nascent social entrepreneurs: does education matter?", *Entrepreneurship and Regional Development*, pp. 1-25.
- Satar, M.S. and John, S. (2016), "A conceptual model of critical success factors for Indian social enterprises", *World Journal of Entrepreneurship, Management and Sustainable Development*, Vol. 12 No. 2.
- Satar, M.S. and Natasha, S. (2019), "Individual social entrepreneurship orientation: towards development of a measurement scale", *Asia Pacific Journal of Innovation and Entrepreneurship*, Vol. 13 No. 1, pp. 49-72.
- Schaltegger, S. and Wagner, M. (2011), "Sustainable entrepreneurship and sustainability innovation: categories and interactions", *Business Strategy and the Environment*, Vol. 20 No. 4, pp. 222-237.
- Sen, P. (2007), "Ashoka's big idea: transforming the world through social entrepreneurship", *Futures*, Vol. 39 No. 5, pp. 534-553.
- Shahid, S.M. and Alarifi, G. (2021), "Social entrepreneurship education: a conceptual framework and review", *The International Journal of Management Education*, Vol. 19 No. 3, 100533.
- Smith, W.K., Besharov, M.L., Wessels, A.K. and Chertok, M. (2012), "A paradoxical leadership model for social entrepreneurs: challenges, leadership skills, and pedagogical

- tools for managing social and commercial demands", *Academy of Management Learning and Education*, Vol. 11 No. 3, pp. 463-478.
- Thomsen, B., Muurlink, O. and Best, T. (2018), "The political ecology of university-based social entrepreneurship ecosystems", *Journal of Enterprising Communities: People and Places in the Global Economy*, Vol. 12 No. 2, pp. 199-219.
- Tran, A.T. and Von Korflesch, H. (2016), "A conceptual model of social entrepreneurial intention based on the social cognitive career theory", *Asia Pacific Journal of Innovation and Entrepreneurship*, Vol. 10 No. 1, pp. 17-38.
- Van Eck, N.J. and Waltman, L. (2014), "Visualizing bibliometric networks", in *Measuring Scholarly Impact: Methods and Practice*, Springer International Publishing, Cham, pp. 285-320.
- Vázquez-Parra, J.C., Amézquita-Zamora, J.A. and Ramírez-Montoya, M.S. (2022), "Student perception of their knowledge of social entrepreneurship: gender gap and disciplinary analysis of an Ashoka Changemaker campus in Latin America", *Journal of Applied Research in Higher Education*, Vol. 14 No. 3, pp. 1224-1241.
- Zupic, I. and Čater, T. (2015), "Bibliometric methods in management and organization", *Organizational Research Methods*, Vol. 18 No. 3, pp. 429-472.

## **PANEL DESCRIPTION**

### **“Public Opinion 2.0: Challenges in the Digital Age and Possible Research Directions”**

This section of contributions belongs to the "Public Opinion 2.0: Challenges in the Digital Age and Possible Research Directions" panel. The speakers note that the concept of public opinion is undergoing a fundamental transformation. In an era marked by challenges such as digital transformation, social polarization, and the diminishing role of traditional democratic institutions and media as intermediaries in public debate, questions arise regarding the methods and necessary future directions for public opinion research. Public Opinion 2.0 is forming within a new digital environment characterized by decentralized content production, non-linear information circulation, and the growing influence of digital platforms. In the Public Opinion 2.0 era, the speed of information dissemination facilitates disinformation and infodemics, which significantly impact democratic processes. The de-institutionalization of public debate mediators (e.g., media, political parties) coincides with discussions about new forms of exclusion, public opinion manipulation, and the real sources of political decisions. Public debate is becoming platformized and algorithmized, affecting the possibility of participation in public discourse - does this foster deliberation, polarization, radicalization, or perhaps "algocracy"? Comprehensive analyses of Public Opinion 2.0 necessitate a cross-interdisciplinary approach that integrates a media studies perspective (media, technology, discourse) with a political science perspective (power, democracy, institutions). Understanding the mechanisms of public opinion formation amidst digital transformation is crucial not only for political and social sciences but also for the future of business environments and economies. Public opinion, especially in the era of network communication and digital media, directly impacts consumer decisions, trust in institutions, and the investment and regulatory climate.

**RETHINKING PUBLIC OPINION STUDIES IN THE DIGITAL ERA: EXPERT DISCOURSE ON THE THEORETICAL AND METHODOLOGICAL CHALLENGES OF PUBLIC OPINION RESEARCH**

**Elena Negrea-Busuioac**

*National University of Political Studies and Public Administration,  
Bucharest, Romania  
elena.negrea@comunicare.ro*

**Agnieszka Hess**

*Jagiellonian University, Krakow, Poland  
agnieszka.hess@uj.edu.pl*

**Agnieszka Stępińska**

*Adam Mickiewicz University, Poznan, Poland  
agnieszka.stepinska@amu.edu.pl*

**Anita Ciunova-Shuleska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
anita.ciunova@eccf.ukim.edu.mk*

**Anna Bączkowska**

*University of Gdansk, Poland  
anna.k.baczkowska@gmail.com*

**EXTENDED ABSTRACT**

**Purpose** Social media platforms provide the most dynamic environment for creating, disseminating, and sharing political information and opinions (Kraft et al., 2020). In addition, they provide fertile ground for increasingly diverse and multidimensional forms of public opinion formation, shaped by various leaders, including those who are not human. The impact of algorithmization and the use of artificial intelligence on shaping the political attitudes of internet users forces scholars to redefine and recategorize public opinion (Gandini et al., 2025). Furthermore, researchers need to resolve the problem of measuring public opinion in the era of digital communication (Baden *et al.*, 2020).

In this paper, we aim to address the theoretical and methodological challenges that scholars studying public opinion face today. Our analysis was guided by the following two research questions: (RQ1) What challenges do experts encounter when conducting research on public opinion in the digital age? (RQ2) How do they envision the future of the field in the evolving landscape?

**Design/methodology/approach** This study employs a survey-based methodology to investigate the experiences and perspectives of public opinion scholars in the context of transformations related to social media and artificial intelligence. A questionnaire was distributed among scholars from the COST OPINION network and respondents recommended by the network through the snowball method. The survey yielded 86 completed questionnaires from experts across 33 countries, representing various disciplines (linguistics, media and

communication, political science, IT, etc.). After initial screening, 82 valid questionnaires were retained for analysis.

The questionnaire consisted of 15 substantive questions and three socio-demographic questions. The first part of the questionnaire was designed to collect information on the profile of scholars studying public opinion. Specifically, we aimed to identify their primary area of professional expertise, their experience in public opinion research, methodological approach, and the methods they employ in their studies. Additionally, we collected information on perceived challenges associated with theoretical concepts and the design of the studies on public opinion, including those triggered by AI and algorithms.

To further explore the challenges experts face, we employed thematic analysis of answers to the open-ended question on the perceived future of the field. Through an inductive coding approach, we identified and categorized themes, providing insights into the methodological and theoretical challenges of public opinion research in the digital era.

**Findings** Our study showed that the more experienced the scholars were, the more they perceived challenges in public opinion research. Senior researchers with more than nine years of experience in their field primarily focus on methodological precision, emphasizing the design of questions and the selection of samples as critical challenges. This suggests that while learning, scholars become more concerned about survey design and representative sampling to trust the validity and reliability of public opinion research. Early-career researchers, including doctoral students, postdoctoral fellows, and those with less than nine years of experience, often face challenges related to respondent bias and dataset complexity as they continue to navigate reducing bias in responses and managing large datasets. Such a strong focus on respondent bias highlights the challenges of creating neutral, unbiased questions and interpreting respondents' responses without the influence of social desirability or cognitive bias. Furthermore, data massification, characterized by an overflow of digital public opinion data, presents a more significant impediment to early-career scholars, who often lack the necessary tools to process and analyze large-scale data.

We also noticed that different academic viewpoints within various disciplines influence the understanding of difficulties and AI-generated risks. Scholars from the fields of law and political science are concerned with regulatory frameworks, democratic processes, and political communication, and their research primarily focuses on misinformation and the dissemination of inauthentic content. The focus of computer scientists on public opinion fragmentation, data-related challenges associated with big data analytics, social media algorithms, and AI-driven content personalization may be attributed to their involvement with big data analysis, social media algorithms, and AI-relevant content personalization. The problem of public discourse fragmentation is a key interest to them because algorithmic sorting and online information bubbles create widespread audience bias and isolate social discussion. However, scholars in the communication and media studies domain manage to bridge these perspectives and regard the dangers of information misinformation and the impact of such narratives via AI as fundamental in the interdisciplinary conversation of AI in the constitution of public opinion.

The thematic analysis highlighted a dual challenge for scholars and experts: addressing technical methodological hurdles while rethinking theoretical frameworks for public opinion research in the digital era. Scholars widely acknowledged the importance of AI and machine learning in processing large datasets from social media, enabling real-time sentiment tracking

and trend identification. However, concerns were raised about over-reliance on these tools, particularly sentiment analysis, due to methodological and theoretical gaps. Consequently, a key debate of the future of public opinion research centers on balancing AI's potential with scientific rigor.

**Originality/value** In this paper, we will discuss challenges and recommendations on data sampling and inclusivity, data access and privacy, data quality, and the use of methods and tools. By collecting experience and ideas from scholars across disciplines, we aim to trace similarities and differences among them. Furthermore, we strive to map the discourse of political communication experts on the challenges of public opinion research.

**Keywords:** *Public opinion, Media study, Expert discourse, COST Action OPINION network*

**JEL classification:** *D79*

**Acknowledgment:** *The research has been supported by a grant from the Faculty of Management and Social Communication under the Strategic Programme Excellence Initiative at Jagiellonian University.*

## REFERENCES

- Baden, C., Kligler-Vilenchik, N., and Yarchi, M. (2020), "Hybrid content analysis: Toward a strategy for the theory-driven, computer-assisted classification of large text corpora", *Communication Methods and Measures*, Vol. 14 No. 3, pp. 165–183. <https://doi.org/10.1080/19312458.2020.1803247>
- Gandini, A., Keeling, S. and Reviglio, U. (2025), "Conceptualising the 'algorithmic public opinion': Public opinion formation in the digital age", *Dialogues on Digital Society*. <https://doi.org/10.1177/29768640251323147>
- Kraft, P. W., Krupnikov, Y., Milita, K., Ryan, J. B., Soroka, S. (2020), "Social media and the changing information environment: Sentiment differences in read versus recirculated news content", *Public Opinion Quarterly*, Vol. 84 No. 1, pp. 195-215, <https://doi.org/10.1093/poq/nfaa015>

## EXPERT DISCOURSE ON KEY THEORIES AND APPROACHES TO STUDYING PUBLIC OPINION: A CONCEPT MAP ANALYSIS

**Agnieszka Hess**

*Jagiellonian University, Krakow, Poland*  
*agnieszka.hess@uj.edu.pl*

**Agnieszka Stępińska**

*Adam Mickiewicz University, Poznan, Poland*  
*agnieszka.stepinska@amu.edu.pl*

**Elena Negrea-Busuioac**

*National University of Political Studies and Public Administration,*  
*Bucharest, Romania*  
*elena.negrea@comunicare.ro*

**Anita Ciunova-Shuleska**

*Ss. Cyril and Methodius University in Skopje,*  
*Faculty of Economics – Skopje, North Macedonia*  
*anita.ciunova@eccf.ukim.edu.mk*

**Anna Bączkowska**

*University of Gdansk, Poland*  
*anna.k.baczkowska@gmail.com*

### EXTENDED ABSTRACT

**Purpose** Public opinion research is undergoing a significant transformation in response to the challenges posed by the digital era, particularly the proliferation of social media and AI-driven tools (Murphy *et al.*, 2014; Kraft *et al.*, 2020). These developments compel scholars to revisit established theories and methodologies for studying opinion formation and expression in dynamic, technology-driven contexts. This study explores the key concepts, models, and perspectives deemed essential by public opinion scholars, using concept map analysis to uncover relationships between these elements and the profiles of experts who employ them. Our study aims to examine: 1) What are the links between key concepts, models, authors, and perspectives that scholars use in their research on public opinion, and 2) How the profile (e.g., expertise, experience, rank, etc.) of the experts maps onto the concepts they use.

**Design/methodology/approach** We collected our data via a survey conducted primarily among members of the COST OPINION network. The questionnaire used in the study comprised 26 questions, two of which were open-ended. A total of 86 questionnaires completed by public opinion scholars and experts from 33 countries were returned. After initial screening, 82 valid questionnaires were further analyzed.

We used quantitative methods to profile experts and qualitative coding of an open-ended question about the concepts frequently used by the respondents in their public opinion-related work. Additionally, we used concept maps to visualize the findings better and to highlight relationships and patterns in our data.

**Findings** Preliminary findings show that the concepts and theories stemming from media studies are most frequently used by respondents in their studies on public opinion, e.g., agenda setting, framing, spiral of silence, and uses and gratification theory. These concepts and theoretical perspectives are often extended to include contemporary approaches to public opinion research in the digital environment, e.g., online processing model, digital public sphere, social media engagement theory, and networked publics. Other approaches reflect the impact of AI tools on public opinion, e.g., sentiment analysis and natural language processing, social network analysis, and algorithmic gatekeeping. Concept map analysis revealed hierarchical clusters of public opinion-related concepts, theories, and models at macro (system theories, public sphere concepts, media environment), meso (discursive mechanisms, attitudes, representation, infrastructure), and micro (emotions, interaction, behavior, socio-psychological processes) levels.

**Originality/value** This study highlights the dynamic interplay between established theories and emerging methodologies in public opinion research, offering insights into how scholars navigate the complexities of opinion formation and expression in the digital age. By mapping expert discourse, our research provides a structured understanding of the evolving landscape of public opinion studies.

**Keywords:** *Public opinion, Expert discourse, Map analysis, COST Action OPINION network*

**JEL classification:** *D79*

**Acknowledgment:** *The research has been supported by a grant from the Faculty of Management and Social Communication under the Strategic Programme Excellence Initiative at Jagiellonian University.*

## **REFERENCES**

- Baden, C., Kligler-Vilenchik, N., and Yarchi, M. (2020), "Hybrid content analysis: Toward a strategy for the theory-driven, computer-assisted classification of large text corpora", *Communication Methods and Measures*, Vol. 14 No. 3, pp. 165–183. <https://doi.org/10.1080/19312458.2020.1803247>
- Kraft, P.W., Krupnikov, Y., Milita, K., Ryan, J.B., Soroka, S. (2020), "Social media and the changing information environment: Sentiment differences in read versus recirculated news content", *Public Opinion Quarterly*, Vol. 84 No. 1, pp. 195-215, <https://doi.org/10.1093/poq/nfaa015>
- Murphy, J., Link, M.W., Childs, J. H., Langer Tesfaye, C., Dean, E., Stern, M., Pasek, J., Cohen, J., Callegaro, M., Harwood, P. (2014), "Social media in public opinion research: Executive summary of the Aapor task force on emerging technologies in public opinion research", *Public Opinion Quarterly*, Vol. 78 No. 4, pp. 788-794. <https://doi.org/10.1093/poq/nfu053>

## **METHODOLOGICAL PROPOSAL FOR PUBLIC OPINION RESEARCH IN THE ERA OF DEMOCRATIC REGRESSION AND DIGITAL TRANSFORMATION**

**Kinga Adamczewska**

*Adam Mickiewicz University, Poznan, Poland  
kinga.adamczewska@amu.edu.pl*

**Roksana Gloc**

*Jagiellonian University, Krakow, Poland  
roksana.gloc@uj.edu.pl*

**Agnieszka Hess**

*Jagiellonian University, Krakow, Poland  
agnieszka.hess@uj.edu.pl*

**Agnieszka Stepínska**

*Adam Mickiewicz University, Poznan, Poland  
agnieszka.stepinska@amu.edu.pl*

**Anna Bączkowska**

*University of Gdańsk, Poland  
anna.baczkowska@ug.edu.pl*

### **EXTENDED ABSTRACT**

**Purpose** The contemporary debate on the state and future of democracy must address the issue of public opinion. This becomes particularly crucial in the context of so-called problematic democracies - political systems exhibiting symptoms of democratic regression. In a digital environment that is dynamic and susceptible to manipulation, the traditional understanding of public opinion proves insufficient. Contemporary digital societies face increasingly complex mechanisms governing the formation and circulation of public opinion. The public sphere, once shaped primarily by institutional media, is now co-created by network users, becoming a dynamic arena where various actors - politicians, journalists, experts, non-governmental organizations, and citizens themselves - compete. This context raises fundamental research questions: How should public opinion be studied today? How can its formation, transmission, and suppression be captured within the digitalized public sphere, especially in countries affected by symptoms of democratic regression?

The aim of this presentation is to outline the research assumptions and methodological proposals for an international project dedicated to analyzing contemporary public opinion. The starting point is to identify both theoretical and practical challenges related to the conceptualization and measurement of public opinion amidst digital transformation, the fragmentation of public debate, the polarization of societies and media, and the weakening of traditional democratic institutions. This methodological proposal thus situates public opinion research at the intersection of democracy, digitalization, and the evolving dynamics of global economic behavior. It highlights how shifts in democratic legitimacy and technological innovation reshape not only the way citizens express their preferences but also how institutions and markets respond to collective attitudes in a rapidly transforming world.

This presentation will reflect on the changing nature of public opinion. As previous research indicates (e.g., Zaller, 1992; Crespi, 1997; Noelle-Neumann, 1974), public opinion is not merely a collection of individual attitudes, but a complex, dynamic process that requires considering mechanisms of social influence and media representation. Foundational works on the crystallization and shifting of the public sphere emphasize that public opinion formats within historically specific communicative and institutional frameworks (Habermas, 1991; Fraser, 1990). Today, the logic of digital platforms reconfigures these frameworks by redefining visibility, authority, and participation within networked spaces. In turn, the contemporary digital environment - encompassing social media, video platforms, and news aggregators - enables rapid information circulation but also introduces new barriers, such as algorithmic reach, content personalization, and the phenomena of polarization and radicalization.

A particularly interesting area for this project is Central and Eastern Europe. The starting point for reflecting on the condition of public opinion in this region is the phenomenon of democratic backsliding, observed for over a decade, whose scale and dynamics have been meticulously captured by projects such as the international V-Dem (Varieties of Democracy) project. This project, based on several hundred indicators of democratic institutional quality, allows for cross-country comparisons while considering their political and media environments. The Liberal Democracy Index (LDI), a synthetic measure of democratic quality, illustrates the diverse situations within the studied countries, for example:

- Poland (LDI, 2023: 0.45, electoral democracy): Problems include the politicization of public media, limited judicial independence, and the strong polarization of public debate.
- Hungary (LDI, 2023: 0.34, electoral autocracy): Monopolization of the media market, restrictions on the opposition, and systemic electoral abuses are observed.
- Serbia (LDI, 2023: 0.26, electoral autocracy): A country plagued by censorship, political pressure on the media, and electoral manipulation.
- Czech Republic (LDI, 2023: 0.80, liberal democracy): A positive example, with free media, an independent judiciary, and strong political pluralism.

In 2009, all the above-mentioned countries were considered liberal democracies – their status indicates dynamic, regressive systemic changes. This phenomenon is increasingly referred to as illiberalism, which – as Štětka and Mihelj (2024) point out – is characterized by:

- a paradoxical attitude towards liberalism – the use of democratic procedures while simultaneously undermining liberal values (e.g., equality, pluralism, freedom of speech);
- ambiguity – existing as both a political ideology and an institutional practice;
- processuality – variability and scalability, which can lead to both a further deepening of authoritarianism and attempts to rebuild democracy.

Furthermore, Uitz and Sajó (2017) note that illiberal systems maintain the pretense of democracy, primarily through elections, while consistently undermining media independence, judicial autonomy, and the activities of civil society institutions. The selected countries represent varying trajectories of democratic backsliding and radicalization within comparable historical, cultural, and media contexts, allowing for a nuanced comparative analysis of how illiberal transformations shape the formation and circulation of public opinion.

Including this context in the analysis of public opinion is essential to understand why and how citizens' opinions can be deformed, blocked, or instrumentally used by dominant political and

media actors. Thus, the proposed research model involves not only mapping content and behaviors but also identifying systemic mechanisms influencing opinion circulation and the quality of public deliberation.

Against this backdrop, the analysis of "blocking" public opinion takes on particular significance - referring to situations where citizens' opinions have no chance of entering public debate due to self-censorship, disinformation, the dominance of powerful actors, or the inaccessibility of communication spaces. Therefore, the research project proposes conceptualizing public opinion as a relational communication network and analyzing the activities of four key groups: experts, journalists, politicians, and social organizations - along with audiences/citizens as co-participants in this interplay. Treating citizens as participants rather than mere audiences requires methodological grounding in the concept of participatory communication and networked publics (Papacharissi, 2010; Couldry and Hepp, 2017). In this project, citizens are understood as active nodes in the opinion network—engaged not only in content consumption but also in its circulation, commentary, and reinterpretation. This approach will be operationalized through a mixed-method design combining survey data (capturing self-reported participation patterns) with qualitative materials (interviews and focus groups revealing experiential dimensions of participation).

**Design/methodology/approach** We propose that the study encompass four main opinion-forming groups: (1) experts (scientists, think tank leaders), (2) politicians, (3) journalists, and (4) NGO representatives. Citizens were considered separately as participants (not just audiences) in the public debate. Each of these groups was treated as a separate research component.

To capture the complexity of the relationships between these actors, the project utilizes diverse research methods. The proposed model is based on a triangulation of quantitative and qualitative methods, enriched with network analysis and a comparative approach.

- A. Surveys: Targeted at citizens and, to some extent, at NGOs and journalists. This will allow for mapping declared attitudes, sources of information, levels of trust, and self-perceptions of participation in public debate.
- B. In-depth interviews: Conducted with representatives of all four main opinion-forming groups. This will serve to further explore motivations, barriers, practices, and perceptions of the state of public debate.
- C. Focus groups: Organized among NGOs and citizens of all ages. This will enable the interactive and contextual understanding of opinion-forming mechanisms.

The network analysis will serve as an integrative layer linking qualitative and quantitative data. On the quantitative side, it will map relational ties (e.g., interaction structures) derived from surveys. On the qualitative side, it will be used interpretively to identify clusters of meaning and influence among experts, journalists, politicians, and NGOs, thus bridging structural and discursive dimensions of opinion formation.

The study will be conducted in selected countries of Central, Eastern, and Southern Europe, considering their specific political contexts, media culture, history, and current political processes. Comparative analysis could be useful to identify common and local patterns in public opinion formation. At the same time, the study aims to answer the question of whether it is possible to develop a universal model of public opinion research that will also be applicable in imperfect or regressive democracies.

**Findings** The project aims to produce a new, realistic framework for public opinion research, considering the complex relationships between opinion elites (experts, politicians, media) and society. This model will be based on three levels:

1. Micro-level: individual opinion formation (attitudes, emotions, exposure to content),
2. Meso-level: group interactions (media, NGOs, relationships with opinion leaders),
3. Macro-level: dominant narratives, institutional frameworks for debate.

The culmination of the project will be a methodology combining the tools of sociology, political science, and media studies, also useful in future studies of complex sociopolitical processes.

**Originality/value** In the problematic democracies - characterized by increasing polarization and distrust in media and state institutions - the ongoing information processes determine the quality of public debate, which forms the foundation of socio-economic stability. Therefore, reflecting on who influences opinion formation, how opinions are transmitted or blocked, and how citizens participate (or are excluded) from information circulation is essential. This understanding is vital not just for comprehending the future of politics, but also for grasping social relations, brand reputation, and the attitudes of citizen-consumers.

**Keywords:** *Public opinion, Methodology, Democratic regression, Digital transformation*

**JEL classification:** *D71, D72, O33*

## REFERENCES

- Couldry, N. and Hepp, A. (2017), *The Mediated Construction of Reality*, Polity.
- Crespi, I. (1997), *The Public Opinion Process. How the People Speak*, Routledge, New York, London.
- Fraser, N. (1990), "Rethinking the public sphere: A contribution to the critique of actually existing democracy", *Social Text*, Vol. 25/26, pp. 56-80.
- Habermas, J. (1991), *The Structural Transformation of the Public Sphere: An Inquiry Into a Category of Bourgeois Society*, The MIT Press, Cambridge, MA.
- Liberal Democracy Index (2023), The Varieties of Democracy (V-Dem) project: <https://www.v-dem.net/data/the-v-dem-dataset/>
- Noelle-Neumann, E. (1974), "The spiral of silence: A theory of public opinion", *Journal of Communication*, 24(2), pp. 43–51.
- Papacharissi, Z. (2010), *A Private Sphere: Democracy in a Digital Age*, Polity.
- Štětka, V., and Mihelj, S. (2024). *The Illiberal Public Sphere. Media in Polarized Societies*, Palgrave Macmillan, Cham.
- Uitz, R. and Sajo, A. (2017), *The Constitution of Freedom: An Introduction to Legal Constitutionalism*, Oxford University Press, Oxford.
- Zaller, J. R. (1992), *The Nature and Origins of Mass Opinion*, Cambridge University Press, Cambridge.

# THE USE OF VR TECHNOLOGY IN POLITICAL COMMUNICATION: PROSPECTS AND CHALLENGES IN SHAPING PUBLIC OPINION

**Malgorzata Durmaj**

*Doctoral School in the Social Sciences,  
Jagiellonian University, Krakow, Poland  
gosia.durmaj@doctoral.uj.edu.pl*

**Anna Kulma**

*Doctoral School in the Social Sciences,  
Jagiellonian University, Krakow, Poland  
anna.kulma@doctoral.uj.edu.pl*

## EXTENDED ABSTRACT

**Purpose** The paper presents the preliminary results of the pilot study and the developed tool, a virtual reality software. The aim is to explore the participatory practices and attitudes of young adults toward implementing technologies in mediatized social communication, including political communication. The authors present their original VR environment and analyze user behavior toward interacting with political candidates and voting in virtual reality. The prospects and challenges associated with the use of immersive tools in shaping public opinion and developing the technology sector, including in Poland, were examined.

**Design/methodology/approach** The research project is framed by the constructivist paradigm. The following research questions have been formulated: Q1: How does the presence of political candidates in virtual reality (VR) shape the behavior and voting decisions of young adults?; Q2: What attitudes do young adults have toward the presence of political candidates in VR?; Q3: What are the prospects and challenges associated with the use of immersive tools in shaping public opinion and developing the technology sector?. The authors conducted empirical research using data analysis triangulation, investigator triangulation, and methodological triangulation. A mixed-method approach was used – the participatory workshop was supplemented with focus group interviews, overt participant observation, and a survey. The non-probability sample included adult undergraduate students of journalism and social communication at the Jagiellonian University, Poland. The following tools were used: original VR software with avatars of fictional political candidates and 3D objects (a ballot box and voting paper), survey and introductory questionnaires, and a FGI script. Inductive coding of FGI discussion transcripts and responses to open-ended survey questions was performed using MAXQDA software.

**Findings** The results have provided a deeper understanding of modifications related to the implementation of new technologies in mediatized public life. The preliminary findings of the pilot study have enriched the scientific dialogue with perspectives on shaping public opinion by VR technology. The attitudes and behaviors of young adults toward the presence of political candidates and the possibility of voting in VR have been identified. The outcomes of the study have highlighted the strengths and limitations of the original VR software and have presented perceived opportunities and threats of using immersive technology in political communication. Reflections on immersive media have shown a blurring of boundaries between "non-digital reality–virtual reality" (*see also* Schweiger and Wimmer, 2023), the effect of VR on the familiarization of political behavior (*see also* Weber *et al.*, 2022), the issue of accessibility, the reduction of the distance between "political candidate–user" and

concurrent concerns about the disinformation, deepfakes, and the loss of naturalness of avatar's speech (*see also* Kroczek and Mühlberger, 2023; Sterna *et al.*, 2019). The results of the study enabled the formulation of guidelines for developers of virtual reality solutions.

**Originality/value** The findings enabled the creation of guidelines for the ethical and effective application of VR tool features, primarily the ones designed for political communication and potential public opinion formation. The results are even more significant considering more than 171 million VR users worldwide (Kumar, 2025) and the developing virtual reality market. The market value is estimated at USD 16.32 billion in 2024, USD 20.83 billion in 2025, and is projected to reach USD 123.06 billion by 2032 (CAGR of 28.9%) (Fortune Business Insight, 2025) and USD 261.92 billion by 2034 (CAGR of 35.60%) (Precedence Research, 2025). Statista Market Insights (n.d.) argues that the VR software market in Poland is expected to reach 1.2 million users by 2029. The estimated VR usage rate in Poland is 2.7% in 2025 and 3% in 2029. In addition, the project is application-oriented in the social sphere. It concentrates on developing awareness and creating approaches to mediatization (Hepp, 2020; Hepp *et al.*, 2024) and the use of VR technology (Guttentag, 2010) in mediatized political communication. The authors propose their original VR environment as a tool for public opinion and participatory practices studies. Due to the digitalization of public debate, the importance of implementing new technologies in research in this area is recognized. Furthermore, the issue of expected differences in attitudes and behaviors among different age groups was raised during the participatory workshop. This approach creates an opportunity to continue the discussion about the immersion in the VR environment with participants of various ages, such as seniors.

**Keywords:** *Political immersion, Virtual reality, New technologies, Public opinion, Mediatization of political communication, Political participation*

**JEL classification:** *D72, D71, O33, O35*

## REFERENCES

- Fortune Business Insight (2025), "Virtual Reality (VR) Market Size, Share & Industry Analysis, By Component (Hardware and Software), By Technology (Semi & Fully Immersive and Non-Immersive) By Device Type (Head Mounted Display (HMD), VR Simulator, VR Glasses, Gloves, and Others), By End-user (Consumer and Commercial/Enterprise), and Regional Forecast, 2025-2032", available at: <https://www.fortunebusinessinsights.com/industry-reports/virtual-reality-market-101378> (accessed 9 July 2025).
- Guttentag, D.A. (2010), "Virtual reality: Applications and implications for tourism", *Tourism Management*, Vol. 31 No. 5, pp. 637-651.
- Hepp, A. (2020), *Deep Mediatization*, Routledge, New York, NY.
- Hepp, A., Bolin, G., Guzman, A.L., and Loosen, W. (2024), "Mediatization and human-machine communication: Trajectories, discussions, perspectives", *Human-Machine Communication*, Vol. 7, pp. 7-22. <https://doi.org/10.30658/hmc.7.1>
- Kroczek, L.O.H. and Mühlberger, A. (2023), "Public speaking training in front of a supportive audience in Virtual Reality improves performance in real-life". *Scientific Reports*, No. 13, 13968. <https://doi.org/10.1038/s41598-023-41155-9>
- Kumar, N. (2025), *Virtual Reality Statistics 2025: Users & Trends*, available at <https://www.demandsage.com/virtual-reality-statistics/> (accessed 9 June 2025).

- Precedence Research (2025), "AR and VR Headsets Market Size, Share, and Trends 2025 to 2034", available at: <https://www.precedenceresearch.com/ar-and-vr-headsets-market> (accessed 9 June 2025).
- Schweiger, M. and Wimmer, J. (2023), "Perceiving media change as a form of media literacy: Using a 360-degree video to analyze the mediatization of social life", *Mediatization Studies*, No. 7, pp. 9-30.
- Statista Research Department (n.d.), "VR Software -Poland", available at: <https://www.statista.com/outlook/amo/ar-vr/vr-software/Poland> (accessed 28 April 2025).
- Sterna, R, Strojny, P. and Rebilas, K. (2019). "Can virtual observers affect our behavior?", *Social Psychological Bulletin*, Vol. 14 No. 3, pp. 1-18.
- Weber, W., Dingerkus, F., Fabrikant, S.I., Zampa, M., West, M., and Yildirim, O. (2022). "Virtual Reality as a Tool for Political Decision-Making? An Empirical Study on the Power of Immersive Images on Voting Behavior", *Frontiers in Communication*, Vol. 7, 842186. <https://doi.org/10.3389/fcomm.2022.842186>

**PUBLIC OPINIONS AND THE CONSTRUCTION OF SELF- AND OTHER-IDENTITY IN ANTI-IMMIGRATION DISCOURSE TOWARDS UKRAINIANS IN POLAND**

**Anna Bączkowska**

*University of Gdańsk, Poland  
anna.baczkowska@ug.edu.pl*

**Agnieszka Hess**

*Jagiellonian University, Krakow, Poland  
agnieszka.hess@uj.edu.pl*

**Artur Lipiński**

*Adam Mickiewicz University, Poznan, Poland  
artur.lipinski@amu.edu.pl*

**Arkadiusz Misztal**

*University of Gdańsk, Poland  
arkadiusz.misztal@ug.edu.pl*

**Aneta Dłutek**

*Poland University of Gdańsk, Poland  
dlutekaneta@gmail.com*

**Joanna Tillack**

*University of Gdańsk, Poland  
joanna.tillack@ug.edu.pl*

**EXTENDED ABSTRACT**

**Purpose** One of the consequences of digital transformation is social polarisation, which has negative effects on discourse polarisation and radicalization, particularly targeting immigrants. These processes are also present in the case of Poland and a flood of Ukrainian immigrants. As a result, scholars seek novel tools and methodologies to grasp the scale of discourse radicalization and society polarization. The content online is expressed particularly freely on social media, which house distributed, decentralized information created in a bottom-up fashion (Yarchi *et al.*, 2020). They have become a substitute for institutionalized public debate (Baden *et al.*, 2025) and often allow the radical content to spread virally. Therefore, the purpose of this study is to demonstrate how this radical, polarized discourse can be analyzed.

The onset of the Russian invasion of Ukraine in February 2022 precipitated massive emigration of Ukrainians to the neighbouring countries, with Poland emerging as the host of the most significant number of immigrants from the conflict-ridden nation. The unprecedented aid and hospitality extended to Ukrainians by Poles were lauded across European media. This initial positive sentiment and welcoming reception, however, have gradually shifted towards a growing dissatisfaction and even hostility towards Ukrainians over the ensuing three years, as evidenced on social media.

The purpose of this research is thus to analyse the social climate of Poles discussing Ukrainian immigrants to Poland based on comments expressed on social media (YouTube) as a reaction to an interview with a Ukrainian leader of the Ukrainian diaspora in Poland.

**Design/methodology/approach** This study aims to analyze comments posted in February 2025 on Polish YouTube regarding the situation and status of Ukrainians in Poland and the perception of Ukrainians by Poles following the widely publicized, controversial statements of Natalia Panchenko, the leader of the Ukrainian diaspora in Poland. The leader's provocative opinions on Poles and the status of Ukrainians in Poland, presented during her interview on the Ukrainian TV Channel 5, sparked extensive commentary by Polish journalists, politicians, and social media users. Even though she later denied expressing these statements, attributing them to Russian manipulation and fake news in several interviews published on YouTube, her words nonetheless provoked a significant backlash, which manifested in an outpouring of grievances, expressing disapproval or even condemnation of some Ukrainians, as well as a severe criticism of the immigration policy adopted by the Polish government.

Following the interview on the Ukrainian TV, she gave another interview on the Super Express YouTube channel (affiliated with one of Poland's private television stations), where she attempted to clarify her views on Polish-Ukrainian relations and to fend off the severe criticism she had received after her controversial remarks on Ukrainian television.

The critical comments analyzed in this study were posted by YouTube users in response to her interview with the Polish journalist. The comments were collected over a one-month period, resulting in a dataset of approximately 7,000 entries. A randomly selected sample of ca. 1200 sample of comments was extracted from the original set and analyzed both qualitatively and quantitatively. For qualitative analysis, Atlas.ti was used to enable research in line with the Grounded Theory (Strauss and Coblin, 1994). A carefully selected sample of the collected comments was next annotated by four expert annotators and a super-annotator after a series of training sessions (with inter-rater agreement oscillating around 0.8). The annotation-based qualitative study is followed by a computerized automatic examination of the data, particularly sentiment analysis, topic modelling, and trends. The tools used in this part of the study stem from corpus linguistics methodology and the latest, cutting-edge neural network-based tools (the so-called deep learning methods) known as Transformers (BERT; Devlin *et al.*, 2019).

Theoretically, the study is framed within theories of identity, Critical Discourse Analysis (e.g., van Leeuwen, 2008), and theories and models of evaluative language. The analysis had the following research questions:

RQ (1) What is the general perception of Ukrainian residents in Poland by Poles 3 years after the war onset, according to the comments under inspection?

RQ (2) How is the Ukrainian identity constructed by the Polish YT users?

RQ (3) How is the Polish identity constructed by the Polish YT users in the context of the Ukrainian refugees' flood?

**Findings** The study has uncovered interesting tendencies and trends in how Polish social media users evaluate immigrants from the war-ridden Ukraine over time, and which factors are crucial in the user-generated, radically negative opinions. Our research also shows the effectiveness of computational methods when applied to media discourse in public opinion research.

**Originality/value** The originality of our study is twofold. It resides in merging two scholarly approaches and traditions: media studies and linguistics. On the other hand, the use of computerized automatic data retrieval and analysis, particularly neural networks represented by deep learning, is a novel approach in media studies. The categories of description that researchers typically resort to in such studies are various types of discourse strategies, well-known categories in linguistic scholarship. In our study, however, we propose novel categories which combine methodological traditions originating in linguistic and media studies.

**Keywords:** *Political polarization, Computational methods, Ukrainians*

**JEL classification:** *J28, J61, J83*

## REFERENCES

- Baden, C. Kligler-Vilenchik, N., Springer, N., Jungblut, M., Zelenkauskaitė, A., Balčytienė, A., Salgado, S., Lipiński, A., Krstić, A. and Bączkowska, A. (2025), "Public opinion negotiations in a digital media ecosystem: A conceptual framework", *International Journal of Public Opinion Research*, Vol. 37 No. 3, edaf049.
- Devlin, J. Chang, M.-W. Lee, K. and Toutanova, K. (2019), "BERT: Pre-training of deep bidirectional transformers for language understanding", *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171-4186.
- Strauss, A., and Corbin, J. (1994), "Grounded Theory Methodology: An Overview", in Denzin, N. and Lincoln, Y. (Eds.), *Handbook of Qualitative Research*, SAGE, Thousand Oaks, CA, pp. 273-284.
- Van Leeuwen, T. (2008), *Discourse and Practice: New Tools for Discourse Analysis*, Oxford University Press, Oxford.
- Yarchi, M. Baden, C. and Kligler-Vilenchik, N. (2020), "Political Polarisation on the Digital Sphere: A Cross-platform, Over-time Analysis of Interactional, Positional, and Affective Polarization on Social Media", *Political Communication*, Vol. 38 No. 1-2, pp. 98-139.

## EXTENDED ABSTRACTS

This section includes extended abstracts presented at the conference that underwent review by a single reviewer and are organized into the following areas: Economics and Marketing, Management & E-business.

### *Economics*

#### **EFFECTS OF LABOR MARKET MEASURES AND POLICIES ON IMPROVING LABOR MARKET PERFORMANCE IN CENTRAL AND SOUTH EASTERN EUROPEAN (CSEE) COUNTRIES**

**Kristijan Kozheski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
kozeski@eccf.ukim.edu.mk*

**Trajko Slaveski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
Trajko.Slaveski@eccf.ukim.edu.mk*

**Predrag Trpeski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
predrag.trpeski@eccf.ukim.edu.mk*

**Borce Trenovski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
borce.trenovski@eccf.ukim.edu.mk*

**Gunter Merdzan**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
gjunter.merdzan@eccf.ukim.edu.mk*

#### **EXTENDED ABSTRACT**

**Purpose** The study aims to examine the rationale, structure, and effectiveness of labor market measures and policies in Central and South-Eastern European (CSEE) countries, with particular attention to the post-transition challenges that have shaped their labor market dynamics. It highlights the critical role of Active Labor Market Policies (ALMPs) in enhancing employability, reducing structural unemployment, and fostering long-term socio-economic inclusion through targeted interventions addressing skill mismatches and persistent unemployment.

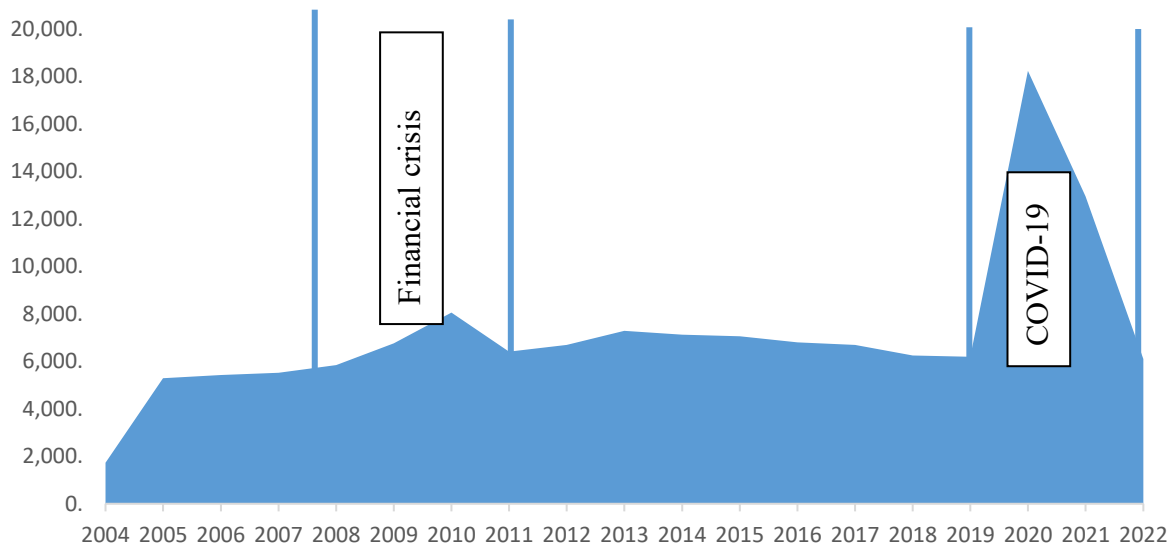
**Design/methodology/approach** The research applies a comparative analytical approach, combining theoretical insights from labor economics with empirical evidence drawn from policy evaluations and statistical indicators for selected CSEE countries. The analysis

categorizes labor market policies into six main groups—public employment services, training programs, employment incentives, direct job creation, entrepreneurship promotion, and financial support for the unemployed—and evaluates their scope, design, and institutional effectiveness across differing economic contexts. In addition, the empirical strategy employs a dynamic panel-data GMM estimator to identify causal effects while addressing endogeneity, unobserved heterogeneity, and potential serial correlation.

**Findings** The findings reveal that the efficiency and sustainability of labor market policies in Central and South Eastern European (CSEE) countries are largely contingent upon institutional quality, policy coordination, and fiscal capacity. Well-targeted active labor market policies (ALMPs)—particularly training programmes and employment incentives—have demonstrated measurable impacts in reducing long-term and youth unemployment (Kluve, 2006; Hijzen and Venn, 2011). In contrast, direct job creation and financial support measures tend to yield only transitory effects unless integrated with active reintegration strategies and skill development initiatives (Card *et al.*, 2018).

The temporal analysis of labor market policy expenditures from 2004 to 2022 identifies distinct phases shaped by post-transition economic transformation, EU accession processes, and exogenous shocks such as the 2008 global financial crisis and the COVID-19 pandemic. During periods of heightened macroeconomic instability, particularly in 2009–2010 and again in 2020–2021, governments in the region expanded ALMP spending through wage subsidy schemes and job retention programs (Hijzen and Venn, 2011). However, the subsequent withdrawal of emergency measures and the re-emergence of fiscal constraints reveal a persistent reliance on reactive, crisis-driven interventions. This contrasts with the more structurally embedded, long-term human capital and productivity strategies observed in advanced EU member states (Card *et al.*, 2018). The findings thus underscore the importance of transitioning toward forward-looking labor market policies that are institutionally grounded, fiscally sustainable, and aligned with broader goals of inclusive and resilient economic growth.

*Figure 1: Total Expenditures on Labor Market Policies and Measures in Selected Central and Eastern European Countries (in million EUR)*



(Source: EUROSTAT)

The data presented in Table 2 indicate pronounced differences in the distribution of total labor market policy expenditures among the selected Central and South Eastern European countries

during the period 2004–2022. Poland consistently dominates the expenditure structure, accounting for between 41.5% and 59.3% of total spending, reflecting both the size of its economy and its long-standing strategic commitment to improving labor market performance and reducing unemployment. Hungary follows with an average share of approximately 15%, although its expenditure dropped sharply to 7.3% in 2020 due to temporary fiscal adjustments and programme restructuring during the COVID-19 pandemic, before recovering in 2022. The Czech Republic ranks third, recording a 23% share of total expenditures in 2022, while Romania and Bulgaria exhibit notably low and declining spending levels, with Romania’s share falling from 8.1% in 2005 to only 1.3% in 2020, suggesting limited fiscal support and weak institutional capacity. Slovenia maintains a relatively stable share of around 4–5%, indicating consistent but modest financing, constrained by its smaller economic base. Croatia, included in the later years of observation, recorded a substantial increase in expenditure in 2020, driven primarily by emergency job-retention programmes introduced during the pandemic. Overall, the data from Table 2 demonstrate that countries with stronger institutions and greater fiscal capacity, such as Poland and the Czech Republic, sustain higher and more stable levels of labor market policy spending, while those with weaker institutional frameworks, such as Romania and Bulgaria, remain constrained by low and fragmented investment in labor market interventions.

*Table 1: Share of Selected Central and Eastern European Countries in Total Expenditures on Labor Market Measures and Policies (in %)*

Country	2005	2010	2015	2020	2022
<b>Bulgaria</b>	3,0	2,6	3,6	4,3	6,2
<b>Romania</b>	8,1	9,3	4,2	1,3	
<b>Slovenia</b>	3,8	5,4	4,2	5,4	4,3
<b>Croatia</b>			4,8	7,7	
<b>Czechia</b>	9,3	13,0	14,6	15,7	23,0
<b>Poland</b>	59,3	45,7	44,9	50,0	41,5
<b>Hungary</b>	12,1	16,4	17,8	7,3	12,0
<b>Slovakia</b>	4,4	7,7	6,0	8,3	13,0

*(Source: ILO STAT)*

The data presented in Table 3 illustrate significant variation in the allocation of expenditures across distinct categories of active labor market measures (ALMMs) in the selected Central and South Eastern European (CSEE) countries over the period 2004–2022. The largest proportion of spending is consistently directed toward employment incentives, which peaked in 2020 at EUR 3,670.26 million and remained elevated in 2021. This trend reflects the widespread implementation of wage subsidy schemes and job-retention programmes introduced as emergency fiscal responses to mitigate mass layoffs and labor market disruptions caused by the COVID-19 pandemic. Expenditures on services provided by public employment institutions also constitute a substantial share of total spending, displaying steady growth from EUR 521.57 million in 2005 to EUR 862.31 million in 2022. This pattern suggests continued institutional investment in employment mediation, counselling, and job placement support. In contrast, spending on training and skills development programmes has declined markedly, from EUR 564.99 million in 2008 to just EUR 86.6 million in 2022, indicating a reduced and inconsistent emphasis on long-term strategies for human capital enhancement. Meanwhile, expenditures for direct job creation and support for self-employment exhibit cyclical dynamics, with noticeable increases during crisis periods—particularly in 2010 and 2020—underscoring their role as short-term countercyclical instruments rather than components of a sustained

developmental agenda. Overall, the evidence from Table 3 suggests that active labor market policy in the region remains primarily oriented toward short-term labor market stabilization, while structural interventions aimed at improving skills, fostering entrepreneurship, and enhancing long-term labor market resilience receive comparatively limited and inconsistent financial prioritization.

Table 2: Expenditures by Categories of Active Labor Market Measures in Selected Central and South Eastern European Countries, 2004–2022 (in million EUR)

Year	Services Provided by Public Labor Market Institutions	Training and Skills Development Programmes	Employment Incentives	Direct Job Creation	Incentives for Starting One's Own Business
2005	521,57	341,21	282,19	314,16	434,37
2006	645,79	387,55	291,73	319,21	447,24
2007	701,49	421,61	381,2	313,88	533,55
2008	758,92	564,99	417,52	342,38	653
2009	716,66	252,92	708,62	456,42	959,84
2010	735,26	292,91	1062,18	660,4	1132,33
2011	762,37	147,07	648,12	376,15	849,89
2012	766,07	147,55	680,57	622,52	641,8
2013	753,08	167,01	829,68	882,92	599,33
2014	802,68	169,42	918,92	977,38	543,58
2015	782,65	149,19	943,17	1091,39	513,27
2016	755,16	120,63	903,9	1114,92	488,38
2017	799,67	130,04	873,21	1014,93	490,37
2018	753,97	145,15	811,52	748,26	472,45
2019	858,15	123,5	707,75	683,09	484,53
2020	883,38	88,71	3670,26	495,88	676,59
2021	835,29	71,85	2577,58	511,61	522,87
2022	862,31	86,6	691,03	422,33	573,75

(Source: ILO STAT)

The results of the econometric analysis indicate that among all categories of active labor market measures, the most substantial and statistically significant positive effect on employment in Central and South Eastern European countries is associated with services provided by public labor market institutions. A 10% increase in fiscal spending on such services is estimated to raise the employment rate by 0.2%, underscoring the critical role of institutional support, including counselling, job matching, and individualized assistance, in facilitating workers' integration into the labor market. Conversely, training and skills development programmes display a statistically significant negative relationship with employment, reflecting their limited efficiency, low participation rates, and insufficient alignment with labor market needs, particularly in countries such as Bulgaria and Romania, where training initiatives often remain formalistic and poorly connected to employer demand. Employment incentives, in contrast, exhibit a positive and significant effect, where a 10% increase in fiscal incentives is expected to raise employment by 0.1%, confirming their relevance as short-term activation tools. The strongest effect, however, is observed for direct job creation measures: a 10% rise in related expenditures increases employment by an estimated 0.4%, suggesting that these interventions are particularly effective during crises, albeit with limited long-term sustainability. Measures promoting self-employment show a negative and significant impact, indicating structural

weaknesses such as weak entrepreneurial ecosystems, regulatory uncertainty, and inadequate institutional support. Overall, the findings suggest that the effectiveness of labor market measures across the region is strongly contingent upon the institutional capacity and operational efficiency of public employment services. Countries with well-developed and adaptable institutions achieve higher and more sustained employment outcomes, while those with fragmented systems and limited fiscal resources face persistent challenges in reducing unemployment and enhancing human capital.

Table 3: Expenditures by Categories of Active Labor Market Measures in Selected Central and South Eastern European Countries, 2004–2022 (in million EUR)

Dynamic Panel Data Estimation Results (Arellano-Bond GMM)		
Number of observations = 132		
Number of countries = 8		
Instruments (total) = 10		
Wald $X^2(4) = 29346$		
Prob > $X^2 = 0.000$		
Dependent variable: employment		
Variable	Coefficient	Robust. Std. Err.
L1.employment	.9304527***	.063456
services	0.0229 ***	0.0041
training	-0.0179 **	0.0029
empl_incentives	0.0109*	0.0017
job_creation	0.0384 **	0.0049
start_up	-0.0233 ***	0.0070
out_work	0.0011	0.0011
cons	49.9130 ***	0.4950
Significance levels: *** p<0.01, ** p<0.05, * p<0.10		
Arellano-Bond test for AR(1) in first differences: z = -1.79 p = 0.074		
Arellano-Bond test for AR(2) in first differences: z = 0.15 p = 0.878		
Sargan test of overid. restrictions: $\chi^2(5) = 3.87$ p = 0.568		
Hansen test of overid. restrictions: $\chi^2(5) = 4.01$ p = 0.547		

(Source: Authors' own calculations)

**Originality/value** This study contributes to the academic discourse by providing a comprehensive and context-specific evaluation of labor market policies in post-transition economies, emphasizing the interaction between policy design, institutional capacity, and macroeconomic environment. It offers valuable insights for policymakers seeking to enhance the long-term resilience of labor markets in CSEE through integrated, evidence-based, and human capital-oriented strategies.

**Keywords:** Labor market policies, Employment incentives, Active labor market measures, Institutional capacity, Central and South Eastern Europe

**JEL classification:** J08, J21, J68

## REFERENCES

Card, D., Kluge, J. and Weber, A. (2018), "What works? A meta analysis of recent active labor market program evaluations", *Journal of the European Economic Association*, Vol. 16 No. 3, pp. 894-931. <https://doi.org/10.1093/jeea/jvx028>

- Eurostat, (2023), *Labor Market Policy Statistics (LMPS): Database*, European Commission. <https://ec.europa.eu/eurostat>
- Hijzen, A. and Venn, D. (2011), *The Role of Short-time Work Schemes during the 2008–09 Recession*, OECD Social, Employment and Migration Working Papers No. 115. OECD Publishing. <https://doi.org/10.1787/5kgkd0bbwvxp-en>
- ILO, (2023). *ILOSTAT database: Labor market indicators and expenditure on active labor market policies*. International Labor Organization. <https://ilostat.ilo.org>
- Kluve, J. (2006), “The effectiveness of European active labor market policy”, *IZA Discussion Paper*, No. 2018. Institute for the Study of Labor (IZA).
- Kuddo, A. (2009), *Employment services and active labor market programs in Eastern European and Central Asian countries*, Social Protection Discussion Paper No. 0918. The World Bank.
- OECD, (2022), *OECD Employment Outlook 2022: Building back more inclusive labor markets*. OECD Publishing. <https://doi.org/10.1787/1bb305a6-en>

## HOW DO FIRMS RESPOND TO MINIMUM WAGE INCREASES IN MACEDONIA?

**Dragan Tevdovski**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
dragan@eccf.ukim.edu.mk*

**Branimir Jovanovikj**

*Vienna Institute for International Economic Studies, Austria  
jovanovic@wiiw.ac.at*

**Viktor Stojkoski**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
vstojkoski@eccf.ukim.edu.mk*

### EXTENDED ABSTRACT

**Purpose** We examine the effects of the 2017 minimum wage reform in North Macedonia on firm-level outcomes, with a focus on employment, productivity, and profitability. Contributing to the broader debate on the consequences of minimum wage increases (see, for example, Petreski *et al.*, 2019; Jovanovikj *et al.*, 2021; Jovanovikj and Naumovski, 2021; Trenovski *et al.*, 2021; Petreski and Pehkonen, 2024), we apply the established concept of the *minimum wage bite*—defined as the share of workers in a firm earning the minimum wage prior to the reform—as a measure of treatment intensity. Our main contribution lies in applying this framework to the full population of registered firms using administrative panel data, allowing for a detailed and representative analysis of policy effects across the entire private sector.

**Design/methodology/approach** We use firm-level panel data covering more than 46,000 registered firms across all sectors, observed in two waves (2016 pre-reform and 2018 post-reform), thus providing a comprehensive two-period firm-level panel. We apply a difference-in-differences (DiD) methodology (Draca *et al.*, 2011; Skedinger, 2014; Yagan, 2015; Saez *et al.*, 2017; Mayneris *et al.*, 2018; Poncet and Zhang, 2018; Gerogiadis *et al.*, 2020), and operationalize treatment intensity through the minimum wage bite, defined as the share of minimum-wage workers in each firm prior to the reform. Firms with a higher bite serve as the treatment group, while firms with little or no bite act as the control group. Mathematically, the *bite<sub>i</sub>* of firm *i* is:

$$bite_i = \frac{\# \text{ of employees in 2016 with minimum wage}}{\# \text{ of employees in 2016}}.$$

Using this variable, we construct the following model:

$$y_{it} = b_0 + b_1 bite_{it} + b_2 D_t bite_{it} + b_3 x_{it} + \mu_t + \eta_i + e_{it}$$

where  $y_{it}$  represents the outcome of interest for the firm  $i$  in year  $t$ , such as total employment, average wages (excluding minimum wage earners), productivity, profitability, or other expenses.  $D_t$  is a post-reform indicator (equal to 1 for 2018),  $x_{it}$  is a vector of control variables (the log of assets, revenues, expenditures, and number of employees, if they are not the

dependent variable),  $\mu_t$  denotes year fixed effects,  $\eta_i$  captures firm fixed effects, and  $e_{it}$  is the error term.

The coefficient of primary interest is  $\beta_2$ , which captures the differential change in outcomes for firms with higher exposure to the minimum wage reform (i.e., those with a larger bite). This interaction term provides a strict measure of the policy's effect, isolating the impact of the reform by comparing firms more and less affected by the increase in the minimum wage, while controlling for time-invariant firm characteristics and broader economic trends.

**Findings** We find that the 2017 minimum wage reform did not result in employment losses among firms more exposed to the policy. The estimated coefficient on the number of employees is negative (-0.421), but statistically insignificant at the 0.01 level, suggesting that firms with a higher minimum wage bite did not reduce their workforce in response to the wage increase. Moreover, we observe statistically significant improvements in labor productivity among more affected firms. The coefficient on productivity is positive (0.0136), although not statistically significant at conventional levels, the direction of the effect is consistent with the hypothesis that the reform may have incentivized firms to reorganize production, enhance efficiency, or motivate workers more effectively. Importantly, no adverse effects are detected on profitability. The estimated coefficient for profitability is negative but very small (-0.003) and statistically insignificant, indicating that firms were largely able to absorb the cost of the wage increase without significant financial strain. We also find a statistically significant reduction in non-wage expenses among affected firms, with an estimated effect of -312.7 ( $p < 0.01$ ). This suggests that firms may have reallocated resources or streamlined other operational costs in response to higher labor expenses. Finally, we detect no statistically significant spillover effects on the wages of other (non-minimum-wage) employees, as the estimated effect is negative but imprecisely estimated (-183.7 with a standard error of 614.2). Taken together, these results point to a reallocation and efficiency-enhancing response by firms, rather than contraction or retrenchment. The findings are consistent with the view that minimum wage policies, when applied in a context such as North Macedonia's, can improve productivity without harming firm performance.

**Originality/value** While previous studies have analyzed the effects of minimum wage reforms in North Macedonia (Petreski *et al.*, 2019; Jovanovikj *et al.*, 2021; Jovanovikj and Naumovski, 2021; Trenovski *et al.*, 2021; Petreski and Pehkonen, 2024), this study is the first to do so using administrative data that covers the entire population of registered firms. This enables a more comprehensive and representative evaluation of the reform's outcomes.

The results of our study indicate that the 2017 minimum wage reform in North Macedonia achieved its objectives without producing significant adverse effects on firm performance. These findings are highly relevant to the broader question of whether the institutional framework for minimum wage setting in the country is adequate. At present, minimum wage increases are determined through legislation following consultations among the government, employers, and workers' representatives. This tripartite social dialogue has been a cornerstone of the process, but debates surrounding each increase often highlight tensions: employers express concerns about potential losses in competitiveness and profitability, while trade unions emphasize the importance of reducing in-work poverty and ensuring fair compensation. Our evidence suggests that both perspectives can be partly justified. Employers' concerns appear mitigated by the fact that firms did not reduce employment or experience profitability losses,

while workers' representatives are supported by the absence of contractionary effects and by signs of efficiency improvements that can benefit the economy more broadly.

For future policy interventions, the implication is that minimum wage adjustments can be a viable instrument of wage policy when accompanied by strong social dialogue and a predictable legislative framework. Maintaining transparency in the determination process, relying on empirical evidence, and ensuring that increases are phased in with sufficient predictability will be crucial for preserving firm adaptability while meeting workers' needs. Our findings further suggest that minimum wage increases may be embedded into a broader sustainable development strategy—particularly in economies with initially low wages—provided that the process remains inclusive of all social partners and is sensitive to the distribution of adjustment costs across sectors. In this sense, the reform experience in North Macedonia highlights the importance of reinforcing institutional mechanisms of dialogue, monitoring firm-level outcomes, and calibrating future wage floors in a way that balances equity with competitiveness.

*Table 1: Difference-in-Differences Estimates of the 2017 Minimum Wage Reform*

Dependent variable ( $y_{it}$ )	Estimate of policy effect ( $\beta_2$ )
Number of Employees	-0.421 (0.206)
Wages of other employees	-183.7 (614.2)
Productivity	0.0136 (0.129)
Other expenses (not related to wages)	-312.7* (97.9)
Profitability	-0.003 0.005

*Note: All models include period and firm fixed effects and control for the log of assets, revenues, expenditures, and number of employees (if they are not the dependent variable). \* denotes significance at 0.01.*

**Keywords:** *Minimum wage reform, Firm-level data, Labor productivity, Wage policy, North Macedonia*

**JEL classification:** *J31, D22, O15, L25*

## REFERENCES

- Draca, M., Machin, S. and Van Reenen, J. (2011), "Minimum wages and firm profitability", *American Economic Journal: Applied Economics*, Vol. 3 No. 1, pp. 129-151.
- Georgiadis, A., Kaplanis, I. and Monastiriotis, V. (2020), "Minimum wages and firm employment: Evidence from a minimum wage reduction in Greece", *Economics Letters*, Vol. 193.
- Jovanovikj, B. and Naumovski, N. (2021), "Minimum wage reform and firms' performance: Evidence from North Macedonia", *Working Paper No. 2*, National Bank of the Republic of North Macedonia.
- Jovanovikj, B., Tevdovski, D. and Stojkoski, V. (2021), "Minimum wage in the Republic of North Macedonia and its effects on the firms' operations", *Friedrich Ebert Stiftung*.

- Mayneris, F., Poncet, S. and Zhang, T. (2018), “Improving or disappearing: Firm-level adjustments to minimum wages in China”, *Journal of Development Economics*, Vol. 135, pp. 20-42.
- Petreski, M. and Pehkonen, J. (2024), “Minimum wage and manufacturing labor share: Evidence from North Macedonia”, *Ekonomiska Misao i Praksa*, Vol. 33 No. 2.
- Petreski, M., Mojsoska-Blazevski, N. and Ouchi, M. (2019), “The minimum wage as a wage equality policy: Evidence from North Macedonia”, *Economic Annals*, Vol. 64 No. 223, pp. 61-81.
- Saez, E., Schoefer, B. and Seim, D. (2019), “Payroll taxes, firm behavior, and rent sharing: Evidence from a young workers' tax cut in Sweden”, *American Economic Review*, Vol. 109 No. 5, pp. 1717-1763.
- Skedinger, P. (2014), “Effects of payroll tax cuts for young workers”, In Rosholm, M. and Svarer, M. (eds.) *Nordic Economic Policy Review: Consequences of Youth Unemployment and Effectiveness of Policy Interventions*, 1. Norden.
- Trenovski, B., Kozheski, K., Tashevska, B. and Peovski, F. (2021), “The minimum wage impact on labor productivity: The case of selected SEE countries”, *Management Research and Practice*, Vol. 11 No. 3, pp. 32-42.
- Yagan, D. (2015), “Capital tax reform and the real economy: The effects of the 2003 dividend tax cut”, *American Economic Review*, Vol. 105 No. 12, pp. 3531-3563.

## THE IMPACT OF ECONOMIC DEVELOPMENT ON ENVIRONMENTAL QUALITY THROUGH TESTING THE ENVIRONMENTAL KUZNETS CURVE HYPOTHESIS IN THE WESTERN BALKAN COUNTRIES

**Dorđe Kotarac**

*Faculty of Agriculture,  
University of Belgrade, Serbia  
djordje.kotarac@agrif.bg.ac.rs*

### **EXTENDED ABSTRACT**

**Purpose** The foundations of sustainable development theory promote economic growth within the limits of preserving the environmental conditions of countries (Lazăr *et al.*, 2019; Jones, 2022; Hatmanu and Cautisanu, 2023). This paper conducts a preliminary examination of the Environmental Kuznets Curve (EKC) by testing its hypothesized inverted U-shaped trajectory (Kisswani *et al.*, 2019; Ansari, 2022; Magazzino *et al.*, 2023). Specifically, the Kuznets hypothesis suggests that, in the short term, increased economic development leads to higher levels of environmental pollution (Mazur *et al.*, 2015; Shahbaz and Sinha, 2019; Simionescu, 2021). However, over the long run, improvements in environmental standards, the implementation of regulatory measures, and the adoption of cleaner technological and production processes contribute to a reduction in environmental degradation (Zhang *et al.*, 2017; Shuai *et al.*, 2017; Kuznets, 2019; Jóźwik *et al.*, 2021). This analysis includes five Western Balkan countries that have achieved candidate status for European Union membership in the post-2000 period. Empirical tests for co-integration among the variables confirmed the existence of a long-run relationship between economic growth and environmental quality. In a study conducted by Armeanu *et al.* (2018), the relationship between variables measuring environmental pollution and the economic growth rates of the EU-27 countries from 1995 to 2014 was empirically tested. The results of the research confirm the Environmental Kuznets Curve (EKC) hypothesis, indicating the existence of a long-term effect in the case of sulphur dioxide (SO<sub>2</sub>) emissions and non-methane volatile organic compounds (NMVOCs). Another significant research in the European context was conducted by Vasylieva *et al.* (2019), examining the relationship between economic growth rate, renewable resource allocation, and greenhouse gas (GHG) emissions during the period 2000-2016. Using FMOLS and DOLS panel estimation techniques, the study confirms the Environmental Kuznets Curve (EKC) hypothesis. Additionally, Dogan and Inglesi-Lotz (2020), conducted a study covering the period from 1980 to 2014, analysing the impact of industrial activity and energy consumption on the increase in carbon dioxide (CO<sub>2</sub>) emissions in the EU-27 countries. The findings of this research indicate a long-term decline in (CO<sub>2</sub>) emissions as a result of more efficient energy use and the adoption of cleaner technologies. The purpose of this paper is to conduct an empirical analysis of the short- and long-term effects of economic growth in Western Balkan countries on the quality of their natural environment. Based on the available data, the study separately examines the impact of economic growth on carbon dioxide (CO<sub>2</sub>) emission levels, in contrast to the effect of economic growth on the increase in total greenhouse gas (GHG) emissions. The Environmental Kuznets Curve hypothesis is tested using data from five Western Balkan countries.

To examine the effect of economic growth on environmental quality and the use of renewable energy in five Western Balkan countries, the following three hypotheses will be tested: H1:

Increases in economic growth rates and per capita energy consumption have a short-term effect on CO<sub>2</sub> and GHG emission levels, as indicators of environmental quality in Western Balkan countries.

H2: Growth in economic activity and total per capita energy consumption exerts a long-term influence on changes in CO<sub>2</sub> and GHG emissions in Western Balkan countries.

H3: Rising economic growth rates and per capita energy consumption lead to a short-term decline in environmental quality, whereas, in the long term, they contribute to improvements in environmental quality.

**Design/methodology/approach** The methodological section of this research focuses on examining the existence of both short- and long-run effects of economic growth and per capita energy consumption on the variation in carbon dioxide (CO<sub>2</sub>) and greenhouse gas (GHG) emissions. The research covers the period from 2000 to 2023 and includes a sample of five Western Balkan countries-candidates for accession to the European Union (EU-27). The analysis explores links between economic growth, energy use, and environmental degradation using panel data methods. To implement the methodological part and ensure the robustness of the empirical results, the stationarity of the time series is tested using unit root tests. Specifically, the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP). The Johansen co-integration test and Vector Error Correction model are applied to determine the order of co-integration of the variables. Prior to conducting the co-integration tests, the optimal lag length is selected based on the various selection criteria: Akaike Information Criterion (AIC), the Schwarz Information Criterion (SIC), and the Hannan-Quinn Information Criterion (HQIC). The study employs the VAR Lag Order Selection Criteria, and based on the obtained results, the existence of a co-integration relationship among the time series is subsequently assessed.

**Findings** The research findings confirm the hypothesis regarding the impact of economic growth and per capita energy consumption on carbon dioxide (CO<sub>2</sub>) emissions and greenhouse gas (GHG) levels over the long term. The results suggest the existence of a long-run relationship between the examined variables across all five Western Balkan countries, including North Macedonia, where a statistically significant effect of the economic growth rate and per capita energy consumption on (GHG) emissions was confirmed. On the other hand, the hypothesis concerning the short-run relationship among the included variables was validated in three Western Balkan countries - Serbia, North Macedonia, and Bosnia and Herzegovina. Finally, the third hypothesis, which pertains to the Environmental Kuznets Curve (EKC) and its inverted Ushape, was empirically supported in the case of Serbia.

**Originality/value** Sustainable economic development is a central topic in current economic research and a core part of the European Union's development strategies: Europe 2020, European Green Deal (Fetting, 2020; Wolf *et al.*, 2021), Agenda for Sustainable Development (Lee *et al.*, 2016; Delbeke *et al.*, 2019; Kryk and Guzowska, 2021; Burgin, 2023). The originality of this paper lies in assessing the significance of the impact of economic growth on environmental conditions through the application of econometric methods, specifically the Vector Error Correction Model. The value of this research stems from analysing the impact of greenhouse gas emissions on economic growth rates – a dimension that previous studies have primarily examined through the impact of carbon dioxide emissions alone. A key limitation of the study concerns the scope of the dataset, particularly the omission of the pre-2000 period. As a recommendation for future research, a comparative analysis between Balkan countries

and a group of EU-27 members is suggested to evaluate differences and similarities in the progress of sustainable development policies across different regions of Europe.

**Keywords:** Sustainable development, Kuznets (EKC) hypothesis, CO<sub>2</sub>/GHG emissions, Panel data methods, Western Balkan countries

**JEL classification:** Q01, O44, C50

## REFERENCES

- Ansari, M. A. (2022), “Re-visiting the Environmental Kuznets curve for ASEAN: A comparison between ecological footprint and carbon dioxide emissions”, *Renewable and Sustainable Energy Reviews*, Vol. 168, pp. 112867.
- Apergis, N., Christou, C. and Gupta, R. (2017), “Are there environmental Kuznets curves for US state-level CO<sub>2</sub> emissions?”, *Renewable and Sustainable Energy Reviews*, Vol. 69, pp. 551558.
- Androniceanu, A. and Georgescu, I. (2023), “The impact of CO<sub>2</sub> emissions and energy consumption on economic growth: a panel data analysis”, *Energies*, Vol. 16 No. 3, pp. 1342.
- Armeanu, D., Vintilă, G., Andrei, J. V., Gherghina, Ș. C., Drăgoi, M. C. and Teodor, C. (2018), “Exploring the link between environmental pollution and economic growth in EU-28 countries: Is there an environmental Kuznets curve?”, *PloS one*, Vol. 13, No. 5, pp. e0195708.
- Bürgin, A. (2023), *The European Commission: A climate policy entrepreneur*. In Handbook on European Union climate change policy and politics, Edward Elgar Publishing, pp. 23-37.
- Churchill, S. A., Inekwe, J., Ivanovski, K. and Smyth, R. (2018), “The environmental Kuznets curve in the OECD: 1870–2014”, *Energy Economics*, Vol. 75, pp. 389-399.
- Dogan, E. and Inglesi-Lotz, R. (2020), “The impact of economic structure to the environmental Kuznets curve (EKC) hypothesis: evidence from European countries”, *Environmental Science and Pollution Research*, Vol. 27, pp. 12717-12724.
- Delbeke, J., Runge-Metzger, A., Slingenberg, Y. and Werksman, J. (2019), “The Paris agreement”, In *Towards a climate-neutral Europe* (pp. 24-45), Routledge.
- Destek, M. A., Ulucak, R. and Dogan, E. (2018), “Analyzing the environmental Kuznets curve for the EU countries: the role of ecological footprint”, *Environmental Science and Pollution Research*, Vol. 25 No. 29, pp. 29387-29396.
- Dar, A. A., Hameed, J., Huo, C., Sarfraz, M., Albasher, G., Wang, C. and Nawaz, A. (2022), “Recent optimization and panelizing measures for green energy projects; insights into CO<sub>2</sub> emission influencing to circular economy”, *Fuel*, Vol. 314, pp. 123094.
- Eriksson, C. (2013), *Economic Growth and the Environment: An Introduction to the Theory*, Oxford University Press, USA.
- Fetting, C. (2020), The European green deal, *ESDN report*, December, Vol. 2 No. 9, p. 53.
- Htike, M. M., Shrestha, A. and Kakinaka, M. (2021), “Investigating whether the environmental Kuznets curve hypothesis holds for sectoral CO<sub>2</sub> emissions: evidence from developed and developing countries”, *Environment, Development and Sustainability*, Vol. 24, pp. 1271212739.

- Hatmanu, M. and Cautisanu, C. (2023), “Investigating the relationships between economic growth and environmental degradation: Evidence from EU15 countries”, *Technological and Economic Development of Economy*, Vol. 29 No. 1, pp. 192-216.
- Huynh, C. M. (2024), “Economic freedom, economic development and income inequality in Asia: an analysis from the Kuznets curve perspective”, *Journal of the Asia Pacific Economy*, Vol. 29, No. 2, pp. 443-462.
- Jones, C. I. (2022), “The past and future of economic growth: A semi-endogenous perspective”, *Annual Review of Economics*, Vol. 14, pp. 125-152.
- Jóźwik, B., Gavryshkiv, A. V., Kyophilavong, P. and Gruszecki, L. E. (2021), “Revisiting the environmental Kuznets curve hypothesis: A case of Central Europe”, *Energies*, Vol. 14 No. 12, pp. 3415.
- Kuznets, S. (2019), “Economic growth and income inequality”, In *The Gap between Rich and Poor* (pp. 25-37), Routledge.
- Kryk, B. and Guzowska, M. K. (2021), “Implementation of climate/energy targets of the Europe 2020 strategy by the EU member states”, *Energies*, Vol. 14 No. 9, pp. 2711.
- Kisswani, K. M., Harraf, A. and Kisswani, A. M. (2019), “Revisiting the environmental Kuznets curve hypothesis: evidence from the ASEAN-5 countries with structural breaks”, *Applied Economics*, Vol. 51 No. 17, pp. 1855-1868.
- Lee, B. X., Kjaerulf, F., Turner, S., Cohen, L., Donnelly, P. D., Muggah, R., ... and Gilligan, J. (2016), “Transforming our world: implementing the 2030 agenda through sustainable development goal indicators”, *Journal of Public Health Policy*, Vol. 37 (Suppl 1), pp. 13-31.
- Lazăr, D., Minea, A. and Purcel, A. A. (2019), “Pollution and economic growth: Evidence from Central and Eastern European countries”, *Energy Economics*, Vol. 81, pp. 1121-1131.
- Mazur, A., Phutkaradze, Z. and Phutkaradze, J. (2015), “Economic growth and environmental quality in the European Union countries—is there evidence for the environmental Kuznets curve”, *International Journal of Management and Economics*, Vol. 45 No. 1, pp.108-126.
- Marinaş, M. C., Dinu, M., Socol, A. G. and Socol, C. (2018), “Renewable energy consumption and economic growth. Causality relationship in Central and Eastern European countries”, *PloS One*, Vol. 13 No. 10, pp. e0202951.
- Magazzino, C., Gallegati, M. and Giri, F. (2023), “The Environmental Kuznets Curve in a longterm perspective: Parametric vs semi-parametric models”, *Environmental Impact Assessment Review*, Vol. 98, 106973.
- Narayan, P. K., Saboori, B. and Soleymani, A. (2016), “Economic growth and carbon emissions”, *Economic Modelling*, Vol. 53, pp. 388-397.
- Simionescu, M. (2021), “The nexus between economic development and pollution in the European Union new member states. The role of renewable energy consumption”, *Renewable Energy*, Vol. 179, pp. 1767-1780.
- Shuai, C., Chen, X., Shen, L., Jiao, L., Wu, Y. and Tan, Y. (2017), “The turning points of carbon Kuznets curve: Evidences from panel and time-series data of 164 countries”, *Journal of Cleaner Production*, Vol. 162, pp. 1031-1047.
- Shahbaz, M. and Sinha, A. (2019), Environmental Kuznets curve for CO2 emissions: A literature survey. *Journal of Economic Studies*, Vol. 46 No. 1, pp. 106-168.
- Simionescu, M., Păuna, C. B. and Niculescu, M. D. V. (2021), “The relationship between economic growth and pollution in some new European union member states: A dynamic panel ARDL approach”, *Energies*, Vol. 14 No.9, pp. 2363.

- Vasylieva, T., Lyulyov, O., Bilan, Y. and Streimikiene, D. (2019), “Sustainable economic development and greenhouse gas emissions: The dynamic impact of renewable energy consumption, GDP, and corruption”, *Energies*, Vol. 12 No. 17, pp. 3289.
- Wang, Q., Wang, X. and Li, R. (2022), “Does urbanization redefine the environmental Kuznets curve? An empirical analysis of 134 Countries”, *Sustainable Cities and Society*, Vol. 76, pp. 103382.
- Wolf, S., Teitge, J., Mielke, J., Schütze, F. and Jaeger, C. (2021), “The European green deal- more than climate neutrality”, *Intereconomics*, Vol. 56 No. 2, pp. 99-107.
- Younsi, M. and Bechtini, M. (2020), “Economic growth, financial development, and income inequality in BRICS countries: does Kuznets’ inverted U-shaped curve exist?”, *Journal of the Knowledge Economy*, Vol. 11 No. 2, pp.721-742.
- Yang, X., Lou, F., Sun, M., Wang, R. and Wang, Y. (2017), “Study of the relationship between greenhouse gas emissions and the economic growth of Russia based on the Environmental Kuznets Curve”, *Applied Energy*, Vol. 193, pp. 162-173.
- Zhang, C., Wang, Y., Song, X., Kubota, J., He, Y., Tojo, J. and Zhu, X. (2017), “An integrated specification for the nexus of water pollution and economic growth in China: Panel cointegration, long-run causality and environmental Kuznets curve”, *Science of the Total Environment*, Vol. 609, pp. 319-328.

## **LIFE EXPECTANCY AND INCOME INEQUALITY IN NORTH MACEDONIA: AN EMPIRICAL ANALYSIS**

**Marijana Cvetanoska**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*marijana.cvetanoska@eccf.ukim.edu.mk*

**Biljana Tashevaska**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*biljana.tashevaska@eccf.ukim.edu.mk*

**Predrag Trpeski**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*predrag.trpeski@eccf.ukim.edu.mk*

### **EXTENDED ABSTRACT**

**Purpose** The aim of this study is to examine the relationship between income inequality, measured by the Gini coefficient, and life expectancy in North Macedonia, applying the Autoregressive Distributed Lag (ARDL) bounds testing approach, using annual data from 2000 to 2022. The analysis includes GDP per capita, education (primary school enrolment), health expenditures, and hospital beds as control variables, all of which affect life expectancy by shaping living standards, knowledge and behaviors, access to health care, and the efficiency of medical services. Income inequality and health outcomes are central aspects of economic and social development. High income inequality has been linked to disparities in health and life expectancy through differences in healthcare access, education, and social cohesion (Wilkinson and Pickett, 2009; Deaton, 2024). The Global Risks Report 2023 highlights that global crises, such as the COVID-19 pandemic and geopolitical tensions, exacerbate economic inequalities, further affecting healthcare access and social outcomes (World Economic Forum, 2023). At the macro level, maintaining and improving population health is recognized as a key policy for sustainable development (Bayati *et al.*, 2013). Life expectancy at birth is a widely used measure of population health and longevity, and, although health is multi-dimensional, it is commonly adopted as a proxy for overall population health (Rabbi Fazle, 2013; Sharma, 2018). This study presents empirical findings that are rarely discussed in existing research and offers valuable guidance for policies aimed at enhancing health outcomes and reducing inequality, which in turn can strengthen human capital and promote sustainable development in North Macedonia.

**Methodology** This study investigates the impact of socio-economic factors on life expectancy in North Macedonia using time-series data. To analyze both short-run and long-run dynamics, the Autoregressive Distributed Lag (ARDL) Bounds modeling approach was employed. The ARDL Bounds approach is a technique for examining linear cointegration, developed by Pesaran (1997), Pesaran and Smith (1998), and Pesaran *et al.* (2001). This method offers several advantages over other cointegration techniques, including the ability to estimate both short-run and long-run parameters within a single model, addressing endogeneity issues, and testing for cointegration among variables with different orders of integration, such as I(0) and I(1). Moreover, it is well-suited for small sample sizes (Jebli and Belloumi, 2017) and provides unbiased long-run estimates (Odhiambo, 2008).

The analysis is based on annual data for North Macedonia covering the period 2000–2022. All variables are obtained from the World Development Indicators of the World Bank database. Specifically, the Gini index is used as a measure of income inequality, while life expectancy at birth (in years) is employed as the dependent variable. Additional control variables include GDP per capita (constant 2015 US\$), primary school enrolment (% of gross), health expenditures (% of GDP), and hospital beds (per 1,000 people). The selection of variables that determine life expectancy is guided by insights from previous research as well as our preliminary estimations. GDP per capita is included as a control variable to account for the general level of economic development, as previous research (Neumayer and Plümer, 2016) indicates that higher income levels are associated with longer life expectancy. Primary school enrolment is incorporated as a control variable to reflect the influence of education on life expectancy, since education enhances health awareness, income, and overall social well-being, all of which contribute to improved health outcomes and increased longevity (Kabir, 2008). Two health supply indicators (health expenditure and the number of hospital beds) are included, given the well-established connection between health care factors and health outcomes (Blazquez-Fernández *et al.*, 2017).

The empirical analysis followed a structured approach to ensure the reliability and validity of the results. Initially, all variables were tested for unit roots to determine their order of integration, ensuring that none were integrated of order two (I(2)), which is a prerequisite for applying the ARDL approach. The Lee-Strazicich LM unit root test was employed to account for potential structural breaks, as conventional tests may produce biased results when such breaks are present, consistent with Perron's (1989) findings on the importance of considering structural changes in time series analysis. Subsequently, several diagnostic tests were conducted to validate the assumptions of the regression model: normality of residuals, serial correlation, heteroskedasticity and the stability of the estimated model. To explore long-run relationships among the variables, the ARDL Bounds Test for cointegration was applied, allowing the identification of a stable equilibrium even with a small sample by comparing the calculated F-statistic to critical bounds for I(0) and I(1) variables. The ARDL model was then reparametrized into an Error Correction Model (ECM) form to distinguish short-run dynamics from long-run equilibrium relationships, with the ECM term capturing the speed of adjustment following deviations from equilibrium. Finally, a robustness check was performed using the Fully Modified Ordinary Least Squares (FMOLS) method, which corrects for potential endogeneity and serial correlation, providing reliable long-run coefficient estimates.

**Findings** First of all, the Lee-Strazicich LM unit root test was employed, and the results are summarized as follows. The test for the *Gini index* indicates the presence of a structural break in 2017. The minimum test statistic is statistically significant ( $-2.36 > -3.49^{*1}$ ), suggesting that the series becomes stationary after first difference. For *life expectancy*, structural breaks were identified in 2018/2019. The findings reveal that the series does not follow a classical unit root process but rather attains stationarity in levels with a structural break ( $-3.35 > -3.49^{*}$ ). In the case of *GDP per capita*, breaks were detected in 2006 and 2017. The test results indicate that the series is stationary in the presence of structural breaks. Given the trending nature of GDP per capita, the Lee-Strazicich test with breaks is considered a more robust approach in this context ( $-2.30 > -4.91^{*}$ ). The *education variable* showed a structural break around 2016/2017. The evidence suggests that the variable is stationary in levels ( $-3.09 > -3.49^{*}$ ). For the *health expenditure* indicator, structural breaks were identified in 2011 and 2019. The test results indicate that the series is stationary in levels ( $-2.62 > -3.49^{*}$ ). Finally, for the *number of hospital*

---

<sup>1</sup> \* critical value 5%

beds, a structural break was observed in 2018, and the series is confirmed to be stationary in levels (I(0)) (-1.59>-3.49\*).

Diagnostic tests showed that the model is free from autocorrelation and heteroscedasticity, the errors are normally distributed, the functional form is properly specified, and the coefficients of the model remain stable. Robustness check using FMOLS showed that GDP per capita positively affects life expectancy, while the Gini index has a marginal negative effect, and education and healthcare expenditures behave as expected. High R<sup>2</sup> and adjusted R<sup>2</sup> values (0.86 and 0.81, respectively) confirm that the model explains a substantial portion of the variation in the dependent variable and demonstrate the robustness of the results.

The ARDL Bounds Test confirms the presence of a significant long-run cointegrating relationship, indicating that the variables share a stable equilibrium over time. Given that the F-statistic is well above the I(1) bounds, the null hypothesis of no long-run relationship can be rejected. According to the results (Table 1), the effect of LIFE\_EXP lag value (-1), DGINI and its lag value (-1), GDP\_PC and its lag value (-1), HEALTH and its lag value (-1), and BEDS and its lag value (-1) on LIFE\_EXP is statistically significant in the short run.

Table 1. Short-run and long-run coefficients

Short-term coefficients and ERC model				
Variable	Coefficient	Standard error	t-statistics	Probability
LIFE_EXP(-1)	-0.299745*	0.097982	-3.059197	0.0156
DGINI	-0.130299*	0.016521	-7.886896	0.0000
DGINI(-1)	-0.173871*	0.017003	-10.225650	0.0000
GDP_PC	0.004271*	0.000274	15.567030	0.0000
GDP_PC(-1)	-0.004152*	0.000267	-15.573520	0.0000
EDU	0.000648	0.012487	0.051909	0.9599
HEALTH	0.790743*	0.075658	10.451530	0.0000
HEALTH(-1)	-0.468239*	0.074928	-6.249180	0.0002
BEDS	-8.226424*	0.663593	-12.396790	0.0000
BEDS(-1)	1.613360*	0.458582	3.518150	0.0079
COINTEQ(-1)	-1.299745 *	0.037341	-34.807730	0.0000
Long-term term coefficients				
Variable	Coefficient	Standard error	t-statistics	Probability
DGINI	-0.234023*	0.019209	-12,18317	0.0000
GDP_PC	0.000091	0.000158	0.574961	0.5811
EDU	0.000499	0.009601	0.051943	0.9598
HEALTH	0.248128*	0.042362	5.857385	0.0004
BEDS	-5.087968*	0.706798	-7.198616	0.0001
C	94.493740	3.986370	23.704210	0.0000

\* Indicates stationarity at the 1% and 5% significance levels, respectively

(Source: Authors' calculations in EViews)

Increases in health expenditures and GDP per capita positively influence life expectancy, whereas increases in income inequality and the number of hospital beds are associated with reductions in life expectancy. Education does not exert a statistically significant effect on life expectancy within the model. One possible explanation is that improvements in education typically generate long-term benefits rather than immediate outcomes.

In the long run, the effect of DGINI, HEALTH, and BEDS on LIFE\_EXP is statistically significant, indicating that changes in these variables have a lasting impact on life expectancy over time. The long-run estimation results reveal that income inequality has a strong negative and highly significant impact on life expectancy, confirming that greater inequality reduces longevity. Health expenditure is positively and significantly associated with life expectancy, emphasizing the importance of investments in healthcare. Conversely, GDP per capita and education show positive but statistically insignificant effects, suggesting that their influence may be indirect or delayed. Interestingly, the number of hospital beds has a significant negative effect, which may reflect inefficiencies in healthcare delivery or the fact that higher bed availability often corresponds to poorer underlying health conditions rather than improved outcomes. Overall, inequality reduction and effective healthcare spending emerge as the key long-term drivers of life expectancy.

The ECM coefficient (COINTEQ=-1.29) indicates rapid adjustment toward the long-run equilibrium following short-term shocks, reflecting that after a short-term shock, life expectancy at birth adjusts rapidly toward the long run but may temporarily oscillate around the equilibrium. This highlights the system's strong tendency to return to equilibrium after short-term fluctuations.

**Originality/value** The findings highlight key determinants of life expectancy in North Macedonia, showing that while short-term changes in inequality, GDP per capita, health spending, and healthcare infrastructure affect outcomes, the system self-corrects toward long-run equilibrium. Income inequality and health expenditure are the most influential factors that improve life expectancy at birth, whereas GDP per capita, education, and hospital beds capacity have limited long-term effects, emphasizing the need for policies targeting inequality reduction and efficient healthcare investment. Based on existing studies, while some research has examined life expectancy trends and regional factors in North Macedonia or its healthcare system development (Miladinov, 2020; Saliu et al., 2022; Milevska Kostova et al., 2024), there is a clear lack of comprehensive analyses investigating the combined short-term and long-term effects of socio-economic determinants, such as income inequality, GDP per capita, education, and health expenditures. This gap highlights the novelty of the present study, which systematically explores these factors using an ARDL approach to uncover both short-run and long-run influences on life expectancy.

**Keywords:** *ARDL bounds testing approach, Income inequality, Life expectancy, North Macedonia*

**JEL classification:** *C32, I14, I15, O52*

## REFERENCES

- Bayati, M., Akbarian, R. and Kavosi, Z. (2013), "Determinants of life expectancy in Eastern Mediterranean Region: A health production function", *International Journal of Health Policy and Management*, Vol. 1, pp. 1-7.
- Blazquez-Fernández, C., Cantarero-Prieto, D. and Pascual-Saez, M. (2017), "Health expenditure and socio-economic determinants of life expectancy in the OECD Asia/Pacific area countries", *Applied Economics Letters*, Vol. 24 No. 3, pp. 167-169.
- Deaton, A. (2024), *The Great Escape: Health, Wealth, and the Origins of Inequality*, Princeton University Press, Princeton, NJ, USA.

- Jebli, M.B. and Belloumi, M. (2017), "Investigation of the causal relationships between combustible renewables and waste consumption and CO<sub>2</sub> emissions in the case of Tunisian maritime and rail transport", *Renewable and Sustainable Energy Reviews*, Vol. 71, pp. 820-829.
- Kabir, M. (2008), "Determinants of life expectancy in developing countries", *The Journal of Developing Areas*, Vol. 41 No. 2, pp. 185-204.
- Miladinov, G. (2020), "Socioeconomic development and life expectancy relationship: Evidence from the EU accession candidate countries", *Genus*, Vol. 76 No. 2, pp. 1-20.
- Milevska Kostova, N., Atanasova, S., Dimkovski, V., Stevanovski, G., Stæhr Johansen, A., Winkelmann, J. and Rechel, B. (2024), "North Macedonia: Health system review", *Health Systems in Transition*, Vol. 26 No. 6, pp. 1-164.
- Neumayer, E. and Plümper, T. (2016), "Inequalities of income and inequalities of longevity: A cross-country study", *American Journal of Public Health*, Vol. 106 No. 1, pp. 160-165.
- Odhiambo, N.M. (2008), "Financial depth, savings and economic growth in Kenya: A dynamic causal linkage", *Economic Modelling*, Vol. 25 No. 4, pp. 704-713.
- Perron, P. (1989), "The great crash, the oil price shock, and the unit root hypothesis", *Econometrica*, Vol. 57 No. 6, pp. 1361-1401.
- Pesaran, M.H. (1997), "The role of economic theory in modelling the long run", *The Economic Journal*, Vol. 107 No. 440, pp. 178-191.
- Pesaran, M.H., Shin, Y. and Smith, R.J. (2001), "Bounds testing approaches to the analysis of level relationships", *Journal of Applied Econometrics*, Vol. 16 No. 3, pp. 289-326.
- Pesaran, M.H. and Smith, R.P. (1998), "Structural analysis of cointegrating VARs", *Journal of Economic Surveys*, Vol. 12 No. 5, pp. 471-505.
- Rabbi Fazle, A.M. (2013), "Imbalance in life table: Effect of infant mortality on lower life expectancy at birth", *Journal of Scientific Research*, Vol. 5 No. 3, pp. 479-488.
- Saliu, H.R., Rusi, M. and Sadiki, S. (2022), "Lifestyle as a factor in longevity in the Pollog region", *Innovative Systems Design and Engineering*, Vol. 12 No. 6, pp. 52-57.
- Sharma, R. (2018), "Health and economic growth: Evidence from dynamic panel data of 143 years", *PLOS ONE*, Vol. 13 No. 10, e0204940.
- Wilkinson, R.G. and Pickett, K. (2009), *The Spirit Level: Why More Equal Societies Almost Always do Better*, Allen Lane, London, UK.
- World Economic Forum (2023), *The Global Risks Report 2023*, World Economic Forum.

## **THE IMPACT OF REMITTANCES ON DOMESTIC INVESTMENT AND CONSUMPTION EXPENDITURES: THE CASE OF NORTH MACEDONIA**

**Bojan Shimbov**

*University Jaume I, Castellon  
shimbov@uji.es*

**Oliver Morrissey**

*University of Nottingham  
oliver.morrissey@nottingham.ac.uk*

**Maite Alguacil**

*University Jaume I, Castellon  
alguacil@uji.es*

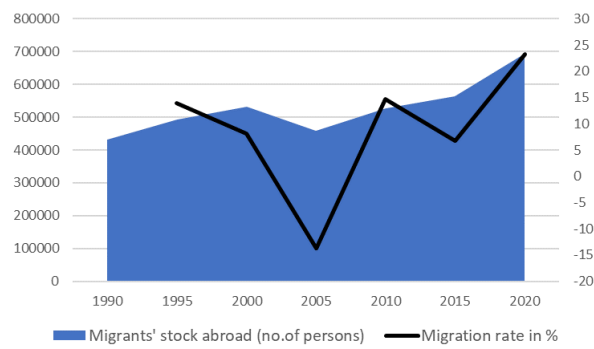
### ***EXTENDED ABSTRACT***

**Purpose** Remittances are a significant source of household income and foreign exchange inflows for the Republic of North Macedonia (MKD) due to the relatively large share of workers who have emigrated, mostly to the EU. This paper analyses data from a survey of almost 2000 households that receive remittances to investigate how households use the income. Although remittances are mostly used for consumption spending, a specific focus of the analysis is on the almost 20% of households that invest or save the money. What are the characteristics of these households, and the determinants of the amount and share invested or saved? The paper then estimates the macroeconomic impact of remittances in terms of the effect on consumption, investment, and saving.

**Design/methodology/approach** This paper makes use of a survey of almost 2000 households receiving remittances in North Macedonia (MKD) in 2021 to empirically investigate how the income is used and if this is related to characteristics of the migrants and households. Unfortunately, the survey was only for receiving households (it is not possible to compare with non-receiving households), data on household characteristics are limited, and it is only a cross-section (although some retrospective questions were included). The decision of how to use remittances received is analyzed by a set of discrete choice models. On the one hand, we have considered the characteristics of the households receiving the remittances (family income, employment, age, gender, ethnicity, settlement type, place of residence), and on the other hand, we have considered factors like: i) remittance timing, ii) trends, and iii) migrant's home country attachment. For robustness checks, we have used PPLM regressions like  $I/S/C (\%)*total\ RR$  and the Zero-inflated Model.

**Findings** North Macedonia is facing a growing trend of migration, which, if not addressed, can potentially cause significant socio-economic disruptions in the country going forward. The stock of migrant population to total population ratio stood at 37 percent in 2020, which is the third highest in the region after Albania and Bosnia and Herzegovina, and well above the average for developing countries. Most of the migrants go to the EU (mostly Germany and Italy) or Switzerland, as well as to Turkey. Finally, migrants from North Macedonia tend to be young people in their most productive years, which may pose serious challenges to development going forward.

Figure 1: Stock of migrants from North Macedonia and migration rate



Note: data refers to the foreign-born population  
(Source: United Nations Population Division and authors' calculations)

Figure 2: First-time residence permits issued by the EU-27 countries

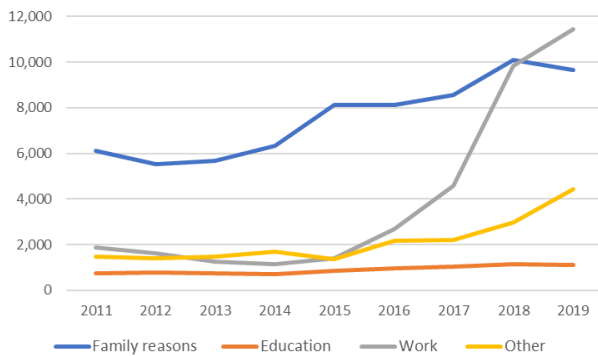
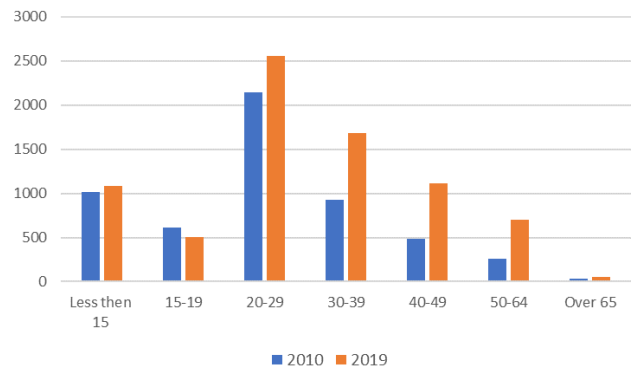


Figure 3: Age structure of immigration from North Macedonia to the EU, by age groups



Note: data for 2020-2021 is available but is not included because of distortions due to the COVID-19 crisis  
(Source: Eurostat and author's own calculations)

This paper analyses data from a survey in 2022 of 2000 households that receive remittances to investigate how households use the income. Although remittances are mostly used for consumption spending, a specific focus of the analysis is on the almost 20% of households that invest or save the money. We try to see what the characteristics of these households and the determinants of the amount and share are invested or saved. Using several specification models and considering the heterogeneity of recipient households, our findings indicate that remittances can significantly impact the financial structure of Macedonian families, suggesting the potential for a broader macroeconomic effect as well (ongoing research). This effect is likely to be stronger when families expect sustained or increasing remittance flows and when the sender has a deep connection to their country of origin.

Table 1: Probit regressions. Binary dependent variable: Use of RR

VARIABLES	(1) Investment	(2) Saving	(3) Consumption	VARIABLES	(1) Investment	(2) Saving	(3) Consumption
log(total RR)	0.537*** (0.000)	0.373*** (0.000)	-0.098* (0.077)	RR as % of HH mid-income	0.031*** (0.000)	0.017*** (0.000)	-0.005 (0.171)
Gender (Female = 0)	-0.177** (0.032)	0.165* (0.062)	0.120 (0.274)	Gender	-0.181** (0.023)	0.139 (0.105)	0.137 (0.208)
Age	0.124*** (0.000)	0.034 (0.297)	-0.130*** (0.005)	Age	0.139*** (0.000)	0.038 (0.221)	-0.109** (0.016)
Family income	0.055 (0.128)	0.120*** (0.002)	-0.205*** (0.000)	Settlement (Small town/Village = 0)	0.266*** (0.002)	0.377*** (0.000)	-0.194* (0.091)
Settlement (Small town/Village = 0)	0.190** (0.031)	0.291*** (0.001)	-0.118 (0.297)	Ethnicity (Other = 0)			
Ethnicity (Other = 0)				- Macedonian	0.006 (0.971)	-0.179 (0.299)	-0.009 (0.966)
- Macedonian	-0.027 (0.879)	-0.223 (0.219)	0.082 (0.701)	- Albanian	-0.032 (0.858)	-0.366** (0.043)	0.482** (0.035)
- Albanian	-0.151 (0.412)	-0.493*** (0.009)	0.511** (0.028)	Employment (Not employed = 0)			
Employment (Not employed = 0)				- Employed	0.315*** (0.002)	0.026 (0.806)	-0.341** (0.034)
- Employed	0.218* (0.051)	-0.112 (0.340)	-0.155 (0.356)	- Self-employed	0.447** (0.014)	0.097 (0.629)	-0.190 (0.457)
- Self-employed	0.392** (0.036)	-0.020 (0.927)	-0.027 (0.917)	- Unemployed	-0.078 (0.527)	-0.313** (0.014)	0.053 (0.779)
- Unemployed	0.058 (0.632)	-0.199 (0.116)	-0.038 (0.842)	RR_year (From 1918 to 2009 = 0)			
RR_year (From 1981 to 2009 = 0)				- From 2010 to 2016	-0.295*** (0.003)	-0.006 (0.954)	-0.072 (0.621)
- From 2010 to 2016	-0.263** (0.011)	0.039 (0.726)	-0.085 (0.574)	- Since 2017 on	-0.194* (0.065)	-0.175 (0.148)	-0.200 (0.166)
- Since 2017 on	-0.103 (0.345)	-0.131 (0.288)	-0.238 (0.117)	RR_trend (No change = 0)			
RR_trend (No change = 0)				- Increased	0.363*** (0.001)	-0.036 (0.794)	0.123 (0.453)
- Increased	0.310*** (0.007)	-0.110 (0.438)	0.137 (0.421)	- Decreased	-0.032 (0.727)	0.124 (0.171)	0.126 (0.317)
- Decreased	-0.027 (0.774)	0.124 (0.179)	0.121 (0.345)	Freq_visit (Several times/year = 0)			
Freq_visit (Several times/year = 0)				- Once a year	-0.313*** (0.000)	-0.056 (0.534)	0.101 (0.392)
- Once a year	-0.256*** (0.004)	0.014 (0.877)	0.100 (0.405)	- Once every few years	-0.256** (0.032)	-0.191 (0.168)	0.128 (0.432)
- Once every few years	-0.150 (0.242)	-0.096 (0.507)	0.119 (0.476)				

Table 2: Zero-inflated ordered probit (3 categories)

Table 12: Zero-inflated ordered probit (3 categories)				
VARIABLES	(1)	(2)	(3)	(4)
	Investment	Investment	Saving	Saving
RR_income	0.002 (0.620)	0.065*** (0.000)	-0.012** (0.035)	0.056*** (0.009)
gender	-0.491*** (0.000)		0.167 (0.203)	
age	0.230*** (0.000)		0.025 (0.737)	
Ethnicity (Other = 0)				
- Macedonian	0.520* (0.055)		-0.696* (0.081)	
- Albanian	0.386 (0.189)		-1.300*** (0.005)	
Employment (Not employed = 0)				
- Employed	0.806*** (0.000)		0.154 (0.504)	
- Self-employed	1.702*** (0.000)		-0.205 (0.502)	
- Unemployed	0.238 (0.182)		-0.325 (0.189)	
RR_year (From 1981 to 2009 = 0)				
- From 2010 to 2016	-0.554*** (0.000)		0.114 (0.514)	
- Since 2017 on	-0.243 (0.170)		-0.109 (0.648)	
Freq_visit (Several times/year = 0)				
- Once a year	-0.395*** (0.004)		0.049 (0.739)	
- Once every few year	-0.439** (0.035)		-0.174 (0.501)	
Region				
- East	0.714 (0.115)	-0.209 (0.596)	0.229 (0.878)	-2.367* (0.078)
- Southwest	0.244 (0.410)	-0.229 (0.503)	0.488 (0.214)	-1.559 (0.267)
- Southeast	0.300 (0.444)	-0.646* (0.068)	1.116*** (0.000)	-1.285 (0.380)
- Pelagonia	0.231 (0.635)	-0.873** (0.026)	1.820** (0.037)	-2.985* (0.055)
- Polog	0.001 (0.997)	-0.026 (0.940)	0.745 (0.127)	-1.599 (0.230)
- Northeast	-0.139 (0.670)	-0.713** (0.042)	2.571*** (0.000)	-2.144 (0.183)
- Skopje	0.638* (0.054)	-0.748** (0.028)	1.262** (0.012)	-2.657* (0.071)
<i>inflate</i>				
RR_trend (No change = 0)				
- Increased		0.449*** (0.005)		0.017 (0.929)
- Decreased		-0.029 (0.813)		0.253** (0.042)
Settlement		0.378*** (0.002)		0.502*** (0.000)

cut1	0.699 (0.131)	-0.178 (0.801)
cut2	1.565 (0.001)	0.913 (0.116)
cut3	2.408*** (0.000)	1.682*** (0.003)
Constant	-0.820** (0.017)	0.546 (0.706)
Observations	1,599	1,599

Robust pval in parentheses  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Originality/value** To the best of our knowledge, remittances and their effects, both from consumption and savings and investment, have been very scarcely analyzed in the case of North Macedonia, so this study may contribute to shedding new light on the subject. In addition to measuring the effects of remittances on investment and consumption expenditures, we are also attempting to quantify their overall macroeconomic impact on the country's gross domestic product, which, to the best of our knowledge, has not been done for North Macedonia. Thus, we are providing insights that can contribute to future policy advice aimed at maximizing these financial transfers.

**Keywords:** *Migration, Remittances, Economic growth*

**JEL classification:** *F22, F24, F43, E21*

## REFERENCES

- Adams, R. H., Jr. and Cuecuecha, A. (2010), "Remittances, household expenditure and investment in Guatemala", *World Development*, Vol. 38, pp. 1626-1641. <https://doi.org/10.1016/j.worlddev.2010.03.003>
- Arrehag, L., Sjöberg, Ö. and Sjöblom, M. (2005), "Cross-border migration and remittances in a post-communist society: Return flows of money and goods in the Korçë District", Albania, *South-Eastern Europe Journal of Economics*, Vol. 3, pp. 9-40
- Bollard, A., McKenzie, D., Morten, M. and Rapoport, H. (2011), "Remittances and the brain drain revisited: The microdata show that more educated migrants remit more", *World Bank Economic Review*, Vol. 25 No. 1, pp. 132-156. <https://doi.org/10.1093/wber/lhr013>
- Castaldo, A. and Reilly, B. (2007), "Do migrant remittances affect the consumption patterns of Albanian households?", *South-Eastern Europe Journal of Economics*, Vol. 5, pp. 25-44 [http://www.asecu.gr/Seeje/issue 08/castaldo.pdf](http://www.asecu.gr/Seeje/issue%2008/castaldo.pdf)
- Clemens, M. and Ogden, T. (2020), "Migration and household finances", *Development Policy Review*, Vol. 38, pp. 3-27. <https://doi.org/10.1111/dpr.12471>
- Clemens, M. and Tiongson, E. (2017), "Split decisions: Household finance when a policy discontinuity allocates overseas work", *Review of Economics and Statistics*, Vol. 99, pp. 531-543. [https://doi.org/10.1162/REST\\_a\\_00657](https://doi.org/10.1162/REST_a_00657)
- Kikkawa, A., Matsumoto, T. and Otsuka, K. (2019), An Inquiry into the Heterogeneous Outcomes of International Migration: Evidence from Rural Households in Bangladesh, *Journal of Development Studies*, 55 (10): 2106-2128
- Kule, D., Mançellari, A., Papapanagos, H., Qirici, S. and Sanfey, P. (2002), "The causes and consequences of Albanian emigration during transition: Evidence from micro data", *International Migration Review*, Vol. 36 No. 1, pp. 229-239

Yang, D. (2011), “Migrant remittances”, *Journal of Economic Perspectives*, Vol. 25 No. 3, pp. 129-152 <https://doi.org/10.1257/jep.25.3.129>

## **THE IMPACT OF POLITICAL INSTITUTIONS ON ECONOMIC GROWTH IN POST-TRANSITION EUROPE**

**Gunter Merdzan**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
gjunter.merdzan@eccf.ukim.edu.mk*

**Predrag Trpeski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
predrag.trpeski@eccf.ukim.edu.mk*

**Daniela Bojadjieva**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
daniela.mamucevska@eccf.ukim.edu.mk*

### **ABSTRACT**

**Purpose** The significance of political institutions in promoting economic growth is well acknowledged within institutional economics; however, empirical results often display considerable variation, especially in European post-transition economies. As noted by North (1990) and Acemoglu and Robinson (2012), institutions influence incentives, mitigate uncertainty, and dictate the efficiency of resource allocation. Nonetheless, evidence from transition countries indicates that the relationship between political institutions and economic outcomes is contingent upon specific contextual factors.

Earlier studies have yielded mixed results regarding the relationship between economic and political freedoms. Piątek *et al.* (2013) found that economic freedom fosters growth in transition countries, while political freedom appears to have a neutral effect, suggesting that political reforms may lag behind market liberalisation. In contrast, Uzelac *et al.* (2020) and Bayar (2016) demonstrate that democracy, political stability, and the quality of governance, particularly in controlling corruption and upholding the rule of law, positively influence economic performance in Central and Eastern Europe. Alexiou *et al.* (2020) report similar long-term advantages associated with institutional quality, indicating that voice and accountability can enhance growth once the institutional framework becomes stable. Additionally, Tamilina and Tamilina (2014) contend that the growth of post-communist countries relies more on the maturity of political institutions than on their economic counterparts, emphasising the significance of institutional evolution over formal design. Conversely, some scholars, including Hodgson (2006) and Bonnal and Yaya (2015), note that democracy can have neutral or even adverse short-term effects, particularly in fragmented or unstable institutional environments.

In this context, the current study aims to contribute to the ongoing debate by empirically re-evaluating the impact of political institutions, assessed through the Freedom House Index (which measures political rights and civil liberties) (Freedom House, 2024) and the Polity IV Index from the Center for Systemic Peace (which evaluates executive constraints and participation) (Marshall *et al.*, 2019), on economic growth in thirteen European post-transition economies from 1996 to 2019. This paper offers updated, dynamic empirical evidence that

resolves long-standing inconsistencies in the literature on institutions and growth in post-transition Europe. The topic remains highly relevant, as political and institutional reforms continue to influence the region's prospects for convergence, resilience, and long-term development. Building on prior research, this paper utilises a system generalised method of moments (system GMM) to explore the relationship between political institutions and economic growth.

The aim of this study is to assess whether advancements in political institutions, characterised by increased democratic participation, improved accountability, and reinforced constraints on executive power, result in higher GDP per capita growth in post-transition Europe. Consequently, the primary hypothesis examined is: Political institutions that promote greater transparency, participation, and checks on executive authority have a positive, statistically significant effect on economic growth in European post-transition economies.

**Design/methodology/approach** The empirical analysis utilises a dynamic panel-data framework, focusing on thirteen European post-transition economies<sup>1</sup> over the period from 1996 to 2019. To address potential endogeneity, the study employs the system GMM estimator developed by Arellano and Bover (1995) and by Blundell and Bond (1998). This methodology facilitates consistent estimation in dynamic contexts where the lagged dependent variable is included as a regressor and the number of time periods is relatively limited compared to the cross-sectional units.

The dependent variable in this analysis is GDP per capita growth, while the primary explanatory variables are political-institutional indicators. These include the Freedom House Index, which assesses political rights and civil liberties (Freedom House, 2024), and the Polity IV Index from the Centre for Systemic Peace, which evaluates executive constraints, political participation, and regime durability (Marshall *et al.*, 2019). Additionally, control variables include the logarithm of GDP per capita in purchasing power parity (PPP), gross fixed capital formation (GFCF) as a measure of physical capital, human capital and population growth.

Before conducting the main regression analysis, the study examines partial correlations between GDP per capita growth and the two institutional indicators, while controlling for all other growth determinants. This approach aims to isolate the relationship between political institutions and economic performance, independent of structural and macroeconomic influences. The validity of the instruments and the absence of serial correlation are assessed using standard diagnostic tests.

**Findings** When accounting for the influences of various macroeconomic variables, including initial income, gross fixed capital formation, human capital, and population growth, the partial correlations between GDP per capita growth and the two institutional indicators remain positive, though modest in magnitude. This pattern suggests that enhancements in political rights, civil liberties, and executive constraints are consistently associated with higher growth, provided that structural and demographic factors are controlled for.

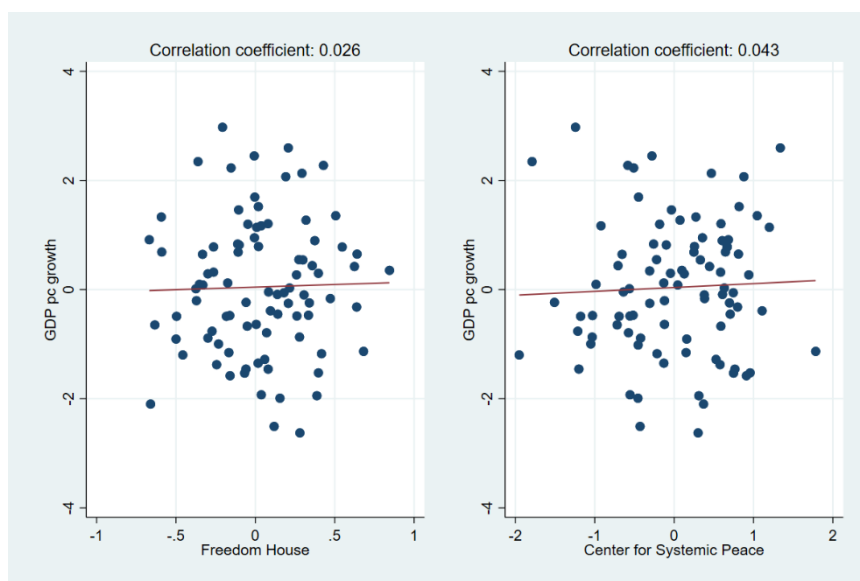
The system GMM estimation reinforces these associations, demonstrating that political institutions have a positive and statistically significant impact on economic growth. The coefficient for the Freedom House Index is 0.421 ( $p < 0.05$ ), while the Polity IV Index from

---

<sup>1</sup> Albania, Croatia, Czechia, Estonia, Hungary, Latvia, Lithuania, North Macedonia, Poland, Romania, Serbia, the Slovak Republic and Slovenia.

the Center for Systemic Peace shows an even stronger and highly significant effect of 0.446 ( $p < 0.01$ ). Diagnostic tests validate the robustness of the model, with results indicating AR(1)  $p = 0.006$ ; AR(2)  $p \approx 0.48$ ; and Sargan  $p \approx 0.45$ . These findings support the validity of the instruments used and confirm the absence of higher-order serial correlation. The lagged GDP per capita term is negative and significant, indicating conditional convergence, whereas human capital demonstrates a substantial positive contribution to growth. Conversely, population growth has a negative effect, and gross fixed capital formation has a negative, significant coefficient, suggesting capital inefficiencies. This aligns with the argument that institutional weaknesses can impede the efficient allocation of investment resources.

Figure 1: Scatter charts of the partial correlation



(Source: Authors' calculations.)

Table 1: Regression results

Variable	Dependent variable: GDP per capita growth	
	Model 1	Model 2
GDP per capita growth (L1)	0.191 (0.119)	0.169 (0.110)
Logarithm of GDP per capita PPP (L1)	-2.468** (0.972)	-2.863*** (0.555)
Gross fixed capital formation (L1)	-0.136*** (0.0355)	-0.122*** (0.0414)
Human capital (L1)	3.029** (1.410)	3.717*** (1.161)
Population growth (L1)	-1.690*** (0.321)	-1.777*** (0.383)
Freedom House (L1)	0.421** (0.182)	
Center for Systemic Peace (L1)		0.446*** (0.143)
Constant	16.44** (6.857)	16.46*** (4.289)
Observations	91	91
Number of groups	13	13

AR(1) p-value	0.00678	0.00619
AR(2) p-value	0.478	0.494
Sargan test p-value	0.500	0.434
Fixed effects in periods	✓	✓

(Source: Authors' calculations)

These findings underscore the significance of political institutions as a crucial factor influencing economic performance in post-transition Europe. They build upon and refine previous empirical results. While Piątek *et al.* (2013) suggested that political freedom was largely neutral for growth when economic freedom was accounted for, our analysis reveals that, through dynamic estimation and appropriate controls, political institutions emerge as a significant catalyst for growth. The pronounced positive impacts of Freedom House and Polity IV align with the conclusions of Uzelac *et al.* (2020) and Bayar (2016), who identified that democracy, political stability, and the control of corruption substantially affect GDP growth in Central and Eastern Europe. Additionally, the long-term positive correlation between institutional quality and growth, as established by Alexiou *et al.* (2020), is reinforced here, particularly with respect to voice, accountability, and the rule of law.

Furthermore, our findings support the theoretical framework proposed by Tamilina and Tamilina (2014), which posits that growth in post-communist economies is more reliant on the maturity and functionality of political institutions than on economic factors. Our evidence reinforces this assertion by showing that, when controlling for endogeneity and persistence, the democratic aspects of institutional quality, such as participation, transparency, and executive accountability, emerge as statistically significant predictors of growth. In contrast to the conclusions drawn by Bonnal and Yaya (2015), who found no meaningful impact of political institutions on global growth rates, our study reveals that, within the specific institutional context of post-transition Europe, these effects are demonstrably positive. Additionally, our results provide a counterargument to Hodgson (2006), who identified a negative relationship between democracy and growth in institutionally fragile settings. The period and countries included in our analysis represent a more consolidated stage of institutional development, which likely alleviated those adverse dynamics.

The combination of positive partial correlations and robust system GMM coefficients provides compelling empirical support for the hypothesis that well-functioning political institutions—marked by democratic accountability, civil liberties, and practical constraints on executive power, facilitate sustainable economic growth in European post-transition economies. This aligns with the theoretical insights of North (1990) and Acemoglu *et al.* (2005), who argue that inclusive political institutions foster credible expectations, enhance governance, and cultivate an environment conducive to long-term development.

These findings carry significant implications for policymakers in post-transition Europe. The demonstrated positive influence of political rights, civil liberties, and executive constraints on economic growth underscores the necessity of continued progress in democratic consolidation and institutional strengthening to sustain development. Reform initiatives should prioritise enhancing judicial independence, ensuring transparent electoral processes, strengthening anti-corruption mechanisms, and promoting broader civic participation. Such measures not only enhance the quality of governance but also cultivate a predictable and trustworthy environment that encourages investment, supports efficient resource allocation, and mitigates the economic costs associated with institutional uncertainty. By reinforcing the maturity and effectiveness of

political institutions, post-transition economies can expedite their convergence with more established democracies and secure more resilient long-term growth trajectories.

While the findings of this study are robust, several limitations should be noted. Due to data constraints, not all post-transition European countries could be included, which may affect the generalizability of the results. Additionally, the analysis is based on two institutional indicators; broader, composite measures might capture additional dimensions of political institutions. Although the empirical strategy employed addresses key endogeneity concerns, alternative methodological approaches could be considered in future studies to further validate and refine the results. Future research could also broaden the range of countries examined and investigate whether different institutional dimensions influence growth through distinct channels.

**Originality/value** This study presents new empirical evidence regarding the relationship between political institutions and economic growth in European post-transition economies, with a specific focus on the quality of democracy, levels of participation, and executive constraints. This approach contrasts with the more commonly analysed factors, such as economic freedom or corruption. The findings indicate that political institutions, as measured by the Freedom House and Polity IV indices, have a positive and significant long-term effect on GDP per capita growth. This underscores the importance of institutional consolidation and democratic accountability as critical drivers of development in the post-transition phase.

The paper contributes originality by reconciling the mixed evidence from previous studies. While Piątek *et al.* (2013) found that political freedom has a neutral impact on growth, and Bonnal and Yaya (2015) indicated a lack of a systematic global relationship, this study demonstrates that, with appropriate dynamic specification, inclusive political institutions actually enhance growth. This supports the assertions of Tamilina and Tamilina (2014) and North (1990) that mature political systems promote credible governance and sustained development. The findings underscore that deepening democracy and reinforcing checks on power are not just political milestones, but essential economic imperatives for achieving long-term growth and convergence in post-transition Europe.

**Keywords:** *Political institutions, Economic growth, European post-transition countries, System GMM*

**JEL classification:** *O11, O43, P20*

## REFERENCES

- Acemoglu, D., Johnson, S. and Robinson, J.A. (2005), "Institutions as the Fundamental Cause of Long-Run Growth", in Aghion, P. and Durlauf, S.N. (Eds), *Handbook of Economic Growth, Vol. 1A*, Elsevier, Amsterdam, pp. 385-472.
- Acemoglu, D. and Robinson, J.A. (2012), *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, Crown Business, New York, NY.
- Alexiou, C., Vogiazas, S. and Solovev, N. (2020), "Economic growth and quality of institutions in 27 postsocialist economies", *Journal of Economic Studies*, Vol. 47 No. 4, pp. 769-787.
- Arellano, M. and Bover, O. (1995), "Another look at the instrumental variable estimation of error-components models", *Journal of Econometrics*, Vol. 68 No. 1, pp. 29-51.
- Bayar, Y. (2016), "Public Governance and Economic Growth in the Transitional Economies of the European Union", *Transylvanian Review of Administrative Sciences*, Vol. 12 No. 48, pp. 5-18.

- Blundell, R. and Bond, S. (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, Vol. 87 No. 1, pp. 115-143.
- Bonnal, M. and Yaya, M.E. (2015), "Political Institutions, Trade Openness, and Economic Growth: New Evidence", *Emerging Markets Finance and Trade*, Vol. 51 No. 6, pp. 1276-1291.
- Freedom House (2024), *Freedom in the World 2024 Data*, Freedom House, Washington, DC.
- Hodgson, G.M. (2006), "Institutions, Recessions and Recovery in the Transitional Economies", *Journal of Economic Issues*, Vol. 40 No. 4, pp. 875-894.
- North, D.C. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge.
- Marshall, M.G., Gurr, T.R. and Jagers, K. (2019), *POLITY IV Project: Political Regime Characteristics and Transitions, 1800-2018. Dataset Users' Manual*, Center for Systemic Peace, available at: <http://www.systemicpeace.org/inscr/p4manualv2018.pdf>
- Piątek, D., Szarzec, K. and Pilc, M. (2013), "Economic freedom, democracy and economic growth: a causal investigation in transition countries", *Post-Communist Economies*, Vol. 25 No. 3, pp. 267-288.
- Tamilina, L. and Tamilina, N. (2014), "Heterogeneity in Institutional Effects on Economic Growth: Theory and Empirical Evidence", *European Journal of Comparative Economics*, Cattaneo University (LIUC), Vol. 11 No. 2, pp. 205-249.
- Uzelac, O., Davidovic, M. and Mijatovic, M.D. (2020), "Legal framework, political environment and economic freedom in Central and Eastern Europe: do they matter for economic growth?", *Post-Communist Economies*, Vol. 32 No. 6, pp. 697-725.

## **THE ROLE OF EMPIRICAL EVIDENCE AND GOVERNMENT PROJECT OFFICE IN THE FORMATION OF THE SLOVENIAN DEVELOPMENT STRATEGY**

**Aleksandar Kešeljević**

*School of Economics and Business,*

*University of Ljubljana, Slovenia*

*saso.keseljevic@ef.uni-lj.si*

### **EXTENDED ABSTRACT**

**Purpose** Financial crisis (2008) and the war in Ukraine (2022-) show that governments have to develop more resilient national development strategies in order to adequately respond to pressing disrupters and unpredictable shifts in the world. The Slovenian government (2014-2018) took Sustainable Development Agenda 2030 (UN) as the baseline platform for its development strategy. The purpose of the article is to show that strong political commitment and a project-based approach (Government Project Office) were the most crucial steps in its implementation, and why certain strategic projects (e.g., green tax reform) were not implemented at the end.

The Slovenian government identified ten strategic projects to be implemented in line with its SDGs strategy. The Government Project Office (GPO) reported at the Government sessions on all projects in progress and drew attention to projects that were lagging behind. One of the key projects was green tax reform (GTR) in order to stimulate structural transformation of the Slovenian economy and its sustainable development.

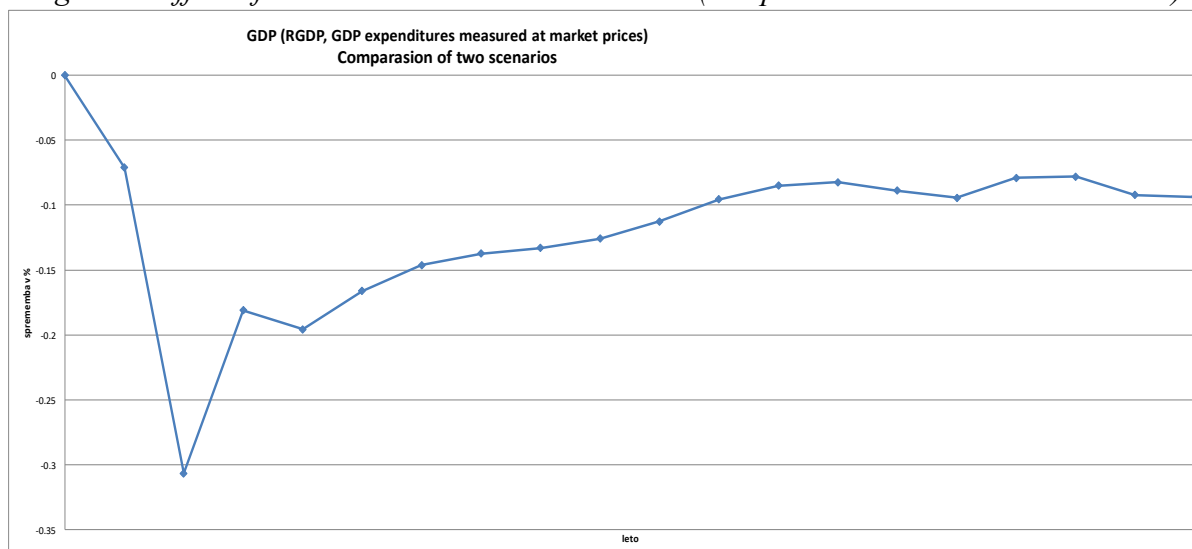
**Design/methodology/approach** GPO was not so much a legal-formal organizational construct but more a way of approaching the management of interdepartmental projects at the government level. The task of the GPO was to monitor the performance of project managers, take care of the project information system, ensure methodologically uniform implementation of projects, draw attention to bottlenecks, and maintain a knowledge base (archive of project documentation, examples of good practices, database of errors). The value added of GPO was its assistance in strategic project promotion, in inter-ministerial coordination and monitoring the implementation process. The office has also trained future project managers and members of their teams by mentoring members and project managers.

The main goal of GTR is to tax bad things (pollution) instead of good ones (income) in order to ensure fiscal neutrality and to improve the situation in an environmental and economic sense (double dividend). We do not go into the topic since conditions for double dividend have been largely discussed elsewhere (Bovenberg and Goulder 1997; Goulder, 1995; Bosquet, 2000; Sokolovska 2020). E3ME model (dynamic simulation econometric model) was used for analyzing the macroeconomic effects of the new green tax. Analysis was based on a comparison of two empirical scenarios, baseline scenario where the green tax was not introduced, and a projection that assumed the introduction of the specific green tax. Comparison between two scenarios was made for some key economic variables (exports, GDP, employment), and in subsequent analyses, where green tax was also recycled in various ways.

**Finding No. 1.** The Effect of the yearly carbon tax in the amount of EUR 15/tCO<sub>2</sub> on GDP is presented in the Figure. Introduction of an extra annual carbon tax would lead to the highest drop in Slovenia's GDP in the second year relative to the baseline scenario with no introduction

of a carbon tax. After the second year, the difference between the two scenarios would gradually decrease. This confirmed our expectations and theoretical findings in the literature. Companies need time to implement new technologies, and consumers need time to adjust their consumption patterns. Thus, policy makers should introduce green tax gradually, transparently, and predictably.

Figure 1: Effect of carbon tax introduction on GDP (comparison to the baseline scenario)



(Source: Kešeljević and Koman, 2014)

**Finding No. 2.** Budget should reflect strategic priorities. The implementation of the Strategy with supporting strategic plan, strategic priorities and projects should be based on medium-term planning and program budget. Experience from Slovenia and some other countries (Finland) show that developing a strong medium-term dimension in budgeting process beyond the traditional annual cycle is a real challenge due to political constraints and short-sighted of many politicians, especially in a coalition government.

**Finding No. 3.** In spite of positive contributions of the GPO and implementation of the ministerial project offices, the GPO was completely disintegrated and dismantled in the next government mandate.

**Originality/Value.** The comparison between different scenarios was carried out in order to identify the optimal fiscal instrument for Slovenia to achieve the double dividend. Analysis showed a relatively small impact of the new green tax on GDP. That was the main reason why GTR was not implemented as one of the key strategic government projects. The major challenge in the creation of the Slovenian development strategy was how to integrate sustainability goals and strategic orientation within the budget (program budget) and how to preserve some good practices as GPO.

**Keywords:** Sustainable development, National development strategy, Green tax reform

**JEL classification:** H23, Q44, Q50

## REFERENCES

- Bosquet, B. (2000), "Environmental tax reform: does it work? A survey of the empirical evidence", *Ecological Economics*, Vol. 34, 19-32.
- Bovenberg, L. & Goulder, L.H. (1997), "Cost of environmentally motivated taxes in the presence of other taxes: general equilibrium analyses", *National Tax Journal*, Vol. 50 No. 1, pp. 59-87.
- Goulder, L.H. (1995), "Environmental taxation and the double dividend: A reader's guide", *International Tax and Public Finance*, Vol. 2 No. 2, pp. 155-182.
- Kešeljević, A. and Koman, M. (2014), "Analysis of the effects of introduction of an additional carbon tax on the Slovenian economy considering different forms of recycling", *Economic and Business Review*, Vol. 16 No. 3, pp. 247-277.
- Sokolovska, I. (2020), *The analysis of the double dividend from the perspective of public and private finance*, Doctoral dissertation, Ljubljana.

## DIGITAL INFRASTRUCTURES AS THE NEW PIPES OF GLOBAL CAPITAL

**Jasna Tonovska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
jasna.tonovska@eccf.ukim.edu.mk*

**Predrag Trpeski**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
predrag.trpeski@eccf.ukim.edu.mk*

### EXTENDED ABSTRACT

**Purpose** Global capital increasingly flows not only through traditional financial markets and institutions, but through digital infrastructures—payment platforms, mobile banking systems, and online transaction networks that have become the new pipes of the global economy. These infrastructures shape how quickly capital enters, how securely it is intermediated, and how abruptly it can reverse. In this sense, digitalization is now part of the architecture of the global economy, conditioning the volatility that defines openness. Traditional analyses of the drivers of capital flows and their volatility, built around push factors (global shocks) and pull factors (domestic fundamentals), cannot fully explain why countries with similar fundamentals often face very different volatility profiles (Koepke, 2015; Cerutti *et al.*, 2015; Fratzscher, 2011). As Carney (2019) noted, the pipes through which capital flows are transmitted matter as much as the drivers themselves. This study builds directly on that insight and provides a systematic empirical evidence that digital infrastructures act as pipes that shape the stability of global capital.

**Design/methodology/approach** We extend the classical push–pull framework into a push–pull–pipes model (Pagliari *et al.*, 2017; Wang, 2019). Push factors include a Global Financial Conditions Index (data sourced from Capital Economics), U.S. real GDP growth volatility (estimated volatility based on data sourced from FRED), and Global Economic Conditions volatility (estimated volatility based on data sourced from KOF Swiss Economic Institute). Pull factors comprise a Domestic Financial Conditions Index (data sourced from Capital Economics, Boraccia *et al.*, 2023), a Financial Development Index (data sourced from the IMF), Capital Account Openness (Chinn–Ito index; Chinn *et al.*, 2008), and Real Effective Exchange Rate volatility (estimated volatility based on data sourced from Bruegel). Pipes are measured by Relative Sovereign Credit Ratings (score versus the global median; De *et al.*, 2020) and a Relative Digital Infrastructure Index (mobile and internet banking penetration per 1,000 adults, benchmarked globally).

$$\sigma_{kt}^{ij} = \alpha_k^{ij} + \beta^{ij} PUSH_t + \gamma^{ij} PULL_{k,t-1} + \delta^{ij} PIPES_{k,t-1} + \varepsilon_{kt}^{ij}$$

In this study, we explicitly embed digital infrastructures within the pipes dimension, thereby expanding the volatility debate beyond incentives to the channels through which capital is transmitted. Regarding the methodology, the analysis uses an unbalanced quarterly panel of advanced and emerging, and developing economies over 2000Q1–2023Q1. The dependent variable  $\sigma_{kt}^{ij}$  is the volatility of gross capital inflows, which is estimated as follows. For each economy and instrument (direct, portfolio, and other investment), we first estimate residuals from an ARIMA (1,1,0) model on the quarterly series; second, the standard deviation of these

residuals is calculated as our volatility measure. The econometric specification employs panel fixed effects to control for unobserved country heterogeneity, with Driscoll–Kraay standard errors to address heteroskedasticity and autocorrelation. We estimate separate regressions for the drivers of volatility in FDI, portfolio, and other-investment inflows. We also run the specifications on the EMDE subsample to examine this group more closely.

**Findings** Across the full sample, push factors dominate: global financial conditions and U.S. cycle volatility account for most of the variation in capital-flow volatility. Pull factors are uneven: stronger financial development and greater openness tend to attract flows but often raise volatility, while REER volatility remains a persistent destabilizer. Within this structure, pipe condition outcomes. Relative sovereign ratings are procyclical—upgrades attract inflows that are often more volatile—and digital infrastructures display a negative and statistically significant association with the volatility of total, portfolio, and especially banking-related inflows, suggesting that countries with deeper digital systems experience smoother financial intermediation and lower exposure to sudden stops.

In EMDEs, however, the estimated coefficients on the Digital Infrastructure Index remain negative but largely insignificant, indicating that the stabilizing potential of digital pipes does not systematically materialize in weaker institutional environments.

This pattern implies that while digitalization may facilitate access and speed of intermediation, it cannot by itself anchor stability without complementary institutional quality and regulatory depth. In such contexts, digital infrastructures operate more as accelerators of integration than as absorbers of volatility—enhancing financial connectivity but offering limited insulation from global shocks.

**Originality/value** For EMDEs, investing in digital infrastructures should be viewed both as a development priority and as a conditional macro-financial resilience strategy. While digital systems can expand access and efficiency, their stabilizing potential remains largely unrealized in the absence of strong institutional and regulatory frameworks. Without these complements, digital infrastructures enhance financial connectivity but offer limited insulation from external shocks, operating more as channels of speed than of stability. Strengthening institutional quality and financial supervision is therefore essential for digital pipes to function as genuine macro-stabilizers.

At the global level, integrating the “pipes” dimension—particularly digital infrastructures—into the IMF’s Integrated Policy Framework would acknowledge their role as structural complements to monetary, fiscal, and macroprudential instruments (Basu et al., 2023; IMF, 2023; IMF, 2022). Doing so would allow global surveillance and policy design to better capture how technological architectures condition the transmission and volatility of capital in an increasingly digitalized financial system (IMF, 2023).

This study reconceptualizes the analysis of capital flow volatility for the digital era by extending the classical push–pull framework into a structural push–pull–pipes perspective. The evidence confirms that while global and domestic macro-financial conditions remain the primary determinants of volatility, the channels through which capital circulates increasingly matter (Reuter *et al.*, 2025). Among these, digital infrastructures emerge as a new driver of capital flow volatility. Across all economies, deeper digital infrastructures are associated with lower volatility in total, portfolio, and banking inflows, reflecting smoother intermediation and improved informational efficiency. In EMDEs, however, the effect remains statistically weak,

implying that digital systems alone cannot substitute for institutional strength. Their stabilizing potential is contingent on regulatory quality, supervisory depth, and the credibility of financial governance. Digital pipes are reshaping the speed, reach, and resilience of capital mobility. Recognizing their dual role—facilitating integration while conditioning volatility—should become central to future macro-financial frameworks.

**Keywords:** *Capital flows volatility, Drivers, Push, Pull, Pipes, Digital infrastructure*

**JEL classification:** *F32, G15, O33*

Table 1: Baseline Regression

	All countries					Emerging and Developing countries						
	Total Inflows	Gross Inflows	Direct Inv Inflows	Portfolio Inv Inflows	Other Inv Inflows	Other Inflows Banks	Total Inflows	Gross Inflows	Direct Inv Inflows	Portfolio Inv Inflows	Other Inv Inflows	Other Inflows Banks
Country Fin. Conditions Index (t-1)	0.167 (0.413)	-0.035 (0.085)	0.052 (0.228)	-0.166 (0.201)	0.030 (0.169)	0.408 (0.355)	0.136 (0.092)	-0.124 (0.194)	-0.297 (0.302)	0.220 (0.154)		
Fin. Development Index (t-1)	4.584*** (0.939)	-0.238 (0.226)	1.275** (0.539)	2.706*** (0.562)	3.237*** (0.500)	3.387** (1.472)	0.312 (0.576)	-0.233 (0.637)	2.806* (1.560)	2.927*** (1.073)		
REER Volatility (t-1)	0.163* (0.097)	-0.049** (0.020)	0.081 (0.051)	0.178** (0.077)	0.036 (0.049)	0.024 (0.089)	-0.108*** (0.029)	0.027 (0.063)	0.193* (0.112)	0.019 (0.044)		
US rGDP Growth (t-1)	0.097* (0.053)	0.023 (0.016)	-0.011 (0.062)	0.122 (0.075)	0.069 (0.062)	0.053 (0.073)	-0.020 (0.025)	-0.028 (0.097)	0.111 (0.093)	0.077 (0.076)		
Chinn - Ito (t-1)	-1.201 (2.006)	0.905 (0.582)	-2.436* (1.444)	-2.147 (1.656)	-1.650* (0.969)	0.544 (2.216)	2.018* (1.069)	0.073 (1.441)	-0.160 (2.077)	1.153 (1.350)		
World Economic Conditions Volatility (t-1)	0.156*** (0.027)	0.017 (0.011)	0.124*** (0.026)	-0.046 (0.033)	-0.035* (0.019)	0.141*** (0.039)	0.011 (0.010)	0.139*** (0.040)	-0.060 (0.040)	-0.039 (0.025)		
World Fin. Conditions Index (t-1)	-0.053 (0.325)	-0.151*** (0.056)	-0.774*** (0.243)	0.483*** (0.163)	0.582** (0.230)	-0.111 (0.395)	-0.272*** (0.093)	-0.417** (0.203)	0.668** (0.282)	0.648*** (0.135)		
Country Relative Rating (t-1)	0.063* (0.037)	-0.015 (0.018)	0.001 (0.034)	0.014* (0.022)	-0.057*** (0.017)	0.015* (0.036)	-0.021*** (0.007)	0.032** (0.020)	0.015 (0.016)	0.011 (0.012)		
Relative Digital Infrastructure Index (t-1)	-0.056** (0.541)	-0.131 (0.248)	-0.563* (0.326)	-0.059* (0.373)	-1.401*** (0.286)	-0.675* (0.535)	-0.260 (0.159)	0.217 (0.286)	0.162 (0.384)	0.092 (0.430)		
Constant	-10.919*** (3.798)	3.270*** (0.947)	0.764 (2.449)	-3.582 (2.529)	-8.343*** (2.130)	-5.921 (5.930)	0.525 (2.322)	4.099 (3.017)	-5.719 (6.249)	-8.931* (5.021)		
Observations	991	991	990	990	990	675	675	674	674	674		
Groups	22	22	22	22	22	15	15	15	15	15		
Country fixed effects	yes	yes	yes	yes	yes	Yes	yes	yes	yes	yes		

Source: Authors' calculations.

## REFERENCES

- Basu, S., Boz, E., Gopinath, G., Roch, F. and Unsal, F. (2023), "Integrated Monetary and Financial Policies for Small Open Economies", working paper 23/161, *International Monetary Fund*, Washington, DC, June.
- Boraccia, G., Espinoza, R.A., Guzzo, V., Lafarguette, R., Jiang, F., Nguyen, V., Segoviano, M.A. and Wingender, F. (2023), "Financial Conditions in Europe: Dynamics, Drivers and Macroeconomic Implications", working paper, *International Monetary Fund*, Washington, DC, 29 September.
- Carney, M. (2019), "Pull, Push, Pipes: Sustainable Capital Flows for a New World Order", available at: <https://www.bankofengland.co.uk>.
- Cerutti, E., Claessens, S. and Puy, D. (2015), "Push factors and capital flows to emerging markets: Why knowing your lender matters more than fundamentals", Working Paper 15/127, *International Monetary Fund*, Washington, DC, June.
- Chinn, M.D. and Ito, H. (2008), "A new measure of financial openness", *Journal of Comparative Policy Analysis*, Vol. 10 No. 3, pp. 309-322.
- De, S., Mohapatra, S. and Ratha, D. (2020), "Sovereign credit ratings, relative risk ratings, and private capital flows", Working Paper 9401, *World Bank Policy Research Working Paper Series*, Washington, DC, September.
- Fratzscher, M. (2011), "Capital flows, push versus pull factors and the global financial crisis", Working Paper No. 1364, *European Central Bank*, Frankfurt, June.
- IMF (2022), "Review of the institutional view on the liberalization and management of capital flows", Policy Paper No. 22/97, *International Monetary Fund*, Washington, DC, March.
- IMF (2023), "Integrated policy framework – Principles for the use of foreign exchange intervention", Policy Paper, *International Monetary Fund*, Washington, DC, 28 November.
- IMF (2023), "Shocks and capital flows: Policy responses in a volatile world", Policy Paper, *International Monetary Fund*, Washington, DC, 27 October.
- Koepke, R. (2015), "What drives capital flows to emerging markets? A survey of the empirical literature", Working Paper No. 62770, *MPRA (Munich Personal RePEc Archive)*, Munich, March.
- Pagliari, M.S. and Hannan, S.A. (2017), "The volatility of capital flows in emerging markets: Measures and determinants", Working Paper No. 17/41, *International Monetary Fund*, Washington, DC, February.
- Reuter, M., Agur, I., Copestake, A., Martínez Pería, M.S. and Teoh, K. (2025), "Payment frictions, capital flows, and exchange rates", Working Paper No. 25/171, *International Monetary Fund*, Washington, DC, January.
- Wang, S. (2019), "Capital flow volatility: The effects of financial development and global financial conditions", Papers pwa945, *Job Market Papers*.

## FROM HUMAN DEVELOPMENT TO INNOVATION OUTCOMES: A CROSS-COUNTRY AND INCOME GROUP ANALYSIS

**Andrijana Bojadjevska Danevska**

*Faculty of Economics and Administrative Sciences  
International Balkan University, Skopje, North Macedonia  
Andrijanab.danevska@ibu.edu.mk*

### EXTENDED ABSTRACT

**Purpose** Innovation and human capital are considered the two main pillars of contemporary economic growth (Aiting *et al.*, 2022). Additionally, certain scholars consider that human capital is an innovation engine (Acemoglu, 1996; Aghion and Howitt, 1998). However, the evidence from developing countries shows that the increased investment in human capital has not brought the corresponding innovation growth (Li and Nan, 2019). In this context, this paper investigates the path from human capital, through innovation inputs and processes to innovation outcomes, by utilizing data from the Digital Evolution Index (2025) for 125 countries over the timespan 2008-2023. More specifically, this research examines whether human development presented as the “state of human condition” and measured within the Digital Evolution Index, translates into higher levels of innovation inputs, processes, and finally innovation outcomes, i.e., technological and innovation outputs that reflect society’s digital progress. In this framework, human development is conceptualized as a set of socio-economic conditions that enable individuals and communities to adopt, adapt, and benefit from digital technologies.

**Design/methodology/approach** Panel regressions, i.e., Fixed Effects and Random Effects models, were applied to control for unobserved country heterogeneity. The Hausman Test statistics with 1.87 and a p-value of 0.76 showed that Random Effects estimators are more consistent. Additionally, to complement the causal inference, a Random Forest Regressor was used in order to capture non-linear patterns and provide feature importance analysis. Also, to account for structural differences across development levels, income group classifications by the World Bank (low, lower-middle, upper-middle, high) were included as additional predictors in the Random Forest Model. In simple terms, to test whether the development stage modifies the path. This study uses a combination of econometrics with machine learning models in order to improve predictive accuracy, provide more accurate and robust predictions (Khan and Wyrwa, 2025).

**Findings** Panel regression confirms that the improvements in human development (measured and presented as human condition) are significantly associated with greater innovation inputs, which in turn foster processes and eventually outcomes. In the preferred Random Effects model, human development, inputs, and processes are all positive and significant predictors of outcomes.

The Random Forest model has a strong predictive performance ( $R^2 > 0.70$ , MAE low relative to outcome scale). It was trained on 80% of the sample and tested on the remaining 20%. The feature importance analysis provides insights into the relative predictive power of each

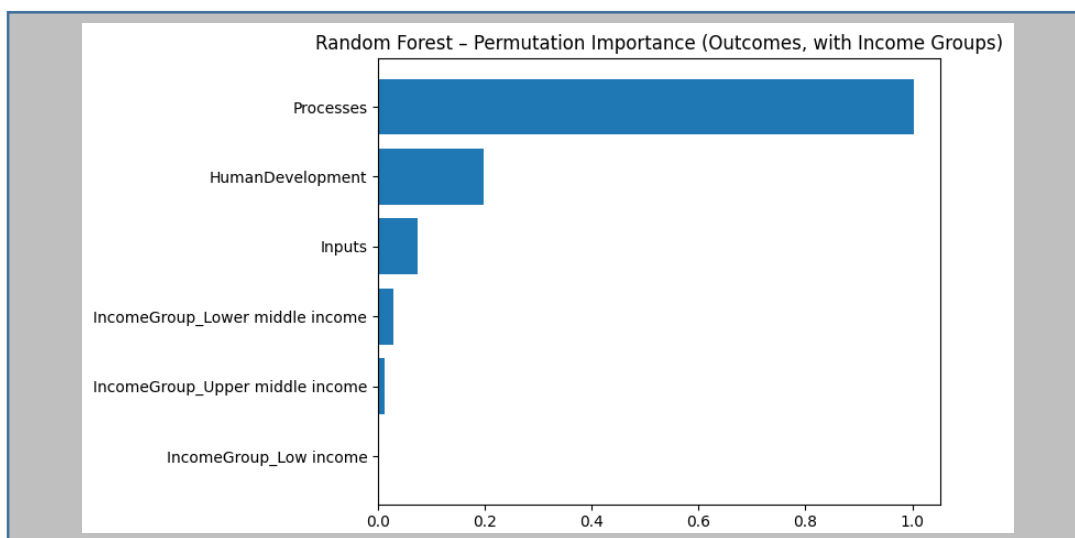
analyzed determinant, where both impurity-based and permutation importance (see Figure 1) confirm that:

- Innovation processes emerge as by far the most influential predictor of innovation outcomes (accounting for the vast majority of predictive power, i.e., close to 1.0 in permutation importance)
- State of human development contributes modestly (~0.20)
- Innovation inputs are weaker (~0.07)
- Income groups contribute modestly, with Lower-middle (~0.03) and Upper-middle (~0.01) income classifications adding small predictive value, while Low-income countries show negligible importance.

In summary, these findings show that while structural development levels matter, actually the strength of processes rather than income group status per se that primarily explains the cross-country differences in innovation outcomes.

**Originality/value** This research combines panel regression and machine learning to empirically trace the path from human development, through innovation inputs and processes to outcomes. Findings suggest that investments in human development and innovation inputs are necessary but insufficient unless supported by robust processes, i.e. “systems in place which can facilitate the development of innovative ideas and practices” (Chakravorti *et al.*, 2025), that translate inputs into tangible outputs, measured through ICT Service Exports (%), ICT Goods Exports (% Total Goods Exports), Apps developed per person, Scientific and Technical Journal Articles etc.

Figure 1: Permutation importance of Predictors for Innovation Outcomes (Random Forest)



(Source: Authors' calculation based on DEI 2025)

Table 1: Panel Regression Results (Dep. variable: Innovation outcomes)

Variable	FE coef. (t)	RE coef. (t)
Human Development	0.311 (3.21)***	0.320 (15.60)***
Inputs	0.101 (2.11)**	0.090 (4.33)***
Processes	0.147 (2.10)**	0.198 (7.69)***
Constant	-0.705 (-0.21)	-2.476 (-2.13)**
R <sup>2</sup> (overall)	0.602	0.630
No.obs	2000	2000

(Source: Authors' calculation based on DEI 2025)

As shown in Table 1, both models fit the data well, indicating that around 60% of the variation in innovation outcomes is explained by the included predictors. Human development is a significant predictor in both FE and RE models, which means that improvements in social readiness, i.e., education, health, and income levels, etc., are strongly associated with higher innovation outcomes. Inputs and processes also show positive and statistically significant effects, highlighting that R&D investments, financial resources, effective institutional and organizational mechanisms, such as digital adoption, collaborative networks, and policy support, are critical in translating human development into tangible innovation results.

**Keywords:** Panel regression, Random forest, Digital evolution, Innovation outcomes, Human development

**JEL classification:** O15, O31, O33

## REFERENCES

- Acemoglu, D. (1996), "A microfoundation for social increasing returns in human capital accumulation", *Quarterly Journal of Economics*, Vol. 111 No. 3, pp. 779-804.
- Aghion, P. and Howitt, P. (1998), "Capital accumulation and innovation as complementary factors in long-run growth", *Journal of Economic Growth*, Vol. 3 No. 2, pp. 111-130.
- Aiting, X., Qiu, K. and Jin, C. (2022), "Regional innovation ability and its inequality: Measurements and dynamic decomposition", *Technological Forecasting and Social Change*, Vol. 180, pp. 121-143.
- Chakravorti, B., Chaturvedi, R.S., Filipovic, C. and Niu, X. (2025), *Digital Evolution Index 2025: From the COVID shock to the AI surge—How 125 digital economies around the world are evolving and changing*. The Fletcher School, Tufts University. Available at: <https://digitalplanet.tufts.edu/wp-content/uploads/2025/dei/Digital-Evolution-Index-2025.pdf> (Accessed: 20 August 2025).
- Khan, A.M. and Wyrwa, A. (2025), "Integrating machine learning and econometric models to uncover macroeconomic determinants of renewable energy production in the selected European countries", *Energy*, Vol. 290, 137266.
- Li, J. and Nan, Y. (2019), "Decision making under human capital mismatch: Prioritizing innovation-driven or industrial upgrading?", *Economic Research – Ekonomska Istraživanja*, Vol. 32 No. 1, pp. 54-70.

## NEUROSCIENCE TECHNIQUES IN HUMAN RESOURCE MANAGEMENT - AN OVERVIEW

**Bojan Kitanovikj**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
bojan.kitanovikj@eccf.ukim.edu.mk*

**Ljupcho Eftimov**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
ljupco.eftimov@eccf.ukim.edu.mk*

### EXTENDED ABSTRACT

**Purpose** Neuroscience is increasingly being applied to management and organizations, giving rise to neuromanagement or organizational neuroscience. Scholars have called for integrating human resource management (HRM) with neuroscience methods to gain a deeper understanding of employee behavior and organizational phenomena (Becker and Menges, 2013; Waldman *et al.*, 2019). This extended abstract summarizes techniques that are applied in the neuro-HRM context. The purpose is to identify how these tools have been used to enhance understanding of HRM topics, more precisely recruitment, training, employee well-being, leadership, and teamwork, and to synthesize key findings. By addressing the question “what’s going on in a person’s head that makes them think or act the way they do?” in the workplace, this review sheds light on the potential of neuroscientific techniques to inform HRM practices. It aims to clarify the state of the art in neuro-HRM and demonstrate how measuring brain activity and other physiological responses can complement traditional HRM methods. Ultimately, the purpose is to highlight the current contributions and challenges of applying neuroscience in HRM and to identify opportunities for future research in this emerging interdisciplinary field.

**Design/methodology/approach** A systematic literature review was conducted to gather peer-reviewed studies at the intersection of neuroscience and HRM. We searched the Scopus database with no date restrictions for articles related to neuromanagement, focusing only on scientific, peer-reviewed articles. The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for identification, screening, and inclusion of studies (Moher *et al.*, 2016). After removing duplicates and unrelated papers, we obtained a final sample of 71 publications. Since this extended abstract refers to research in progress, in the next phases, we will expand with additional keywords in the search query related to neuroscience, organizations, management, and similar. Each article was analyzed to extract its research design, the neuroscience techniques used, the HRM domain or issue addressed, and key findings. We categorized the studies according to the major HRM functional areas of the employee life cycle. This methodological approach ensured that preliminary results are drawn from a broad evidence base of peer-reviewed sources, and the analysis distills common themes and outcomes across studies.

**Findings** Neuroscience techniques adopted in HRM research are diverse, ranging from brain-imaging methods to physiological and behavioral measurements. These tools have found a growing application as they enable capturing the employees’ subconscious or involuntary responses, complementing traditional self-report surveys by providing physiological and neural

data in real time (Waldman *et al.*, 2019). To begin with, electroencephalography (EEG) and related neurophysiological measures are the most frequently used techniques in organizational neuroscience studies, considering their appeal lies in their millisecond-level temporal resolution for tracking brain activity states like engagement or stress in real time (Zazon *et al.*, 2023). Functional magnetic resonance imaging (fMRI), with high spatial resolution, has also been applied, though less often, to map which brain regions are activated by managerial stimuli, mostly in the context of ethical dilemmas or leadership scenarios (Choi *et al.*, 2022). A newer portable imaging method, functional near-infrared spectroscopy (fNIRS), appears in a few studies to monitor cortical activity during tasks in more natural settings (Zhang *et al.*, 2021). In addition, several HRM studies incorporate eye-tracking (ET) technology, which records where and how long individuals direct their gaze. Eye-tracking is used to examine visual attention patterns, for instance, in how job candidates or employees scan information on career websites or during training materials (Mönke *et al.*, 2025). Similarly, facial expression analysis is employed to capture emotional responses either through software or sensors that detect subtle muscle movements (Su *et al.*, 2021). These reflexive measures of gaze, pupil dilation, and facial cues provide a more objective direction in terms of individuals’ reactions and affective states when exposed to HR-related stimuli. Furthermore, studies have used physiological sensors such as electrodermal activity (EDA) or galvanic skin response to gauge arousal and stress levels, heart rate or ECG for workload and stress, and even hormone levels in a few cases, although such measures are less common in HRM literature (Zito *et al.*, 2021). Some research incorporates behavioral response metrics like reaction time tests to infer cognitive processes and implicit biases relevant to HR decisions (Waldman *et al.*, 2019). Table 1 provides an overview of these neuroscientific techniques and their application in HRM literature.

Table 1: Neuroscientific techniques in human resource management, applications and key findings

Technique	Typical measures	HRM application areas	Illustrative findings
EEG	Brainwave activity (alpha, beta, theta) indicating attention, engagement, stress, mental workload	<ul style="list-style-type: none"> <li>Recruitment and selection</li> <li>Training and learning</li> <li>Leadership</li> <li>Fatigue monitoring</li> </ul>	<ul style="list-style-type: none"> <li>Detecting candidate stress peaks during interviews and identifies engaging discussion topics (Zito <i>et al.</i>, 2021),</li> <li>Tracking real-time attention and cognitive load during training sessions to optimize content delivery (Wang <i>et al.</i>, 2021),</li> <li>Revealing neural signatures of inspirational leadership (Waldman <i>et al.</i>, 2019).</li> </ul>
fMRI	Blood-oxygenation (BOLD) signals identifying activated brain regions	<ul style="list-style-type: none"> <li>Leadership</li> <li>Decision-making</li> <li>Employee ethics</li> </ul>	<ul style="list-style-type: none"> <li>Mapping neural correlates of moral reasoning and vision communication in leaders, highlighting activity in prefrontal and limbic regions (Balthazard <i>et al.</i>, 2012).</li> </ul>
fNIRS	Cortical oxygenation levels during naturalistic tasks	<ul style="list-style-type: none"> <li>Decision-making,</li> <li>Workload monitoring</li> </ul>	<ul style="list-style-type: none"> <li>Showing frontal cortex synchrony can predict team coordination quality under pressure (Zhang <i>et al.</i>, 2021).</li> </ul>
ET	Fixation duration,	<ul style="list-style-type: none"> <li>Recruitment (resume</li> </ul>	<ul style="list-style-type: none"> <li>Revealing recruiter gaze patterns on resumes and candidate profiles,</li> </ul>

	saccades, pupil dilation	screening and website design) <ul style="list-style-type: none"> <li>• Training material evaluation</li> </ul>	guiding layout optimization (Pina et al., 2023), <ul style="list-style-type: none"> <li>• Detecting attention hotspots in e-learning interfaces (Mönke et al., 2025).</li> </ul>
Facial expression analysis / facial EMG	Micro-expressions, muscle activation linked to discrete emotions	<ul style="list-style-type: none"> <li>• Recruitment interviews</li> <li>• Employee engagement</li> <li>• Customer-service training</li> </ul>	<ul style="list-style-type: none"> <li>• Identifying subtle affective responses to interview questions or training stimuli, complementing self-report measures (Su et al., 2021).</li> </ul>
EDA	Skin conductance reflecting arousal and stress	<ul style="list-style-type: none"> <li>• Recruitment</li> <li>• Workload monitoring</li> <li>• Stress monitoring</li> </ul>	<ul style="list-style-type: none"> <li>• Capturing stress peaks during critical job-interview phases (Zito et al., 2021),</li> <li>• Signaling rising stress during back-to-back virtual meetings, supporting break scheduling (Becker &amp; Menges, 2013).</li> </ul>
Heart-rate/ECG and other peripheral sensors	Heart-rate variability, respiration	<ul style="list-style-type: none"> <li>• Well-being</li> <li>• Fatigue detection</li> </ul>	<ul style="list-style-type: none"> <li>• Providing early warning of cognitive overload in safety-critical occupations (Caldwell et al., 2019).</li> </ul>

(Source: Authors' work)

Still, across methods, threats to validity, generalizability, and laboratory bias persist. Some of these challenges stem from small samples, task artificiality, analytic flexibility, and over-interpretation of proxies, such as EEG indices of engagement, and EDA peaks as stress without behavioral triangulation (Massaro and Baljević, 2022). These concerns are widely noted in the literature and call for hypotheses' pre-registration, power and sensor justifications, and field validation beyond laboratory contexts (Radtke Caneppele *et al.*, 2022). Thus, recent works echo the lab-to-field gap, and portability promises for EEG/fNIRS and wearables, urging multi-site replications and explicit reporting of sensor quality and preprocessing choices (Radtke Caneppele *et al.*, 2022; Waldman *et al.*, 2019).

**Originality/value** This review is among the first to systematically map the neuroscience techniques applied in HRM and to integrate findings across recruitment, training, leadership, and well-being research. Its value lies in demonstrating how methods such as EEG, fMRI, fNIRS, eye-tracking, and physiological sensors can pinpoint subconscious drivers of employee behavior and provide evidence-based guidance for HR professionals and managers in general. Hence, the implications for managers and HR professionals point out that bio-signals can be applied as process-improvement decision support rather than individual gatekeeping. Illustratively, pilot eye-tracking and EDA can be used to redesign hiring UX and interview pacing, workload, and attention indicators can be utilized to pace training and coach teams, and employees' well-being can be monitored at the aggregate level to guide task rotation and recovery policies, coupled with performance and participant feedback (Awumey *et al.*, 2024; Mönke *et al.*, 2025; Wang *et al.*, 2021). To balance benefits with risks, strict ethical guardrails should be imposed by management to purpose limitation, obtain informed, revocable consent, vouch for local processing and short retention of data, implement role-based access and third-party audits, as well as ban covert use or disciplinary decisions based on physiology, to name a few (Hurley *et al.*, 2024). Considering this, the study extends HRM theory to the biological

level of analysis while offering practitioners novel tools for improving employee experience and organizational effectiveness (Waldman *et al.*, 2019).

By consolidating scattered studies, the review identifies that some areas, like diversity management and employee satisfaction, remain underexplored, joining the call for increased multidisciplinary collaboration. Namely, we propose some future research trends: 1) field-based trials with mobile EEG or fNIRS and wearables inside authentic HR workflows like assessment centers, onboarding and recruitment stages, while evaluating downstream outcomes; 2) benchmark models across genders, age groups, cultures, and integrate subgroup performance to mitigate bias in physiological or affective inference; 3) longitudinal facial-EMG markers can be connected with job satisfaction and engagement trajectories for employees instead of just momentary states; 4) hyperscanning of EEG or fNIRS can be scaled in hybrid teams to test how inter-brain synchrony relates to psychological safety, adaptation, and performance under pressure (Wang *et al.*, 2021; Réveillé *et al.*, 2025; Massaro *et al.*, 2025).

**Keywords:** *Neuromanagement, Human resource management, Neuro-human resource management, Neuroscience techniques*

**JEL classification:** *M12, M54, O15*

## REFERENCES

- Awumey, E., Das, S. and Forlizzi, J. (2024), “A systematic review of biometric monitoring in the workplace: analyzing socio-technical harms in development, deployment and use”, In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 920-932). <https://doi.org/10.1145/3630106.3658945>
- Balthazard, P. A., Waldman, D. A., Thatcher, R. W. and Hannah, S. T. (2012), “Differentiating transformational and non-transformational leaders on the basis of neurological imaging”, *The Leadership Quarterly*, Vol. 23 No. 2, pp. 244-258. <https://doi.org/10.1016/j.leaqua.2011.08.002>
- Becker, W. J. and Menges, J. I. (2013), “Biological implicit measures in HRM and OB: A question of how not if”, *Human Resource Management Review*, Vol. 23 No. 3, pp. 219-228. <https://doi.org/10.1016/j.hrmr.2012.12.003>
- Caldwell, J. A., Caldwell, J. L., Thompson, L. A. and Lieberman, H. R. (2019), “Fatigue and its management in the workplace”, *Neuroscience & Biobehavioral Reviews*, Vol. 96, pp. 272-289. <https://doi.org/10.1016/j.neubiorev.2018.10.024>
- Choi, W. D., Kelley, T., Walden, E., McCool, B. and O’Boyle, M. (2022), “Emotional dissonance among frontline hospitality employees: An exploratory study examining habituation theory using fMR”, *Journal of Hospitality & Tourism Research*, Vol. 46 No. 1, pp. 174-188. <https://doi.org/10.1177/10963480211011633>
- Hurley, M. E., Sonig, A., Herrington, J., Storch, E. A., Lázaro-Muñoz, G., Blumenthal-Barby, J. and Kostick-Quenet, K. (2024), “Ethical considerations for integrating multimodal computer perception and neurotechnology”, *Frontiers in Human Neuroscience*, Vol. 18, 1332451. <https://doi.org/10.3389/fnhum.2024.1332451>
- Massaro, S. and Baljević, D. (2022), “Organizational Neuroscience”, *Oxford Research Encyclopedia of Business and Management*. Retrieved 21 Oct. 2025, from <https://oxfordre.com/business/view/10.1093/acrefore/9780190224851.001.0001/acrefore-9780190224851-e-324>

- Massaro, S., Kang Yang, T. Y. and Christopoulos, G. (2025), “Wearable devices methodology: Opportunities and challenges in human resource management”, *Human Resource Management Journal*. <https://doi.org/10.1111/1748-8583.70014>.
- Moher, D., Stewart, L. and Shekelle, P. (2016), “Implementing PRISMA-P: recommendations for prospective authors”, *Systematic Reviews*, Vol. 5 No. 1, p. 15. <https://doi.org/10.1186/s13643-016-0191-y>
- Mönke, F. W., Bürger, A. S., Steinbrecher, J., Heinemann, H., Prüßmeier, P. L. and Schäpers, P. (2025), “When less is not more: Incomplete information in LinkedIn assessments and the moderating role of applicants’ résumé”, *Journal of Business and Psychology*, pp. 1-27. <https://doi.org/10.1007/s10869-025-10032-9>
- Pina, A., Petersheim, C., Cherian, J., Lahey, J. N., Alexander, G. and Hammond, T. (2023), “Using machine learning with eye-tracking data to predict if a recruiter will approve a resume”, *Machine Learning and Knowledge Extraction*, Vol. 5 No. 3, pp. 713-724. <https://doi.org/10.3390/make5030038>
- Radtke Caneppele, N., Ribeiro Serra, F. A., Contreras Pinochet, L. H. and Ramos Ribeiro, I. M. (2022), “Potential and challenges for using neuroscientific tools in strategic management studies”, *RAUSP Management Journal*, Vol. 57 No. 3, pp. 235-263. <https://doi.org/10.1108/RAUSP-01-2021-0014>
- Réveillé, C., Vergotte, G., Dray, G., Jean, P. A., Perrey, S. and Bosselut, G. (2025), “Trajectories of interbrain synchrony during teamwork: links with team composition and performance”, *Social Cognitive and Affective Neuroscience*, Vol. 20 No. 1, nsaf081. <https://doi.org/10.1093/scan/nsaf081>
- Su, Y. S., Suen, H. Y. and Hung, K. E. (2021), “Predicting behavioral competencies automatically from facial expressions in real-time video-recorded interviews”, *Journal of Real-Time Image Processing*, Vol. 18 No. 4, pp. 1011-1021. <https://doi.org/10.1007/s11554-021-01071-5>
- Waldman, D. A., Wang, D. and Fenters, V. (2019), “The added value of neuroscience methods in organizational research”, *Organizational Research Methods*, Vol. 22 No. 1, pp. 223-249. <https://doi.org/10.1177/1094428116642013>
- Wang, D., Waldman, D. A., Balthazard, P. A., Stikic, M., Pless, N. M., Maak, T., ... and Richardson, T. (2021), “Applying neuroscience to emergent processes in teams”, *Organizational Research Methods*, Vol. 24 No. 3, pp. 595-615. <https://doi.org/10.1177/1094428120915516>
- Zazon, D., Fink, L., Gordon, S. and Nissim, N. (2023), “Can NeuroIS improve executive employee recruitment? Classifying levels of executive functions using resting state EEG and data science methods”, *Decision Support Systems*, Vol. 168, 113930. <https://doi.org/10.1016/j.dss.2023.113930>
- Zhang, M., Jia, H. and Wang, G. (2021), “Interbrain synchrony of team collaborative decision-making: an fNIRS hyperscanning study”, *Frontiers in Human Neuroscience*, Vol. 15, 702959. <https://doi.org/10.3389/fnhum.2021.702959>
- Zito, M., Bilucaglia, M., Fici, A., Gabrielli, G. and Russo, V. (2021), “Job assessment through bioelectrical measures: A neuromanagement perspective”, *Frontiers in Psychology*, 12, 673012. <https://doi.org/10.3389/fpsyg.2021.673012>

## TRIGGERS OF STRATEGIC TRANSITION THROUGH SENSEMAKING IN NASCENT ENTREPRENEURSHIP

**Tea Josimovska**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
tea.josimovska@ek-inst.ukim.edu.mk*

**Daniel Nauck**

*h2 Magdesburg-Stendal University of Applied Sciences, Germany  
daniel.nauck@h2.de*

**Vladimir Trajkovik**

*Faculty of Computer Science and Engineering,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
trvlado@finki.ukim.mk*

### EXTENDED ABSTRACT

**Purpose** Decision-making under conditions of uncertainty has long been a central concern in various disciplines focusing on behavioral research (Bjurström and Bakken, 2022; Johnson and Busemeyer, 2010), particularly within the field of entrepreneurship (Shepherd *et al.*, 2015; Zayadin *et al.*, 2023). Entrepreneurial intention among emerging adults is not shaped by personality alone but is also strongly influenced by contextual enablers such as perceived support and opportunity structures (Apasieva *et al.*, 2024). In this context, sensemaking emerges as a multidimensional concept explaining how entrepreneurs interpret ambiguous signals, reassess assumptions, and construct meaning to guide pivotal decisions (Niemi *et al.*, 2022). Despite growing interest among researchers in entrepreneurial sensemaking *processes* and their antecedents, particularly within nascent entrepreneurial ventures, remains underexplored. This is particularly important since the nascent entrepreneurs are lacking established routines guiding their decision-making processes. In that vein, the sensemaking process is a meaningful mechanism for identifying opportunities, assessing them, and adapting to unexpected challenges (Maitlis and Christianson, 2014; Snowden, 2011; Sandberg and Tsoukas, 2015). It is a communicative process through which individuals construct meaning by using language, talk, and interaction to interpret ambiguous situations (Weick *et al.*, 2005). The paper focuses on the entrepreneurial sensemaking process during pivoting from a nascent venture to a sustainable business model. The aim is to explore the interaction mechanisms between the triggers - a disruption that challenges existing understanding and initiates sensemaking (Kocielnik and Hsieh, 2017; Obear, 2007) and enablers - routines, practices, and structures that facilitate the sensemaking process (Maitlis and Lawrence, 2007).

**Design/methodology/approach** This paper is grounded on a single case study from North Macedonia. Data were collected between January and March 2025 through a written interview and an in-depth, in-person interview with both founders. The analysis was based on the Gioia qualitative inductive research approach (Gioia *et al.*, 2013). In the first stage, 1st-order coding was applied to capture informant-centric terms and experiences of the founders with third-party actors, generating a wide array of descriptive codes related to emotional reactions, pivot decisions, and stakeholder interactions. In the second stage, 2nd-order coding grouped these categories into abstract themes reflecting deeper patterns in the entrepreneurial sensemaking process. Finally, these themes were organized into aggregate dimensions to identify underlying

mechanisms, particularly the role of specific triggers and enablers shaping how sense was made during the venture’s strategic transition.

**Findings** The analysis revealed distinct types of triggers, such as institutional, individual peer signals, market silence, as well as enabling conditions like urgency, credibility, and relational trust, shaping how and when entrepreneurial sensemaking was initiated and sustained. Additionally in the analysis were identified three distinct mechanisms through which triggers and enablers interact to shape sensemaking were identified: enabler-driven, trigger-driven, and interpretative breakdown. In enabler-driven mechanisms, sensemaking only occurred when the trigger was reinforced by internal or contextual enablers like technical knowledge or urgency. Trigger-driven mechanisms emerged when the trigger was strong or clear enough that sensemaking would likely occur in the same way with or without the presence of an enabler. In contrast, interpretative breakdown mechanisms reflected moments where the absence of clear signals or supportive enablers led to the new concept of the *sensemaking gap*. Finally, we reveal two previously underexplored dimensions of entrepreneurial sensemaking: *self-driven sensemaking* - emerging from pure internally motivated reflections, and *reactive sensemaking* - emerging from authoritative time-bound external pressures. Nevertheless, the findings have some limitations because they are based on a single case, which can limit their broader use and leave out other relevant aspects, such as cross-cultural factors.

**Originality/value** This paper offers original insight into how distinct types of triggers, such as institutional, individual peer signals, or market silence, interact with enabling conditions like urgency, credibility, and relational trust to shape entrepreneurial sensemaking in nascent entrepreneurial ventures. It moves beyond linear models of decision-making by emphasizing social dimensions through which meaning is constructed and actions are justified. The value lies in providing an understanding of how founders navigate the uncertainty, offering guidance on how support mechanisms can be better tailored to trigger productive sensemaking and timely strategic adaptation. In practical terms, the findings suggest that policy-makers should strengthen this process by supporting the development of entrepreneurship support organizations that offer mentorship programs, peer-learning platforms, and help build trust-based initiatives that can ultimately foster productive sensemaking in new ventures.

*Table 1: Triggers, Enablers, and Outcomes of the Sensemaking Process in Nascent Entrepreneurship*

Stage	Trigger	Enabler	Mechanism	Outcome
Early incubation phase	Environmental jolt	Technical knowledge/founder activism	Enabler Driven Mechanism	Immediate self-driven sensemaking
Market exploration phase				
Early transition phase	Accelerator invitation: B2B transition requirement	Stagnation and urgency (internal crisis)/Conditional invitation	Enabler Driven Mechanism	Reframed retrospective sensemaking
	Institutional	Authoritative, time-intensive	Trigger-Driven Mechanism	Immediate reactive sensemaking
	Individual peer trigger	Relational connection	Enabler Driven Mechanism	Relational sensemaking

	Absence of a signal (inaction or silence)	/	Interpretative Breakdown Mechanism	Sensemaking gap
	Industry disinterest	Knowledge about validation	Enabler-driven Mechanism	Reflective sensemaking
Late transition phase	Recognition of strategic misalignment	Temporal distance	Trigger-Driven Mechanism	Deliberative collective sensemaking
	Post accelerator investor excitement	Founders' self-confidence	Trigger-Driven Mechanism	Prospective sensemaking
	Market signal of willingness to pay	Legitimacy	Trigger-Driven Mechanism	Ongoing sensemaking

(Source: Author's work)

**Keywords:** Nascent entrepreneurship, Sensemaking, Pivot, Qualitative process research

**JEL classification:** L26, D83, M13

## REFERENCES

- Apasieva, T.J., Mitreva, M. and Čiković, K.F. (2024), "Crafting future entrepreneurs from emerging adults: what matters more - personality or context?", *International Journal of Education Economics and Development*, Vol. 15 No. 1-2, pp. 31-56. <https://doi.org/10.1504/IJEED.2023.10048768>
- Bjurström, E. and Bakken, B.T. (2022), "Decision-making under uncertainty: A Brehmerian approach", *Journal of Behavioural Economics and Social Systems*, Vol. 4 No. 2. <https://doi.org/10.54337/ojs.bess.v4i2.7749>
- Gioia, D.A., Corley, K.G. and Hamilton, A.L. (2013), "Seeking qualitative rigor in inductive research: Notes on the Gioia methodology", *Organizational Research Methods*, Vol. 16 No. 1, pp. 15-31. <https://doi.org/10.1177/1094428112452151>
- Johnson, J.G. and Busemeyer, J.R. (2010), "Decision making under risk and uncertainty", *Wiley Interdisciplinary Reviews: Cognitive Science*, Vol. 1 No. 5, pp. 736-749. <https://doi.org/10.1002/wcs.76>
- Kocielnik, R. and Hsieh, G. (2017), "Send me a different message: utilizing cognitive space to create engaging message triggers", In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*, Portland, OR, 25 February–1 March 2017. New York: ACM, pp. 2193-2207. <https://doi.org/10.1145/2998181.2998324>
- Maitlis, S. and Christianson, M. (2014), "Sensemaking in organizations: Taking stock and moving forward", *Academy of Management Annals*, Vol. 8 No. 1, pp. 57-125. <https://doi.org/10.1080/19416520.2014.873177>
- Maitlis, S. and Lawrence, T.B. (2007), "Triggers and enablers of sensegiving in organizations", *Academy of Management Journal*, Vol. 50 No. 1, pp. 57-84. <https://doi.org/10.5465/AMJ.2007.24160971>
- Niemi, L., Stenholm, P., Hakala, H. and Kantola, J. (2022), "Immanent sensemaking by entrepreneurs and the interpretation of consumer context", *International Small Business Journal*, Vol. 40 No. 8, pp. 966-990. <https://doi.org/10.1177/02662426221087870>

- Obear, K. (2007), “Navigating triggering events: Critical skills for facilitating difficult dialogues”, *The Diversity Factor*, Vol. 15 No. 3, pp. 23–29. <https://doi.org/10.1080/2194587X.2015.1091360>
- Sandberg, J. and Tsoukas, H. (2015), “Making sense of the sensemaking perspective: Its constituents, limitations, and opportunities for further development”, *Journal of Organizational Behavior*, Vol. 36 No. S1, pp. S6-S32. <https://doi.org/10.1002/job.1937>
- Shepherd, D.A., Williams, T.A. and Patzelt, H. (2015), “Thinking about entrepreneurial decision making: Review and research agenda”, *Journal of Management*, Vol. 41 No. 1, pp. 11-46. <https://doi.org/10.1177/0149206314541153>
- Snowden, D.J. (2011), “Naturalizing sensemaking”, In Mosier, K.L. & Fischer, U.M. (eds.) *Informed by Knowledge: Expert Performance in Complex Situations*. New York: Psychology Press, pp. 223-234. <https://doi.org/10.4324/9780203847985-22>
- Weick, K.E., Sutcliffe, K.M. and Obstfeld, D. (2005), “Organizing and the process of sensemaking” *Organization Science*, Vol. 16 No. 4, pp. 409-421. <https://doi.org/10.1287/orsc.1050.0133>
- Zayadin, R., Zucchella, A., Anand, A., Jones, P. and Ameen, N. (2023), “Entrepreneurs’ decisions in perceived environmental uncertainty”, *British Journal of Management*, Vol. 34 No. 2, pp. 831-848. <https://doi.org/10.1111/1467-8551.12612>

## APPENDIX

Table 2: Entrepreneurial transition process

	MyAir	AirXXX (active developed)		AirXXX (support only), EarthXXX.ai	
<b>phase</b>	<b>early incubation phase,</b> 2014 - 2018	<b>market exploration phase,</b> 2018 - 2019	<b>early transition phase,</b> 2020 - 2023	<b>late transition phase,</b> 2023-2024	<b>pre-seed phase,</b> 2025 - (...)
<b>status</b>	pre-startup venture creation			post-startup venture creation	
<b>motivation</b>	development focus		business model focus, customer orientation		
<b>main objective</b>	development of service application	internationalisation of service	business model incubation and experimentation, prototyping, customer research	business model product development (platform)	business model development

(Source: Author’s work)

## MAPPING RESEARCH ON AI AND CONSUMER PURCHASE INTENTION: BIBLIOMETRIC INSIGHTS (2009–2025)

**Snezana Ristevska-Jovanovska**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
snezana.ristevska-jovanovska@eccf.ukim.edu.mk*

**Ivona Serafimovska**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
ivona.serafimovska@eccf.ukim.edu.mk*

**Irena Bogoevska-Gavrilova**

*Faculty of Economics-Skopje,  
Cyril and Methodius University in Skopje, North Macedonia  
irena.bogoevska.gavrilova@eccf.ukim.edu.mk*

### EXTENDED ABSTRACT

**Purpose** Artificial Intelligence (AI) has rapidly become a core constituent of digital marketing, revolutionizing how companies interact with consumers. AI-driven tools, like chatbots and algorithmic product recommendations, personalized ads, and predictive analytics, give companies unprecedented capabilities to understand and influence consumer behavior in personalized, efficient, and scalable ways (Shoib and Hermawan, 2025). Reflecting this potential, industry analyses report that the use of AI in marketing escalated (e.g. increasing by an estimated 84% in 2020 alone) (Gera and Kumar, 2023). Research increasingly shows that AI technologies shape purchase intentions by leveraging adaptive machine learning, real-time data processing, and multimodal data integration (Zhang *et al.*, 2025). These capabilities allow firms to predict consumer behavior, personalize services, and enhance experiences, often fostering satisfaction and willingness to buy (Erliana, 2025; Lopes *et al.*, 2024; Guo *et al.*, 2024), positioning AI at the center of strategies designed to influence consumer behavior (Gansser and Reich, 2021; Kumar *et al.*, 2024). However, the influence of AI is not uniformly positive. Studies also highlight challenges like privacy concerns, security risks, and algorithmic biases that can undermine trust, with consumer attitudes, notably trust and perceived risk, acting as critical mediators of AI acceptance (Riandhi *et al.*, 2025). Although the corpus on AI in marketing has expanded markedly, extant reviews remain fragmented or overly broad, leaving the purchase-intention focus under-synthesized (Chen and Prentice, 2024; Lee *et al.*, 2023). With research output accelerating after 2020 and intensifying post-ChatGPT (2022), this study conducts a bibliometric and content analysis of literature from 2009–2025, which aims to map out publication trends, thematic concentrations, and emerging insights in this research stream. Specifically, our study addresses the following key research questions:

**RQ1:** What are the publication trends and patterns in research on AI consumer purchase intention from 2009 to 2025?

**RQ2:** What are the main themes and topics explored in the literature linking AI to consumer purchase intention?

**RQ3:** How have these research themes evolved, particularly in the last three years (2022–2025) following recent AI advancements (e.g., ChatGPT)?

**Design/methodology/approach** To ensure transparency in conducting the systematic literature

review (Lim and Rasul, 2022), we followed the four stages outlined in the PRISMA protocol: identification, screening, eligibility, and inclusion (Moher *et al.*, 2009). The initial search was performed on August 25, 2025, using the Web of Science database. We targeted article titles (TITLE), abstracts (ABS), and keywords (KEY) with the terms: "artificial intelligence" OR "machine learning" OR "deep learning" OR "natural language processing" AND "purchase intention" OR "buy\* intention" OR "willingness to buy" OR "intention to purchase" OR "intention to buy". Only journal articles published between 2009 and 2025 were considered, yielding 261 documents. After excluding one non-English article, 43 ineligible items (e.g., reviews, books, editorials), and 32 papers during abstract screening, the final dataset comprised 183 journal articles. This refined sample was subsequently analyzed using text mining in the latest version (1.6.20) of VOSviewer, producing visualizations of keyword co-occurrence across the whole period and the last three years, country co-authorship networks, publication trends, leading journals, and the most cited works. The three-year focus reflects the impact of ChatGPT's 2022 introduction, and although 2025 is ongoing, the results indicate a continuing upward trajectory.

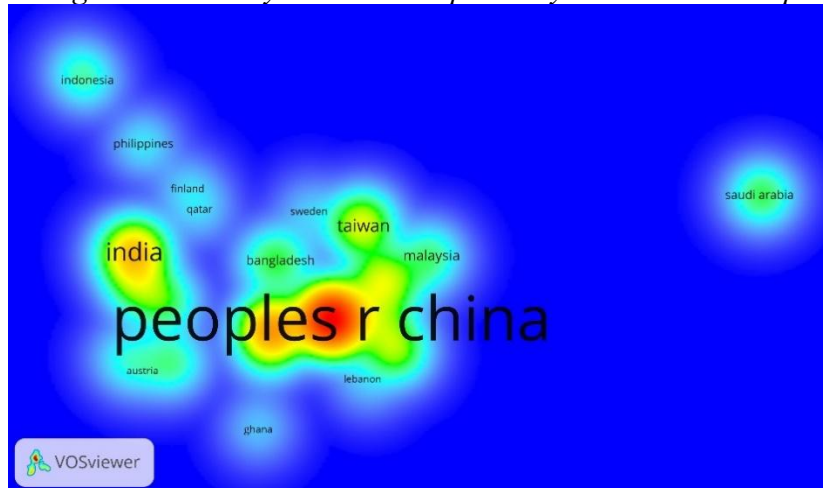
## Results and analysis

**Descriptive analytics** In this section, following the framework of Donthu *et al.* (2021), it is necessary to focus on two main techniques - performance analysis and science mapping. Table 1 provides the performance metrics, such as publication metrics, citation metrics, as well as a combination of both. The corpus comprises 183 publications across 16 active years (NAY), resulting in an average productivity per active year (PAY) of 11.44. Authorship patterns indicate a highly collaborative domain: 11 single-authored papers (6.0%) versus 172 co-authored (94.0%), with 602 contributing authors (NCA) and a collaboration index (CI) of 3.63. Impact indicators reinforce the field's visibility. The set has accumulated 15,195 total citations (TC), averaging 83.03 citations per document (AC), 83.61% of items are cited at least once (PCP), and citations per cited publication (CCP) equal 29.62. Collaboration intensity is further reflected in the collaboration coefficient (CC = 0.9399), confirming the dominance of multi-authored work and the need for complementary expertise. Altogether, these metrics portray a productive, influential, and highly collaborative research stream.

## Country co-authorship analysis

A country co-authorship map identified 52 countries, with the largest connected network comprising 43 nations across seven clusters (Figure 1). China leads with 73 publications (Cluster 3), followed by the United States (32, Cluster 6) and India (22, Cluster 1). This dominance of China aligns with broader AI research trends showing China as a leading producer of AI scholarship (Li and Rohayati, 2025), while the US and India's strong output is consistent with their established roles in technology and marketing research (Hue and Hung, 2025). Other notable contributors include South Korea (12), Taiwan (11), France (7), Germany, England, Japan, and Malaysia (6 each), and Australia (3). The clusters highlight strong European collaboration (Austria, France, Germany, Spain) alongside cross-regional groupings such as England–China–Japan, Canada–Bangladesh–Poland, Gulf/Asian partnerships, and links like USA–South Africa–Ghana and Egypt–Jordan–Saudi Arabia. Such patterns mirror observations in related bibliometric studies, where international collaboration is seen to bridge diverse research communities in AI applications (Li and Rohayati, 2025). Altogether, the country network suggests that consumer-AI research is highly international, with powerhouse countries driving output and fostering cross-border partnerships that bring together complementary knowledge and market contexts.

Figure 1: Country co-authorship density visualization map



(Source: Authors' depiction)

### **Keywords co-occurrence analysis based on text mining in the abstracts**

In this segment, we provide two network visualization maps depicting keyword co-occurrence patterns for the entire examined period (2009-2025) and the most recent years since the introduction of ChatGPT (2022-2025).

#### *Keywords co-occurrence analysis based on text mining in the abstracts for the whole analyzed period*

Through text mining of 183 abstracts (excluding structured abstract labels and copyright notices), a visualization map was constructed. The analysis produced 273 keywords meeting the minimum threshold, of which 31 were excluded, resulting in 242 keywords grouped into four clusters (Figure 2). In the network visualization, item size reflects frequency, lines represent co-occurrence strength, and selected labels are omitted to avoid overlap, revealing key thematic insights. The red cluster is dominated by “artificial intelligence” (49 occurrences) and closely related terms such as “acceptance”, “user acceptance”, “adoption”, “attitude”, and “perceived value”. These co-occurrences suggest that scholarships primarily investigate how consumers perceive and adopt AI-driven technologies. For example, Sohn and Kwon (2020) emphasize that traditional models must be adapted for novel AI products, and Gansser and Reich (2021) extend UTAUT specifically for AI contexts. Recent research similarly underscores the mediating role of trust and perceived quality in these models (Pathak and Bansal, 2024; Riandhi *et al.*, 2025), which is consistent with the cluster’s strong ties between “AI” and attitude-oriented keywords. In the green cluster, “purchase intention” (71 occurrences) serves as the most central keyword, indicating its pivotal role within this body of research. Its strong co-occurrence with terms such as “artificial intelligence” and “consumer behaviour” highlights a growing interest in how AI-driven technologies shape consumer decision-making processes (Lopes *et al.*, 2024). Connections with “engagement” and “social media” suggest that studies frequently examine interactive and digital environments as key contexts influencing purchase-related outcomes (Chen and Prentice, 2025). Meanwhile, links to “experience” and “trust” emphasize that both experiential factors and perceptions of credibility remain critical antecedents of purchase intention (Verhagen *et al.*, 2006; Riandhi *et al.*, 2025). The blue cluster centers on “behavior”, “machine learning”, “brand”, “information”, and “consumer satisfaction”. It reflects how machine learning methods are applied to analyze consumer actions, predict decision-making, and generate insights for brand strategy. The strong links to “information” and “consumer satisfaction” emphasize the importance of information quality and post-purchase evaluations, positioning this cluster at the intersection of behavioral





## REFERENCES

- Balakrishnan, J., and Dwivedi, Y. K. (2024), "Conversational commerce: Entering the next stage of AI-powered digital assistants", *Annals of Operations Research*, Vol. 333 No. 2, pp. 653-687. <https://doi.org/10.1007/s10479-021-04049-5>
- Chen, Y., and Prentice, C. (2025), "Integrating artificial intelligence and customer experience", *Australasian Marketing Journal*, Vol. 33 No. 2, pp. 141-153.
- de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., and Parasuraman, R. (2016), "Almost human: Anthropomorphism increases trust resilience in cognitive agents", *Journal of Experimental Psychology: Applied*, Vol. 22 No. 3, pp. 331-349. <https://doi.org/10.1037/xap0000092>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., and Lim, W. M. (2021), "How to conduct a bibliometric analysis: An overview and guidelines", *Journal of Business Research*, Vol. 133, pp. 285-296.
- Erliana, S. (2025), "A systematic literature review on artificial intelligence features driving purchase intention on web commerce: Insights into customer experience and trust using Python-based analysis", *Journal of Digital Business and Innovation Management*, Vol. 4 No. 2, pp. 442-454.
- Gansser, O. A., and Reich, C. S. (2021), "A new acceptance model for artificial intelligence with extensions to UTAUT2: An empirical study in three segments of application", *Technology in Society*, Vol. 65, 101535.
- Gera, R., and Kumar, A. (2023), "Artificial intelligence in consumer behaviour: a systematic literature review of empirical research papers published in marketing journals (2000-2021)", *Academy of Marketing Studies Journal*, Vol. 27 No. 1, pp. 1-15.
- Guo, Y., Tang, Q., Wang, H., Jia, M., and Wang, W. (2024), "Artificial intelligent housekeeper based on consumer purchase decision: A case study of online e-commerce", *Industrial Management and Data Systems*, Vol. 124 No.8, pp. 2588-2612.
- Hue, T. T., and Hung, T. H. (2025), "Impact of artificial intelligence on branding: A bibliometric review and future research directions", *Humanities and Social Sciences Communications*, Vol. 12 No. 1, p. 1-11. <https://doi.org/10.1057/s41599-025-04488-6>
- Kumar, V., Ashraf, A. R., and Nadeem, W. (2024), "AI-powered marketing: What, where, and how?", *International Journal of Information Management*, Vol. 77, 102783.
- Lee, W.-J., Lee, H.-S., and Cha, M.-K. (2023), "AI like ChatGPT, users like us: How ChatGPT drivers and AI efficacy affect consumer behaviour", *Virtual Economics*, Vol. 6 No. 4, pp. 44-59.
- Li, M., and Rohayati, M. I. (2024), "A bibliometric analysis of artificial intelligence applications in global higher education", *International Journal of Information System Modeling and Design*, Vol. 16 No. 1, pp. 1-24. <https://doi.org/10.4018/IJISMD.365913>
- Lim, W. M., and Rasul, T. (2022), "Customer engagement and social media: Revisiting the past to inform the future. *Journal of Business Research*, Vol. 148, pp. 325-342.
- Lopes, J. M., Silva, L. F., and Massano-Cardoso, I. (2024), "AI meets the shopper: Psychosocial factors in ease of use and their effect on e-commerce purchase intention", *Behavioral Sciences*, Vol. 14, No. 7, 616.
- Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. (2009), "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement", *Annals of Internal Medicine*, Vol. 151 No. 4, pp. 264-269.
- Mustak, M., Salminen, J., Plé, L., and Wirtz, J. (2021), "Artificial intelligence in marketing: Topic modeling, scientometric analysis, and research agenda", *Journal of Business Research*, 124(C), pp. 389-404. <https://ideas.repec.org/a/eee/jbrese/v124y2021icp389-404.html>

- Pathak, A., and Bansal, V. (2024), “Factors influencing the readiness for artificial intelligence adoption in Indian insurance organizations”, In S. K. Sharma, Y. K. Dwivedi, B. Metri, B. Lal, and A. Elbanna (Eds.), *Transfer, Diffusion and Adoption of Next-Generation Digital Technologies* (Vol. 698, pp. 43–55). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-50192-0\\_5](https://doi.org/10.1007/978-3-031-50192-0_5)
- Riandhi, A. N., Arviansyah, M. R., and Sondari, M. C. (2025), “AI and consumer behavior: Trends, technologies, and future directions from a Scopus-based systematic review”, *Cogent Business and Management*, Vol. 12, No. 1, 2544984.
- Shoib, A. A., and Hermawan, A. (2025), “The impact of artificial intelligence (AI) on consumer behavior in digital marketing: A systematic literature review”, *International Journal of Finance and Business Management*, Vol. 3 No. 3, pp. 161-176.
- Sohn, K., and Kwon, O. (2020), “Technology acceptance theories and factors influencing artificial intelligence-based intelligent products”, *Telematics and Informatics*, Vol. 47, 101324. <https://doi.org/10.1016/j.tele.2019.101324>
- Verhagen, T., Meents, S., and Tan, Y.-H. (2006), „Perceived risk and trust associated with purchasing at electronic marketplaces”, *European Journal of Information Systems*, Vol. 15 No. 6, 542-555. <https://doi.org/10.1057/palgrave.ejis.3000644>
- Zhang, Y., and Wang, S. (2023), “The influence of anthropomorphic appearance of artificial intelligence products on consumer behavior and brand evaluation under different product types”, *Journal of Retailing and Consumer Services*, Vol. 74, 103432.

## THE ROLE OF MESSAGE CREDIBILITY IN BUILDING INFLUENCER LOYALTY AND INCREASING PURCHASE INTENTIONS

**Irena Bogoevska-Gavrilova**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*irena.bogoevska.gavrilova@eccf.ukim.edu.mk*

**Anita Ciunova-Shuleska**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*anita.ciunova@eccf.ukim.edu.mk*

### EXTENDED ABSTRACT

**Purpose** Companies employ influencer marketing by collaborating with key social media figures whose content reaches potential consumers (Jiménez-Castillo and Sánchez-Fernández, 2019). Influencers with values aligned to the brands they represent help companies identify and attract target audiences (More and Lingam, 2017). By creating and sharing brand-related content, they can shape followers' attitudes and behaviors, as well as influence broader audiences (Liu *et al.*, 2015; Grafström *et al.*, 2018). However, the effectiveness of such influence also depends on how individuals perceive the message credibility, which can shape their attitudes (Appelman and Sundar, 2016). This attitude change depends on attention, comprehension, and willingness to accept information (Hovland *et al.*, 1953) and is influenced by factors related to source, message, receiver, and channel (Hovland *et al.*, 1953; Petty and Cacioppo, 1981). According to the Elaboration Likelihood Model, message credibility is central to persuasion when individuals engage in high cognitive effort through the central route (Petty and Cacioppo, 1986). Message credibility dimensions such as clarity, argument strength, emotional appeals, and presentation style affect both credibility perceptions and attitude change (Slater and Rouner, 1996). While different authors highlight different dimensions, message credibility is generally linked to content features such as quality, consistency, and trustworthiness (Wathen and Burkell, 2002). Related studies identify dimensions such as source, content, format, timeliness, accuracy, and delivery (Rieh and Belkin, 1998; Metzger *et al.*, 2003), with more recent work emphasizing argument strength and information quality as key dimensions of message credibility. (Li and Suh, 2015). These dimensions are also relevant in the context of human branding. Human brand theory (Thomson, 2006) offers insight into how a media figure can evolve into a human brand that fosters audience attachment and drives positive marketing outcomes by meeting followers' needs (Ki *et al.*, 2020). A human brand is a well-known person with unique qualities (like a distinctive personality or talent) who can be promoted just like a product. Such individuals build strong connections with their followers and gain loyalty by satisfying their needs (Duffy, 2005; Thomson, 2006). Additionally, social exchange theory emphasizes reciprocity, where interactions are shaped by the value of exchanged rewards (Gouldner, 1960; Malinowski, 1922; Blau, 1964; Molm, 1990). In the influencer-follower context, influencers share content that attracts attention, while followers respond through engagement, thereby reinforcing the relationship. When these exchanges remain balanced, they foster ongoing commitment and strengthen psychological ties, motivating individuals to maintain the relationship (Kim and Kim, 2021). Thus, reciprocity contributes to loyalty, as balanced exchanges enhance the stability and continuity of the influencer-follower relationship (Kim and Kim, 2021). Loyalty reflects the strong bond between consumers and companies, driving repeat purchases and shaped by customer

experience (Serman and Sims, 2022; Alam *et al.*, 2012; Bilgihan, 2016; Chang, 2017). Rooted in attachment theory, it involves forming lasting connections that foster commitment (Drigotas and Rusbult, 1992; Hazan and Shaver, 1994). In this study, influencer loyalty refers to social media users' dedication to both influencers and their brand-related posts (Bogoevska-Gavrilova, 2023).

The research aims to explore the influence of message credibility, operationalized through argument strength and information quality, on influencer loyalty, and to assess the subsequent effect of loyalty on consumers' purchase intention. This paper is the first to examine message credibility through the dimensions of argument strength and information quality in the context of influencer marketing, with a specific focus on influencer loyalty and purchase intention. By doing so, it addresses a key research gap concerning how message credibility fosters influencer loyalty and how this loyalty subsequently drives consumers' purchase intentions in the context of the central route of persuasion. We propose the following research hypotheses:

H1: Argument strength positively affects influencer loyalty.

H2: Information quality positively affects influencer loyalty.

H3: Influencer loyalty positively affects purchase intentions.

**Design/methodology/approach** To examine the proposed hypotheses, a quantitative approach was employed through structural equation modeling (SEM), conducted in two stages: confirmatory factor analysis (CFA) to validate the measurement model, and structural model testing using AMOS version 20.0. SEM is widely applied in marketing research due to its capacity to test models of consumer and market behavior (MacLean and Gray, 1998). The analysis was based on data from an online survey of social media users in North Macedonia, where 380 initial responses were collected and refined to a final sample of 307 valid cases, including only those who actively follow influencers. The sample is predominantly female (79.15%) and composed mainly of highly educated Generation Z respondents ( $\leq 27$  years) who are active social media users, especially on Instagram, spending 1.5-3.5 hours daily. Information quality and argument strength were assessed using a five-point Likert scale, with four items for each construct adapted from Li and Suh (2015). Influencer loyalty was measured through a seven-point Likert scale comprising four items adapted from Kim and Kim (2021).

Finally, purchase intention was captured with four items measured on a seven-point Likert scale, adapted from Gomes *et al.* (2022). All scales were translated from English into Macedonian and pilot tested before data collection to ensure clarity and validity. To address potential common method bias (CMB), Harman's single-factor test was conducted, and the unrotated factor analysis revealed no evidence of a dominant factor, as the single-factor solution explained less than 50% of the variance (in line with the recommendations of Podsakoff *et al.* (2003) and Fuller *et al.* (2016)). Cronbach's alpha ensured internal consistency for each construct, with all coefficients exceeding the recommended 0.7 threshold (Hair *et al.*, 2010). During CFA, squared multiple correlations, standardized covariance residuals, and standardized regression coefficients were carefully examined (Hair *et al.*, 2006). The results confirmed that the measurement model adequately fits the data. Table 1 reports the standardized regression coefficients and estimates for the hypothesized paths.

Table 1: Structural model estimates

	Estimates
H1: Argument strength → Influencer loyalty	0.370***
H2: Information quality → Influencer loyalty	0.251**
H3: Influencer loyalty → Purchase intention	0.556***
Notes: The estimates are standardized, and the level of significance is: ***p < 0.001, **p < 0.01, *p < 0.05, ns – not significant	

(Source: Authors' calculations)

**Findings** The results presented in Table 1 confirm all three proposed hypotheses. Argument strength exerts the strongest positive effect on loyalty, followed by information quality, while influencer loyalty significantly enhances customers' purchase intentions. The findings of this study align with a growing body of research in marketing and communication that highlights message credibility as a critical determinant of consumer responses. Prior work has demonstrated that high information quality in online messages enhances consumer loyalty (Pearson *et al.*, 2012), while more recent evidence shows that message context, particularly its informative value, strengthens followers' perceptions of influencer content and fosters loyalty (Zainudin *et al.*, 2023). Similarly, argument quality has been found to promote user loyalty by increasing satisfaction (Handayani *et al.*, 2020). Furthermore, loyalty consistently exerts a strong positive influence on purchase intention (Malik *et al.*, 2013; Khan *et al.*, 2019). Taken together, these studies confirm the pivotal role of information quality and argument strength as drivers of loyalty, positioning loyalty as a key mechanism through which message credibility translates into consumer purchase behavior.

**Originality/value** Based on the literature reviewed, no prior studies have explicitly examined message credibility as a two-dimensional construct (comprising argument strength and information quality) in relation to consumers' loyalty to an influencer and its subsequent effect on purchase intention. Instead, earlier research has predominantly focused on the direct influence of message credibility on purchase intention (Hocevar *et al.*, 2017; Balaji *et al.*, 2021; Chakraborty, 2019). This study is the first to explore message credibility in the context of influencer marketing with a specific emphasis on influencer loyalty. The findings extend current knowledge by showing that message credibility, operationalized through argument strength and information quality, is a key driver of loyalty, which in turn significantly increases purchase intentions. For marketing managers, these results underscore the need to go beyond creating entertaining content, emphasizing instead the development of influencer-generated messages that are both persuasive and rich in reliable, high-quality information. By prioritizing well-structured arguments and credible informational value, managers can strengthen followers' loyalty to influencers and, ultimately, enhance consumer purchase behavior in social media environments. Lastly, the study acknowledges its single-country focus as a cultural limitation. Yet, it offers a focused understanding of influencer marketing within one cultural setting and provides a basis for future cross-country comparisons.

**Keywords:** *Influencer marketing, Message credibility, Information quality, Argument strength, Influencer loyalty, Purchase intention*

**JEL classification:** *M31, M37*

## REFERENCES

- Alam, A., Usman Arshad, M. and Adnan Shabbir, S. (2012), "Brand credibility, customer loyalty and the role of religious orientation", *Asia Pacific Journal of Marketing and Logistics*, Vol. 24 No. 4, pp. 583-598.
- Appelman, A. and Sundar, S.S. (2016), "Measuring message credibility: Construction and validation of an exclusive scale", *Journalism & Mass Communication Quarterly*, Vol. 93 No. 1, pp. 59-79. <https://doi.org/10.1177/1077699015606057>
- Balaji, M.S., Jiang, Y. and Jha, S. (2021), "Nanoinfluencer marketing: How message features affect credibility and behavioral intentions", *Journal of Business Research*, Vol. 136, pp. 293-304. <https://doi.org/10.1016/j.jbusres.2021.07.049>
- Bilgihan, A. (2016), "Gen Y customer loyalty in online shopping: An integrated model of trust, user experience and branding", *Computers in Human Behavior*, Vol. 61, pp. 103-113.
- Blau, P. (1964), *Exchange and Power in Social Life*, Wiley, New York.
- Bogoevska-Gavrilova, I. (2023), "Unraveling Influencer Loyalty: Examining the Impact of Source Credibility", in Trpeska, M. (Ed.), *4<sup>th</sup> International Scientific Conference Economic and Business Trends Shaping the Future, 9-10 November 2023*, Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje, Skopje, North Macedonia, pp. 60-71. <http://doi.org/10.47063/EBTSF.2023.0005>
- Chakraborty, U. (2019), "The impact of source credible online reviews on purchase intention: The mediating roles of brand equity dimensions", *Journal of Research in Interactive Marketing*, Vol. 13 No. 2, pp. 142-161. <https://doi.org/10.1108/JRIM-06-2018-0080>
- Chang, H.H. (2017), "Gender differences in leisure involvement and flow experience in professional extreme sport activities", *World Leisure Journal*, Vol. 59 No. 2, pp. 124-139.
- Drigotas, S.M. and Rusbult, C.E. (1992), "Should I stay or should I go? A dependence model of breakups", *Journal of Personality and Social Psychology*, Vol. 62 No. 1, pp. 62-87.
- Duffy, D.L. (2005), "The evolution of customer loyalty strategy", *Journal of Consumer Marketing*, Vol. 22 No. 5, pp. 284-286. <https://doi.org/10.1108/07363760510611716>
- Fuller, C.M., Simmering, M.J., Atinc, G., Atinc, Y. and Babin, B.J. (2016), "Common methods variance detection in business research", *Journal of Business Research*, Vol. 69 No. 8, pp. 3192-3198.
- Gomes, M.A., Marques, S. and Dias, Á. (2022), "The impact of digital influencers' characteristics on purchase intention of fashion products", *Journal of Global Fashion Marketing*, pp. 1-18.
- Gouldner, A.W. (1960), "The norm of reciprocity: A preliminary statement", *American Sociological Review*, Vol. 25 No. 2, pp. 161-178.
- Grafström, J., Jakobsson, L. and Wiede, P. (2018), *The Impact of Influencer Marketing on Consumers' Attitudes*, Thesis, Jönköping University, International Business School.
- Hair, J.F., Black, B., Babin, B., Anderson, R.E. and Tatham, R.L. (2006), *Multivariate Data Analysis*, Prentice Hall, New Jersey.
- Hair, J.F., Black, W.C., Balin, B.J. and Anderson, R.E. (2010), *Multivariate Data Analysis*, Maxwell Macmillan International Editions.
- Handayani, P.W., Gelshirani, N.B., Azzahro, F., Pinem, A.A. and Hidayanto, A.N. (2020), "The influence of argument quality, source credibility, and health consciousness on satisfaction, use intention, and loyalty on mobile health application use", *Informatics in Medicine Unlocked*, Vol. 20, p.100429.
- Hazan, C. and Shaver, P.R. (1994), "Attachment as an organizational framework for research on close relationships", *Psychological Inquiry*, Vol. 5 No. 1, pp. 1-22.
- Hocevar, K.P., Metzger, M. and Flanagin, A.J. (2017), "Source credibility, expertise, and trust in health and risk messaging", in *Oxford Research Encyclopedia of Communication*, Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228613.013.287>.

- Hovland, C.I., Janis, I.K. and Kelley, H.H. (1953), *Communication and Persuasion*, Yale University Press, New Haven.
- Jiménez-Castillo, D. and Sánchez-Fernández, R. (2019), “The role of digital influencers in brand recommendation: Examining their impact on engagement, expected value and purchase intention”, *International Journal of Information Management*, Vol. 49, pp. 366-376. <https://doi.org/10.1016/j.ijinfomgt.2019.07.009>
- Khan, M.M., Memon, Z. and Kumar, S. (2019), “Celebrity endorsement and purchase intentions: The role of perceived quality and brand loyalty”, *Market Forces*, Vol. 14 No. 2.
- Ki, C.W.C., Cuevas, L.M., Chong, S.M. and Lim, H. (2020), “Influencer marketing: Social media influencers as human brands attaching to followers and yielding positive marketing results by fulfilling needs”, *Journal of Retailing and Consumer Services*, Vol. 55, p.102133.
- Kim, D.Y. and Kim, H.Y. (2021), “Trust me, trust me not: A nuanced view of influencer marketing on social media”, *Journal of Business Research*, Vol. 134, pp. 223-232.
- Li, R. and Suh, A. (2015), “Factors influencing information credibility on social media platforms: Evidence from Facebook pages”, *Procedia Computer Science*, Vol. 72, pp. 314-328.
- Liu, S., Jiang, C., Lin, Z., Ding, Y., Duan, R. and Xu, Z. (2015), “Identifying effective influencers based on trust for electronic word-of-mouth marketing: A domain-aware approach”, *Information Sciences*, Vol. 306, pp. 34-52. <https://doi.org/10.1016/j.ins.2015.01.034>
- MacLean, S. and Gray, K. (1998), “Structural equation modelling in market research”, *Journal of the Australian market research society*, Vol. 6 No. 1, pp. 17-32.
- Malik, M.E., Ghafoor, M.M., Iqbal, H.K., Riaz, U., Hassan, N., Mustafa, M. and Shahbaz, S. (2013), “Importance of brand awareness and brand loyalty in assessing purchase intentions of consumer”, *International Journal of Business and Social Science*, Vol. 4 No. 5, pp.167-171.
- Malinowski, B. (1922), *Argonauts of the Western Pacific*, Holt, Reinhart and Winston, New York.
- Metzger, M.J., Flanagin, A.J., Eyal, K., Lemus, D.R. and McCann, R.M. (2003), “Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment”, *Communication Yearbook*, Vol. 27 No. 1, pp. 293-335. [https://doi.org/10.1207/s15567419cy2701\\_10](https://doi.org/10.1207/s15567419cy2701_10)
- Molm, L.D. (1990), “Structure, action, and outcomes: The dynamics of power in social exchange”, *American Sociological Review*, Vol. 55 No. 3, pp. 427-447.
- More, J.S. and Lingam, C. (2019), “A SI model for social media influencer maximization”, *Applied Computing and Informatics*, Vol. 15 No. 2, pp. 102-108. <https://doi.org/10.1016/j.aci.2017.11.001>
- Pearson, A., Tadisina, S. and Griffin, C. (2012), “The role of e-service quality and information quality in creating perceived value: antecedents to web site loyalty”, *Information Systems Management*, Vol. 29 No. 3, pp. 201-215. <https://doi.org/10.1080/10580530.2012.687311>
- Petty, R.E. and Cacioppo, J.T. (1981), *Attitudes and Persuasion: Classic and Contemporary Approaches*, William C. Brown, Dubuque, IA. <https://doi.org/10.4324/9780429502156>
- Petty, R.E. and Cacioppo, J.T. (1986), “Elaboration likelihood model of persuasion”, in Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, Academic Press, San Diego, CA, pp. 123-205.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. and Podsakoff, N.P. (2003), “Common method biases in behavioral research: a critical review of the literature and recommended remedies”, *Journal of Applied Psychology*, Vol. 88 No. 5, pp. 879–903.

- Rieh, S.Y. and Belkin, N.J. (1998), "Understanding judgment of information quality and cognitive authority in the WWW", in *Proceedings of the 61st Annual Meeting of the American Society for Information Science*, pp. 279-289. <https://onlinelibrary.wiley.com/doi/10.1002/asi.10017>
- Serman, Z.E. and Sims, J. (2022), "Source credibility theory: SME hospitality sector blog posting during the Covid-19 pandemic", *Information Systems Frontiers*, pp. 1-18.
- Slater, M.D. and Rouner, D. (1996), "How message evaluation and source attributes may influence credibility assessment and belief change", *Journalism & Mass Communication Quarterly*, Vol. 73 No. 4, pp. 974-991. <https://doi.org/10.1177/107769909607300415>
- Thomson, M. (2006), "Human brands: investigating antecedents to consumers' strong attachments to celebrities", *Journal of Marketing*, Vol. 70 No. 3, pp. 104-119. <https://doi.org/10.1509/jmkg.70.3.104>
- Wathen, C.N. and Burkell, J. (2002), "Believe it or not: Factors influencing credibility on the Web", *Journal of the American Society for Information Science and Technology*, Vol. 53 No. 2, pp.134-144. <https://doi.org/10.1002/asi.10016>
- Zainudin, M.I., Wei, E.S. and Mangsor, M.I. (2023), "The relationship between endorsers' message context, credibility and brand loyalty among young generations' fashion consumers", in Zainudin, M.I. and Rahmat, H. (Eds.), *Proceedings of the 4th International Conference on Communication, Language, Education and Social Sciences (CLESS 2023), Advances in Social Science, Education and Humanities Research*, Vol. 819, Atlantis Press, pp. 17-32. [https://doi.org/10.2991/978-2-38476-196-8\\_3](https://doi.org/10.2991/978-2-38476-196-8_3)

## DOCTORAL SYMPOSIUM PAPERS

This section features research contributions by doctoral candidates presented at the Doctoral Symposium.

### DIGITAL TRANSFORMATION IN TERMS OF IMPROVING THE PERFORMANCE OF COMPANIES IN NORTH MACEDONIA

**Bojana Antikj**

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*bojana.antik@yahoo.com*

#### EXTENDED ABSTRACT

**Purpose** The digital economy has become a key driver of economic growth in the modern world, influencing industries, business practices, and consumer behavior (Raji *et al.*, 2024). For the Republic of North Macedonia, digital transformation (Casey, 2015) offers significant potential for modernizing its economy, improving productivity, and increasing competitiveness globally (Thurow, 2000). This paper explores the contribution of digitalization, how it affects online sales, the importance of digital skills that employees should possess, and the contribution of social media to increasing sales. By examining the key elements (Sutinen and Närvänen, 2022) of companies' digitalization, their presence on social media, and the level of online sales, we can see the effect of digitalization (Gonzalez-Verona *et al.*, 2021) on the sales success (Gupta and Israney, 2025) achieved by companies. The high level of digitalization is considered a key factor that positively reflects the increase in online sales. Furthermore, we can process the impact of digitalization (Abou-Foul *et al.*, 2020) on the broader economy, including the creation of innovations and investments (Lytras and Serban, 2025) in the development of employees' digital skills (Stofkova *et al.*, 2022), which directly affect the level of effectiveness of companies. Digital skills (Audrin *et al.*, 2024) are reflected in the overall operation, and a special focus is on sectors such as digital finance, e-commerce (Suganya *et al.*, 2025), and manufacturing. However, all these positive effects are based on digital tools (Helsper and Van Deursen, 2015) and technological innovations that enable significantly better and faster exchange of information and more efficient communication. In fact, this communication and contact with customers is often carried out through social media (Mountinga *et al.*, 2011). It also highlights the opportunities for performing on social media (Nafees *et al.*, 2020), through Instagram and Facebook profiles, where companies are presented and followed by consumers (Cheung *et al.*, 2009). In this way, companies present themselves on the market and get closer to customers (Manthiou *et al.*, 2014). Buyers have the opportunity to purchase products and satisfy their desires and needs in this facilitated way. Of course, digital transformation brings a number of challenges and opportunities related to the digital transformation in North Macedonia. The digital economy has the potential to be a central component of North Macedonia's long-term economic strategy, fostering sustainable growth, improving living standards, and positioning the country as a more competitive player in the market.

Based on the above, the aim of this study is to examine the effect of digitalization on the development and improvement of the economy.

The paper is based on three hypotheses:

H1: The level of digitalization of companies is positively related to the level of online sales.

H2: The level of companies' presence on social media is positively correlated with the level of online sales.

H3: The level of digital skills of employees is positively related to the level of online sales.

**Design/methodology/approach** An online questionnaire will be distributed to business owners, managers, entrepreneurs, and IT professionals in North Macedonia. 100 respondents will actively participate in the survey and will answer identical questions. In order to test the hypotheses, a quantitative analysis of the collected data will be conducted. The instrument for data collection will be a survey questionnaire of the type of closed questions according to the Likert scale principle, with a total of five levels. The survey will be composed of three parts, namely: Questions related to digital development, such as whether and how much companies are investing in the digital development of companies, whether managers are striving to implement digital technologies, i.e. whether they see the future here, how online sales are monitored, whether there are analytical tools, whether the company has a strategy for digital transformation and improvement. The second part will be based on social media as a way of presenting the company. So, the questions would be of the type of whether social media is part of the operation, what type of posts are placed by the company, whether communication with users is regular through comments and messages, the attitude and willingness to advertise the company through paid ads on Facebook and Instagram. The third part will be based on the independent variable - digital skills of employees. The questionnaire will be aimed at understanding the level of knowledge of basic digital skills, the frequency of training intended for digital technologies, the readiness for online data analysis, the success of online sales, and whether it is without downtime and major technical problems. The analysis will be based on structural equation modeling (SEM) in two stages: Confirmatory factor analysis (CFA), and the structural model will be tested using software (SPSS).

Through structural equation modeling (SEM) (Malhotra *et al.*, 2017), an assessment of the structural relationships between different constructs will be obtained, which will be represented by several different indicators in an integrated and defined model. The statistical method of regression will be used in order to determine the influence of the dependent variable on the independent variable.

A potential limitation of questionnaires is the reliability of the data obtained. The accuracy of the data obtained can always be questioned. Not all respondents will be open to giving honest answers for various reasons. Some will want to hide their strategy, some will misrepresent themselves, and so on.

**Expected findings** We can conclude that in the first hypothesis, the dependent variable is online sales, and the independent variable is the level of digitalization. The questionnaire determines how much digitalization contributes to increasing sales, and the information can be obtained through online sales reports. In fact, through the answers to the number of orders placed and sales made via the internet, the questionnaire allows us to see the percentage of sales increase after the implementation of digital elements. With this, we see the effectiveness of digitalization on the level of sales increase. The research for the second hypothesis is conducted by researching the company's social media profiles and the number of followers on them. It is a kind of indicator of the level of interest in the brand, the recognition that leads to sales. The fact that someone follows a brand profile speaks of interest. The third hypothesis is based on the level of influence of employees' digital skills on increasing online sales. The questionnaire will assess the readiness and awareness of employees, as well as the company's investments in their education for digitalization. This can be seen through the effects they

achieved before and after the education, whether the education led them to think about implementing new digital solutions in order to increase sales. In this way, it will be seen whether and how digitalization affects the performance and operations of companies and their economic results. Online sales are not just a technological process but the result of a combination of technology, human factors, and communication.

**Originality/value** Digitalization has been considered for years as a process that infiltrates all parts of human life and work. There are numerous studies that consider digitalization as a process that affects different processes separately, but not as an integrated process that connects all economic sectors. The research is expected to expand the scientific framework for digital marketing, and it will achieve this by integrating the three elements: digital development, social media, and employee skills. The survey that will be conducted among companies will provide empirical evidence on the degree of digitalization of companies and its impact on sales. The results will also show how successful companies present themselves on social media, and how for them social media is a way to connect with customers in order to increase sales. It will be seen which content reaches customers the most. On the other hand, emphasis will be placed on employee education, because only employees who know the importance of digitalization and of course, possess the knowledge to implement changes in the implementation of digital practices in the daily operations of the company, contribute the most to increasing sales. We can conclude that digitalization reaches its maximum if it is applied in all sectors; in this way, full capacity will be achieved, and the economy will grow. The paper will help to consolidate theoretical knowledge and practical application in the daily operations of companies, providing a clear picture of the effects of the application of the independent variables, digital development, social media, and digital skills of employees on the dependent variable, online sales.

**Keywords:** *Digital economy, Economic growth, Digitalization, Transformation, Infrastructure*

**JEL classification:** *O33, M15, L25*

## REFERENCES

- Abou-Foul, M., Ruiz-Alba, J.L. and Soares, A. (2021), "The impact of digitalization and servitization on the financial performance of a firm: an empirical analysis", *Production Planning & Control*, Vol. 32 No. 12, pp. 975-989.
- Audrin, B., Audrin, C. and Salamin, X. (2024), "Digital skills at work – conceptual development and empirical validation of a measurement scale", *Technological Forecasting and Social Change*, Vol. 202, 123279.
- Casey, M. (2015), "Why media preservation can't wait: the gathering storm", *IASA Journal*, No. 44, pp. 14-22.
- Cheung, C.M., Luo, C., Sia, C.L. and Chen, H. (2009), "Credibility of electronic word-of-mouth: informational and normative determinants of online consumer recommendations", *International Journal of Electronic Commerce*, Vol. 3 No. 4, pp. 9-38.
- González-Varona, J.M., López-Paredes, A., Poza, D. and Acebes, F. (2021), "Building and development of an organizational competence for digital transformation in SMEs", *Journal of Industrial Engineering and Management*, Vol. 14 No. 1, pp. 15-24.
- Gupta, S.L. and Israney, R.S. (2025), *Service Marketing: Text and Cases*, Sultan Chand and Sons, New Delhi.

- Helsper, E.J. and Van Deursen, A.J.A.M. (2015), "Digital skills in Europe: research and policy", in Andreasson, K. (Ed.), *Digital Divides: The New Challenges and Opportunities of E-Inclusion*, CRC Press, Boca Raton, pp. 125-149.
- Lytras, M.D., Serban, A.C. and Ntanos, S. (2025), "Exploring distance learning in higher education: satisfaction and insights from Mexico, Saudi Arabia, Romania, Turkey", *Journal of Business Economics and Management*, Vol. 26 No. 1, pp. 1-20.
- Madanchian, M. (2024), "The impact of artificial intelligence marketing on e-commerce sales", *Systems*, Vol. 12 No. 10, p. 429.
- Malhotra, N.K., Nunan, D. and Birks, D.F. (2017), *Marketing Research: An Applied Approach*, Pearson Education.
- Manthiou, A., Lee, S., Tang, L.R. and Ching, L. (2014), "The experience economy and the visitors' behavioral consequences: an empirical study on VEISHEA event", in *The 16th Annual Graduate Student Research Conference in Hospitality and Tourism, Houston, TX, USA*.
- Mountinga, D., Moorman, M. and Smit, E. (2011), "Introducing COBRAs: exploring motivations for brand-related social media use", *International Journal of Advertising*, Vol. 30 No. 1, pp. 13-46.
- Nafees, L., Cook, C.M. and Stoddard, J.E. (2020), "The impact of the social media influencer power on consumer attitudes toward the brand: the mediating/moderating role of social media influencer source credibility", *Atlantic Marketing Journal*, Vol. 9 No. 1, p. 3.
- Raji, M.A., Olodo, H.B., Oke, T.T., Addy, W.A., Ofodile, O.C. and Oyewole, A.T. (2024), "E-commerce and consumer behavior: a review of AI-powered personalization and market trends", *GSC Advanced Research and Reviews*, Vol. 18 No. 3, pp. 066-077.
- Stofkova, J., Poliakova, A., Stofkova, K.R., Malega, P., Krejnus, M., Binasova, V. and Daneshjo, N. (2022), "Digital skills as a significant factor of human resources development", *Sustainability*, Vol. 14 No. 20, 13117.
- Suganya, R., Dhanaraj, V.T. and Ajithkumar, R. (2025), *E-commerce and Technology*, Wordsswiggle Publication, New Delhi.
- Sutinen, U.-M. and Närvänen, E. (2022), "Constructing the food waste issue on social media: a discursive social marketing approach", *Journal of Marketing Management*, Vol. 38 Nos. 3-4, pp. 239-240.
- Thurow, L.C. (2000), "Globalization: the product of a knowledge-based economy", *The Annals of the American Academy of Political and Social Science*, Vol. 570 No. 1, pp. 19-31.

## **SOCIAL CAPITAL AND THE ROLE OF SOCIAL BROKERS IN AI (NON) ADOPTION IN DEVELOPING COUNTRIES**

**Mijalche Santa**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
mijalche.santa@eccf.ukim.edu.mk*

**Blerton Zejneli**

*Faculty of Economics-Skopje,  
Ss. Cyril and Methodius University in Skopje, North Macedonia  
blertonzejneli@gmail.com*

### **ABSTRACT**

*This research explores how social capital supports the adoption of artificial intelligence (AI) in developing countries, focusing on the role of "social brokers." A social broker is a trusted individual who occupies a unique position within a network, connecting individuals from different networks or maintaining connections with a larger number of individuals within the existing network. Based on input from the initial phase of the project, conducted in a developing country with high internet use but low AI adoption, we use qualitative research methods to better understand the practical aspects of AI adoption. Our early findings suggest that AI adoption goes beyond the right technology or skills and is strongly influenced by trusted communities and networks that shape decisions about AI adoption. "Social brokers" play a key role in this process. They help close knowledge gaps, address concerns of people who have not adopted AI or have adopted it at a low level, and show how AI can be relevant and useful for specific jobs and tasks. These "social brokers" are often seen as trusted friends, technology influencers, former colleagues, or respected local industry experts. Their presence and activities in tightly connected social networks appear to be very important for reducing the gap in AI adoption. The next phase of this research will focus on identifying the aspects of social capital that influence AI adoption, understanding the relationships that help overcome resistance to adopting AI, and developing strategies that use social capital to encourage faster AI adoption in developing countries.*

**Keywords:** *AI, Technology adoption, Social brokers, Developing countries*

**JEL classification:** *O31, O32, O33*

### **1. INTRODUCTION**

Artificial Intelligence (AI) is rapidly transforming various industries around the globe and presents a tremendous opportunity for development across industries, professions, and national boundaries (Agrawal *et al.*, 2019). As generative AI is truly a general-purpose technology (McAfee, 2024), it has already begun to significantly alter our ways of living and working (Gordon and Gunkel, 2024). AI's global impact stems from its accessibility, affordability, and ease of use. This provides a unique opportunity for developing countries to overcome any existing gaps with developing countries and gain a competitive position (Alonso *et al.*, 2020; Aly, 2020; Fan and Qiang, 2024) or at least minimize inequality due to technological change (Alonso *et al.*, 2022; Freire, 2024). However, there is a very limited adoption of AI among the developing countries (Khan *et al.*, 2024). Thus, the question is how AI adoption in developing

countries can be increased? This is important because “the ultimate impact of generative AI on the economy depends on how quickly and intensively the technology is adopted” (Bick *et al.*, 2024).

The existing research has identified key reasons for delayed AI adoption in developing countries, such as technological infrastructure, economic stability, policies, financial constraints, and workforce readiness (Ali *et al.*, 2024; Al-Zahrani and Alasmari, 2025). However, developing or improving these conditions might take years, delaying the adoption and impact of artificial intelligence on these economies. Nevertheless, early research in the adoption of AI in the United States identifies inertia and adjustment costs as barriers to AI adoption (Eastwood, 2024; McElheran *et al.*, 2023). Overcoming this and identifying how AI can be embedded in everyday work practices may speed up the adoption. In this direction, we are conducting research on how social capital can be used to overcome the inertia of AI adoption.

Despite widespread connectivity, AI uptake remains concentrated among a minority of firms and individuals, especially large “high performers”, while SMEs lag due to skills gaps, weak data readiness, uncertain ROI, compliance risk, and integration costs. Enabling factors include leadership commitment, workforce training, robust data infrastructure, and access to collaborative ecosystems (Eurostat, 2025; McKinsey, 2024; OECD, 2025). This paradox also appears in the European Union countries, where only about one in seven enterprises used AI in 2024, even though digital penetration is near universal, with roughly a threefold gap between large and small firms (Eurostat, 2025; OECD, 2025). Regarding social capital, the scholarly consensus, drawing on classic theory, is that strong ties within and between organizations build trust, ease knowledge sharing, and raise absorptive capacity, which in turn improve AI readiness and adoption. Recent empirical studies find that SMEs with richer relational networks are better positioned to identify use cases and manage cyber risk, and that collaboration ties to peers with AI experience accelerate uptake in scientific contexts (Nahapiet and Ghoshal, 1998; Ode *et al.*, 2025; Bianchini *et al.*, 2025). At the same time, social capital is not unambiguously beneficial; overly dense inward-looking networks can reduce autonomy and slow product innovation, which cautions against closed or hierarchical configurations when implementing AI (Wang *et al.*, 2025).

Social capital refers to the resources individuals or groups acquire through their relationships within social networks, highlighting the mutual benefits that arise from trust, norms and collaboration, which strengthen social cohesion and cooperation (Nahapiet and Ghoshal, 1998; Putnam, 1994).

Through this research, we seek to address the following questions and provide insights into their implications:

RQ1: How do the structural, cognitive, and relational dimensions of social capital influence AI adoption in organizations?

RQ2: What specific elements of social capital contribute to overcoming inertia in AI adoption within organizations?

RQ3: What strategies leveraging social capital can be employed to accelerate AI adoption in organizations in developing countries?

We are conducting our research in North Macedonia, a developing country with high internet penetration and competitive average net salaries. Given this context, the very low adoption of AI can neither solely be attributed to a lack of digital access nor the low purchasing power.

Instead, we believe that it points to more complex factors that we aim to explore.

## **2. METHODOLOGY**

At this stage, we are performing exploratory interviews with a diverse set of professionals to get a better understanding of their awareness of AI and benefits, concerns, usage patterns, and motivation. The professionals selected for the exploratory interviews were chosen because they represent diverse industries, roles, and levels of familiarity with digital technologies, which provides a broad spectrum of perspectives on AI adoption. By engaging individuals with varying experiences, expertise, and organizational contexts, the research captures a more comprehensive understanding of awareness, perceived benefits, concerns, usage patterns, and motivations surrounding AI. This diversity enhances the reliability of the findings, ensures that the analysis is not narrowly confined to one sector or professional background, and allows for richer insights into the social and organizational dynamics that influence AI adoption.

In the next stage of interviews, we will explore the nuances of social capital. By focusing on qualitative data and two cycles of interviews, we aim to uncover insights into the perceptions and behaviors influencing adoption. We utilize Reflexive Thematic Analysis (RTA) (Braun *et al.*, 2019) to analyse and interpret qualitative data collected through interviews. Interpretive qualitative research provides a comprehensive approach to exploring human experiences in particular contexts (Rahman, 2016), allowing for a deeper understanding of how social capital impacts adoption.

Reflexive Thematic Analysis is appropriate because it offers a flexible yet rigorous approach for identifying patterned meanings across interviews while preserving the contextual richness of participants' accounts. It aligns with our exploratory, interpretivist stance and two-cycle design by supporting iterative movement between data and analysis, allowing themes to evolve as understanding deepens. RTA emphasizes the researcher's active role in knowledge production and requires reflexivity, which is essential when examining how social capital shapes perceptions, motivations, and behaviors. Its compatibility with semi-structured interviews, transparent auditability through coding notes and analytic memos, and capacity to capture both semantic and latent meanings make it well-suited to unpack the nuanced mechanisms through which relationships, trust, and norms influence AI adoption.

## **3. INITIAL FINDINGS**

While our initial findings reaffirm the existing key challenges in AI adoption, including a lack of awareness and privacy concerns, they surface a unique factor named "social brokers" that play a significant role in bridging the adoption gap. Through our analysis, we have identified the emergence of a main theme and three sub-themes:

- Main Theme: Social brokers are key catalysts for AI adoption: Social brokers facilitate and shape attitudes towards AI.
- Subtheme 1: Social brokers bridging information gaps: Social brokers reduce uncertainty about AI by providing trusted information.
- Subtheme 2: Social brokers showing "how": Social brokers offer hands-on experiences on AI for specific tasks.
- Subtheme 3: Social brokers alleviating perceived risks: Social brokers help reduce perceived risks in adopting AI.

#### 4. CONCLUSION

Our initial findings about the key role of “social brokers” bring to light a new perspective of looking at AI adoption that goes beyond the traditional. Many developing countries are characterized by strong social ties where trust-based networks often drive decision-making. Our findings on the adoption of AI in a developing country are unique in the specific context but align with the general literature on social influence on technology adoption (Vannoy and Palvia, 2010; Hasija and Esper, 2022).

While these findings provide valuable insights in a developing country context, their generalisability is currently limited due to the geographical limitation of this project. Future research should extend this study across other developing countries. As we continue, we will explore the social capital dimensions and their impact on AI adoption. By focusing on qualitative data and using two cycles of interviews, this study aims to uncover nuanced insights into the perceptions and behaviors influencing AI adoption.

This study contributes to the literature on technology adoption by introducing the concept of social brokers as a critical mechanism for AI adoption in developing-country contexts. Existing frameworks such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) emphasize perceived usefulness, ease of use, and social influence as key determinants of adoption. Our findings refine these models by showing that in dense, trust-based networks, social influence operates less through diffuse peer norms and more through the mediation of structurally central and trusted actors. Brokers act as translators, validators, and endorsers of new technologies, thereby shaping perceptions of usefulness and ease of use and mitigating perceived risks.

This perspective also challenges assumptions within diffusion of innovation theory, which traditionally highlights the role of early adopters and innovators. Instead, our evidence suggests that adoption is often accelerated by brokers who bridge otherwise disconnected groups, facilitating cross-cluster diffusion even when technological uncertainty is high. Integrating insights from social capital theory, we argue that the structural position (brokerage), relational trust, and cognitive alignment of these actors jointly create conditions for effective adoption. Finally, in contexts where institutional trust and formal evaluation mechanisms are weak, brokers serve as substitutes for missing intermediaries, highlighting boundary conditions under which their influence is most pronounced. Collectively, these contributions extend and refine dominant theories by foregrounding the structural and relational mechanisms that drive AI adoption under uncertainty in developing countries.

#### REFERENCES

- Agrawal, A., Gans, J.S. and Goldfarb, A. (2019), "Exploring the impact of artificial intelligence: prediction versus judgment", *Information Economics and Policy*, Vol. 47, pp. 1-6.
- Ali, H., ul Mustafa, A. and Aysan, A.F. (2025), "Global adoption of generative AI: what matters most?", *Journal of Economy and Technology*, Vol. 3, pp. 166-176.
- Alonso, C., Berg, A., Kothari, S., Papageorgiou, C. and Rehman, S. (2022), "Will the AI revolution cause a great divergence?", *Journal of Monetary Economics*, Vol. 127, pp. 18-37.
- Alonso, C., Kothari, S. and Rehman, S. (2020), "How artificial intelligence could widen the gap between rich and poor nations", *IMF Blog*.
- Aly, H. (2020), "Digital transformation, development and productivity in developing countries:

- is artificial intelligence a curse or a blessing?", *Review of Economics and Political Science*, Vol. 7 No. 4, pp. 238-256.
- Al-Zahrani, A.M. and Alasmari, T.M. (2025), "A comprehensive analysis of AI adoption, implementation strategies, and challenges in higher education across the Middle East and North Africa (MENA) region", *Education and Information Technologies*, pp. 1-51.
- Bianchini, S., Müller, M. and Pelletier, P. (2025), "Drivers and barriers of AI adoption and use in scientific research", *Technological Forecasting and Social Change*, Vol. 220, p. 124303.
- Bick, A., Blandin, A. and Deming, D.J. (2024), "The Rapid Adoption of Generative AI", working paper no. 32966, National Bureau of Economic Research.
- Braun, V. and Clarke, V. (2021), "Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches", *Counselling and Psychotherapy Research*, Vol. 21 No. 1, pp. 37-47.
- Braun, V., Clarke, V., Hayfield, N. and Terry, G. (2019), "Thematic analysis", in Liamputtong, P. (Ed.), *Handbook of Research Methods in Health Social Sciences*, Springer, pp. 843-860.
- Eastwood, B. (2024), "The who, what, and where of AI adoption in America", MIT Sloan Ideas Made to Matter, 7 February, available at: <https://mitsloan.mit.edu/ideas-made-to-matter/who-what-and-where-ai-adoption-america>
- Eurostat (2025), "Usage of AI Technologies Increasing in EU Enterprises", 23 January.
- Fan, Q. and Qiang, C. (2024), "Tipping the scales: AI's dual impact on developing nations", *World Bank Blogs*.
- Freire, C. (2025), "Is this time different? Impact of AI in output, employment and inequality across low, middle and high-income countries", *Structural Change and Economic Dynamics*, Vol. 73, pp. 136-157.
- Gordon, J.S. and Gunkel, D.J. (2025), "Artificial intelligence and the future of work", *AI & Society*, Vol. 40 No. 3, pp. 1897-1903.
- Hasija, A. and Esper, T.L. (2022), "In artificial intelligence (AI) we trust: a qualitative investigation of AI technology acceptance", *Journal of Business Logistics*, Vol. 43 No. 3, pp. 388-412.
- Khan, M.S., Umer, H. and Faruqe, F. (2024), "Artificial intelligence for low income countries", *Humanities and Social Sciences Communications*, Vol. 11 No. 1, pp. 1-13.
- McAfee, A. (2024), "Generally Faster: The Economic Impact of Generative AI", Google Technology & Society Visiting Fellow, available at: <https://your-url-here.com>
- McElheran, K., Li, J.F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L.S. and Zolas, N. (2023), "AI Adoption in America: Who, What, and Where", working paper no. 31788, National Bureau of Economic Research.
- McKinsey & Company (2024), "The State of AI in Early 2024: Gen AI Adoption Spikes and Starts to Generate Value", 30 May.
- Nahapiet, J. and Ghoshal, S. (1998), "Social capital, intellectual capital, and the organizational advantage", *Academy of Management Review*, Vol. 23 No. 2, pp. 242-266.
- OECD (2025), "Emerging Divides in the Transition to Artificial Intelligence", OECD Publishing.
- OECD, BCG and INSEAD (2025), "The Adoption of Artificial Intelligence in Firms", OECD Publishing.
- Ode, E., Awolowo, I.F., Nana, R. and Olawoyin, F.S. (2025), "Social capital and artificial intelligence readiness: the mediating role of cyber resilience and value construction of SMEs in resource-constrained environments", *Information Systems Frontiers*, pp. 1-21.
- Putnam, R.D. (1994), "Social capital and public affairs", *Bulletin of the American Academy of Arts and Sciences*, Vol. 47 No. 8, pp. 5-19.
- Rahman, M.S. (2016), "The advantages and disadvantages of using qualitative and quantitative approaches and methods in language 'testing and assessment' research: a literature review",

*Journal of Education and Learning*, Vol. 6 No. 1, pp. 102-112.

Vannoy, S.A. and Palvia, P. (2010), "The social influence model of technology adoption", *Communications of the ACM*, Vol. 53 No. 6, pp. 149-153.

Wang, Z., McNally, R., Lenihan, H. and Li, C. (2025), "Social capital as a barrier to innovation: the mediating role of job autonomy", *Finance Research Letters*, Vol. 72, 106530.

## CLASSIFICATION OF A FOREIGN EXCHANGE DIFFERENCE FROM AN INTRAGROUP MONETARY LIABILITIES AND ASSETS IN MULTINATIONAL TELECOMUNICATION COMPANIES UNDER IFRS18

**Miroslav Serafimoski**

*Deutsche Telekom Services Europe (DTSE)*

*Faculty of Economics-Skopje,*

*Ss. Cyril and Methodius University in Skopje, North Macedonia*

*serafimoskim@telekom.de*

### **ABSTRACT**

*This paper analyses the classification of foreign exchange differences arising from intragroup monetary items under IFRS 18 Presentation and Disclosure in Financial Statements. IFRS 18 introduces a new structure for the statement of profit or loss, requiring income and expenses to be classified into operating, investing, financing, and other specified categories, with paragraph B65 mandating that foreign exchange differences be classified consistently with the income and expenses of the underlying item. So, the aim of this paper is to analyze and make a recommendation on how the foreign exchange differences from intragroup transactions will be classified in the IFRS18 structure of the Statement of Profit and Loss. However, intragroup income and expenses are eliminated under IFRS 10, while IAS 21 requires exchange differences on intragroup monetary items to remain recognized in consolidated profit or loss — creating a classification challenge. The paper evaluates five theoretical options and explains why the IFRS Interpretations Committee rejected Options 2, 3, and 5. The analysis concludes that only two classification outcomes are acceptable: (i) Option 1 — classify in the same category as the underlying intragroup income or expense would have been classified if not eliminated, or (ii) Option 4 — default to the operating category if that cannot be determined without undue cost or effort. Given that Option 1 relies on interpretative reasoning rather than explicit IFRS 18 requirements, this paper argues that, under the standard's current wording, the operating category is likely to become the prevailing classification in practice. The paper highlights the need for further clarification by the standard-setter to avoid future inconsistency across reporting entities and preserve IFRS 18's objective of enhanced transparency and comparability.*

**Keywords:** *IFRS18, Exchange differences classification, Operating, financing and investing category; Profit and loss statement structure; Accounting presentation*

**Jel Classification:** *M41*

### **1. INTRODUCTION**

In April 2024, the International Accounting Standards Board (IASB) issued IFRS 18 *Presentation and Disclosure in Financial Statements*, marking one of the most significant developments in financial reporting since the introduction of IFRS. The new standard replaces IAS 1 and seeks to substantially enhance the structure and clarity of financial communication, with a particular emphasis on improving the presentation and disaggregation of financial performance in the statement of profit or loss. Although IFRS 18 does not change how financial performance is measured, it introduces fundamental changes to how results are presented and disclosed, aiming to provide investors with more comparable, transparent, and decision-useful

information across industries and jurisdictions. The standard becomes effective on 1 January 2027, with early application permitted.

IFRS 18 aims to improve financial reporting, i.e., increasing transparency and comparability of the financial statements by:

- requiring additional defined subtotals in the statement of profit or loss;
- requiring disclosures about management-defined performance measures; and
- adding new principles for grouping (aggregation and disaggregation) of information.

The second and third requests above are not in focus of this paper, so will not be further elaborated.

Adding defined subtotals to the statement of profit or loss makes companies' financial performance easier to compare and provides a consistent starting point for investors' analysis. Items of income and expense shall be classified into categories in the statement of profit or loss:

- operating
- investing
- financing
- income tax, discontinued operations.

The *operating category* is the default category that:

- comprises all income and expenses arising from a company's operations, regardless of whether they are volatile or unusual in some way. Operating profit is not a measure of 'persistent' or 'recurring' operating performance. It provides a complete picture of the results from a company's operations for the period.
- includes, but is not limited to, income and expenses from a company's main business activities. Income and expenses from other business activities, such as income and expenses from additional activities, are also classified in the operating category if those income and expenses do not meet the requirements to be classified in any of the other categories.

The *investing category* comprises income and expenses from:

- investments in associates, joint ventures, and unconsolidated subsidiaries;
- cash and cash equivalents; and
- other assets that generate a return individually and largely independently of the company's other resources.

The *financing category* comprises:

- income and expenses from liabilities arising from transactions that involve only the raising of finance, and
- interest income and expenses, and the effects of changes in interest rates from liabilities arising from transactions that do not involve only the raising of finance.

Among the other classification discussions, one of the most emphasized interpretations is for the classifications of the foreign exchange differences from an intragroup transaction.

The Interpretation Committee received a request about the classification of a foreign exchange difference from an intragroup monetary liability (or asset). Paragraph B65 of IFRS 18 Presentation and Disclosure in Financial Statements requires an entity to 'classify foreign exchange differences included in the statement of profit or loss applying IAS 21 [The Effects

of Changes in Foreign Exchange Rates] in the same category as the income and expenses from the items that gave rise to the foreign exchange differences...' (IASB, 2025).

Multinational enterprises (MNEs) in the telecommunications industry operate at the intersection of heavy capital intensity, complex revenue arrangements, and structurally diverse corporate groups. Typical features include multi-jurisdictional footprints, treasury and financing centres, extensive intercompany service arrangements, and material exposures to foreign exchange (FX) risk stemming from cross-border procurement, spectrum license payments, roaming settlements, and intra-group funding. Against this backdrop, the introduction of IFRS 18 Presentation and Disclosure in Financial Statements reconfigures how performance is portrayed by requiring consistent classification of income and expenses into operating, investing, and financing categories, together with clearer subtotals and disaggregation principles. A pivotal aspect—particularly salient for telecom MNEs with frequent intra-group monetary positions—is how FX differences are classified in profit or loss when they arise from items whose related income and expenses are eliminated on consolidation. IFRS 18 directs entities to align the classification of FX differences with that of the underlying item; when consolidation eliminates the underlying income/expense, practice questions emerge about default versus look-through classification approaches.

This case examines a large, anonymized telecommunications conglomerate with operating subsidiaries across both developed and emerging markets. The conglomerate deploys centralized treasury policies, routinely issues intercompany loans to fund network roll-out and spectrum acquisitions, and settles substantial roaming and infrastructure-sharing payables between group entities. These arrangements are denominated in multiple currencies and frequently retranslated, generating recurring FX differences. Under IFRS 10, intra-group income and expenses are eliminated, whereas IAS 21 requires FX differences on intra-group monetary items to remain in consolidated profit or loss—producing a presentation challenge precisely addressed by IFRS 18's categorization model and its guidance on FX classification. The issue is not merely technical; for an industry where external stakeholders scrutinize operating profit and cash-based performance indicators, the placement of FX effects (operating vs. investing vs. financing) can materially influence performance narratives, KPI targets, loan covenants, and management compensation design.

## **2. LITERATURE REVIEW**

Identifying and measuring FX risk is far from easy. Few CFOs, even among the best, measure more than a small portion of their overall risk. Many companies manage only visible and easily quantifiable risks, such as exposure to foreign currency liabilities. And CFOs seldom comprehend all the risks in their business. They may not understand, say, the relationship between exchange rates and the local prices in the markets where they sell their products (Coppe *et al.*, 1996). The application of the provisions of the IAS 21 standard, if it is necessary to recalculate financial reporting indicators into foreign currencies, will allow Russian companies to generate the most correct information that will be useful to users to assess their financial situation, considering the influence of currency factors. At the same time, the reporting organization should further disclose in its accounting policy the methods for converting financial reporting indicators, including those recommended by IFRS (Kryatova *et al.*, 2022). The objective of IAS 21 is to prescribe the basis for selecting an entity's functional currency and the accounting treatment for the recognition of, and subsequent measurement of, transactions denominated in a foreign currency, and the process of translating financial statements denominated in a foreign currency. The principal issues are which exchange rate(s)

to use and how to report the effects of changes in exchange rates in the financial statements (Muthupandian, 2009). As the interaction of companies across borders increases, accounting for foreign exchange transactions as well as translation of financial statements has become an important topic. An entity is exposed to foreign exchange gains or losses through investments or balances in a foreign currency or ownership in a foreign operation. In order to assess these risks and the related impact, foreign currency transactions and balances should be accounted for accurately. IAS 21 aims to set the framework for the inclusion of foreign currency transactions and foreign operations in the financial statements of an entity and translation of financial statements into a presentation currency (Cayirli, 2020). The International Accounting Standards Board (IASB) is poised to introduce a new standard, IFRS 18 Presentation and Disclosure in Financial Statements, scheduled to come into effect on January 1, 2027. IFRS 18 intends to define operating profit as residual income after deducting the investing and financing categories from the total (Lee, 2024). Moreover, the widespread use of outsourced accountants may also pose challenges in terms of knowledge transfer and training. Companies may need to invest in educating their outsourced accounting teams about the requirements of IFRS 18 and providing ongoing support to ensure accurate and consistent implementation (Neves, 2024).

### **3. RESULTS AND ANALYSIS**

The main question is how an entity, when applying IFRS 18, classifies a foreign exchange difference arising from intragroup monetary liabilities and assets.

Two main transaction types should be analyzed as the most prominent examples in the multinational telecommunication companies:

- a) Intragroup loans to the subsidiaries
- b) In-house cash pooling

The first example arises when a parent entity provides an intragroup loan to its subsidiary and the two entities have different functional currencies. The loan is in the functional currency of either the parent or the subsidiary. On initial recognition, the entity measures the loan using the spot exchange rate at the transaction date. At each subsequent reporting date, the loan is retranslated at the closing exchange rate, with any resulting exchange differences recognized in profit or loss. The payable and receivables will be eliminated in the consolidated Financial Statement and also the interest income and expenses. Nevertheless, the exchange differences arising from those transactions should stay in the PL.

Another prominent example is cash management or in-house cash pooling in multinational telecommunications companies. In-house cash (IHC) is a centralized internal bank arrangement within a multinational group, typically managed by the Group Treasury or a shared service center. It allows subsidiaries to settle intercompany payables and receivables and optimize group liquidity without using external banks for every transaction.

The IHC facilitates Centralized funding of subsidiaries (internal lending), Efficient use of surplus cash, Centralized FX and interest-rate management, and also reduction of external transaction costs.

During consolidation:

- All intercompany receivables/payables and related FX effects are eliminated.

- FX gains/losses arising from intra-group monetary items cancel out, except where balances exist between entities with different functional currencies and the transaction is not fully reciprocal at period-end.

In accordance with paragraph B86(c) of *IFRS 10 Consolidated Financial Statements* the entity should eliminate in full the intragroup loan and any resulting intragroup income and expenses: *“Eliminate in full intragroup assets and liabilities, equity, income, expenses, and cash flows relating to transactions between entities of the group (profits or losses resulting from intragroup transactions that are recognized in assets, such as inventory and fixed assets, are eliminated in full). Intragroup losses may indicate an impairment that requires recognition in the consolidated financial statements. IAS 12 Income Taxes applies to temporary differences that arise from the elimination of profits and losses resulting from intragroup transactions”* (IASB, 2011).

However, in accordance with paragraph 45 of IAS 21 the entity should not eliminate the exchange difference:

*“The incorporation of the results and financial position of a foreign operation with those of the reporting entity follows normal consolidation procedures, such as the elimination of intragroup balances and intragroup transactions of a subsidiary (see IFRS 10 Consolidated Financial Statements). However, an intragroup monetary asset (or liability), whether short-term or long-term, cannot be eliminated against the corresponding intragroup liability (or asset) without showing the results of currency fluctuations in the consolidated financial statements. This is because the monetary item represents a commitment to convert one currency into another and exposes the reporting entity to a gain or loss through currency fluctuations. Accordingly, in the consolidated financial statements of the reporting entity, such an exchange difference is recognised in profit or loss...”* (IASB, 2025).

The main question that arises here is how to apply paragraph B65 of IFRS 18 when classifying, in its consolidated statement of profit or loss, the exchange difference if the income and expenses from the item that gave rise to that exchange difference have been eliminated on consolidation.

Theoretically, five options exist:

**Option 1** - in the same category as the income and expenses from the intragroup loan would have been classified if those income and expenses had not been eliminated on consolidation.

**Option 2** - in the financing category because the exchange difference results from a transaction that, in substance, relates to the raising of (intragroup) finance.

**Option 3** - in the investing category because the exchange difference, in substance, arises due to a transfer of cash from one currency into another currency for a period of time.

**Option 4** - in the operating category. Because the income and expenses from the intragroup loan have been eliminated on consolidation, the parent cannot apply paragraph B65 of IFRS 18 to classify the exchange difference. Consequently, applying paragraph 52 of IFRS 18, the parent classifies the exchange difference (by default) in the operating category.

**Option 5** - by applying any of the first four views consistently following an accounting policy choice because IFRS 18 does not have explicit requirements on how to classify the exchange difference.

The IFRS Interpretation Committee discussed this topic in September 2025 and came to the following conclusion.

Paragraph B65 of IFRS 18 (reproduced in paragraph 9 of this paper) contains requirements for classifying exchange differences in the statement of profit or loss. Based on our analysis, it can be concluded that a reasonable reading of the applicable requirements results in the entity classifying the exchange difference as either:

- (a) in the same category in which it would have classified the income and expenses from the intragroup loan if those income and expenses had not been eliminated on consolidation, or, if doing so would involve undue cost or effort, in the operating category; or
- (b) in the operating category as the default category applying paragraph 52 of IFRS 18 (IC, 2025).

### **Options 2, 3, and 5**

This means the IFRS IC disagrees with Option 2 and Option 3. As mentioned, the exchange difference arises from the intragroup loan itself and not from a cash transfer which means is contrary to Option 3. Furthermore, the exchange difference is recognized by the group entity for which the loan is a foreign currency transaction — whether borrower or lender — and therefore cannot be viewed as arising solely from the raising of intragroup finance, which means is contrary to Option 3.

Accordingly, IFRS IC also disagrees with Option 5, which assumes Options 2 and 3 are acceptable alternatives. Because Options 2 and 3 are not acceptable, Option 5 is also not acceptable.

### **Option 1**

The argument for this option is based on that the elimination of intragroup loans does not mean that the loan does not exist in the consolidated financial statements, but the intragroup loan is eliminated for presentation purposes only.

Although paragraph 7.10 of the Conceptual Framework is in the context of items that are offset in the statement of financial position,

*“Offsetting occurs when an entity recognizes and measures both an asset and a liability as separate units of account, but groups them into a single net amount in the statement of financial position. Offsetting classifies dissimilar items together and therefore is generally not appropriate”* (IASB, 2018)

is reasonable to consider that this concept of offsetting also applies to income and expenses that are offset in the statement of profit or loss. Therefore, IC stated that ...it is reasonable to conclude that eliminating income and expenses arising from an intragroup loan on consolidation does not mean that those income and expenses do not exist; rather, those income and expenses have been offset for presentation purposes...

Applying this approach to the example mentioned above in this paper, if the loan is granted in the functional currency of the parent company, the exchange differences that are recognized by the subsidiary, should be presented in the parents consolidated financial statements as exchange difference in the financing category because any interest expense on the loan (which has been eliminated against the corresponding interest income on consolidation) would have been classified in the financing category.

Vice versa, if the loan is granted from the parent company to the subsidiary in the functional currency of the subsidiary then the exchange differences will be recognized by the parent entity and should be presented in the parents consolidated financial statements as exchange difference

in the investing category because any interest income on the loan (which has been eliminated against the corresponding interest income on consolidation) would have been classified in the investing category.

#### **Option 4**

The parent entity can apply paragraph B65 of IFRS 18 only if there is a ‘category’ within which income and expenses arising from the item that gave rise to the exchange difference, but in the example given above, the income and expenses arising from the intragroup loan have been eliminated on consolidation and are consequently not presented in the consolidated statement of profit or loss. In this situation, the parent entity reasonably could conclude that there is no ‘same’ category within which it can classify the exchange difference and, therefore, it classifies the exchange difference in the operating category as the default category applying paragraph 52 of IFRS 18

#### **4. CONCLUSION**

The classification of foreign exchange differences arising from intragroup monetary items under IFRS 18 requires the application of paragraph B65, even when the related intragroup income and expenses are eliminated on consolidation in accordance with IFRS 10. There are 5 theoretical options of classification and as analyzed above it is reasonable such exchange differences should be classified either (i) in the same category in which the income and expenses from the intragroup loan would have been presented had they not been eliminated (Option 1), or (ii) in the operating category as the default category (Option 4, applying paragraph 52 of IFRS 18).

Options 2 and 3 are not acceptable because the exchange difference arises from the intragroup loan itself — not from the raising of finance or a cash transfer. Consequently, Option 5, which assumes these to be acceptable alternatives, is also not acceptable.

So, at the moment 2 Options are available where reading the IFRS 18 standard and especially the Paragraph B65 of IFRS 18 requires an entity to ‘classify foreign exchange differences included in the statement of profit or loss applying IAS 21 [The Effects of Changes in Foreign Exchange Rates] in the same category as the income and expenses from the items that gave rise to the foreign exchange differences...’ but as the income and expenses from the items that gave rise to the foreign exchange differences does not exist and are eliminated the entity should apply paragraph 52 of IFRS 18 and classify the exchange differences in the operating category as the default category. So, if the standard setters do not amend the IFRS18, it can be concluded that reading the current standard is leading to classification in Operating category because Option 1 is not based on the explicit requirements of IFRS18 but is based just on interpretation of IFRS 10 and Conceptual framework that the elimination of intercompany transactions does not mean the transactions does not exist from Group point of view in the consolidated financial statements.

An additional topic for discussion is what if the standard setters do not come with further clarification. That may result in setting accounting policy by the entities and different companies to have a different approach, which, on the other side, conflicts with the main aim of the standard mentioned above in this paper that IFRS 18 should result in more transparent and comparable consolidated financial statements.

## REFERENCES

- Cayirli, O. (2020). "IAS 21 - The Effects of Changes in Foreign Exchange Rates: A Review of Concepts and Related Issues", available at: [papers.ssrn.com/sol3/Delivery.cfm/SSRN\\_ID3519984\\_code1331214.pdf?abstractid=3519984&mirid=1&type=2](https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3519984_code1331214.pdf?abstractid=3519984&mirid=1&type=2) (accessed 12 September 2025).
- Coppe, B., Graham, M., and Koller, T. M. (1996). "Are you taking the wrong FX risk?", *The McKinsey Quarterly*, Vol. 1 No. 80, available at: <https://link.gale.com/apps/doc/A55503386/AONE?u=anon~86422dc7&sid=googleScholar&xid=da36bba3> (accessed 3 September 2025).
- IASB. (2011). IFRS10 Consolidated financial statements, available at: <https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards/english/2021/issued/part-a/ifrs-10-consolidated-financial-statements.pdf> (accessed 7 September 2025)
- IASB. (2018). Conceptual Framework, available at: <https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards/english/2021/issued/part-a/conceptual-framework-for-financial-reporting.pdf> (accessed 7 September 2025)
- IASB. (2025). IAS 21 The effects of Change in Foreign Exchange Rates. IFRS.org, available at <https://www.ifrs.org/> (accessed 7 September 2025)
- IC, I. (2025). Staff paper on FX differences.
- Kryatova, L.A., Nurgaliev, R.N., Amirova, R.I., Gamilovskaya, A.A., Ivanova, E.V. (2022). Features of the Application of IAS 21 "Impact of Exchange Rate Changes". In: Bogoviz, A.V., Suglobov, A.E., Maloletko, A.N., Kaurova, O.V. (eds) Cooperation and Sustainable Development. Lecture Notes in Networks and Systems, Vol 245. Springer, Cham.
- Lee, S. (2024). "Establishment of IFRS 18 Presentation and Disclosure in Financial Statements: its Impact and Implications. Capital Market Focus: Korean Capital Market Institute", available at: [https://www.kcmi.re.kr/kcmifile/webzine\\_content/OPINION/6305/webzinepdf\\_6305.pdf](https://www.kcmi.re.kr/kcmifile/webzine_content/OPINION/6305/webzinepdf_6305.pdf) (accessed 13 September 2025).
- Muthupandian. (2009). "IAS 21 The Effects of Changes in Foreign Exchange Rates - A Closer Look", *The Management Accountant*, Vol. 44 No. 10.
- Neves, H. D. (2024). "IFRS 18 Implementation in Brazilian Enterprises: Challenges and Opportunities", *International Journal of Business Administration*, Vol. 15, No. 2, pp. 25669-25669.

## **COOPERATIVE VS. NON-COOPERATIVE GAMES: A COMPARISON OF SOLUTIONS AND IMPLICATIONS**

**Stefan Mladenoski**

*Faculty of Economics, American University of Europe-FON*

*North Macedonia*

*mladenoski.stefan@outlook.com*

### **ABSTRACT**

*Game theory is considered a key framework in understanding the market economy. That is exactly why the comparison between cooperative and non-cooperative games and their outcomes is significant. Namely, when it comes to what constitutes a cooperative game, as the name suggests, it is a game in which two or more players make decisions together. In contrast, a non-cooperative game is one in which each player independently decides how to play their move. The aim of this paper is to compare the solutions and implications of cooperative and non-cooperative games, which will be explained through theory and practical examples. The methodological approach used in this paper is based on theoretical analysis, modeling of games, and examining the solutions obtained from the structure of the games. Cooperative and non-cooperative games both involve strategic decision-making by rational players, but they differ fundamentally in the role of communication and the possibility of binding agreements. While cooperative games allow for coalition formation and the achievement of efficient outcomes, non-cooperative games focus on individual strategies in conditions of limited trust, which often lead to less optimal but more realistic equilibria. The results of this comparison highlight the importance of selecting the appropriate game model when analyzing strategic situations, depending on the level of trust, the potential for cooperation, and the nature of interests among players. This opens space for further research toward the development of hybrid models and their practical application in areas such as economic markets or international relations.*

**Keywords:** *Cooperative games, Non-cooperative games, Nash equilibrium*

**JEL classification:** *C71, C72, D43*

### **1. INTRODUCTION**

When defining game theory, it is important to note that it represents an analytical tool for explaining the phenomenon of strategic decision-making, that is, situations in which the decisions of one player directly depend on the expected or actual choices of others. The models used in game theory are mostly abstract. Nevertheless, this abstraction becomes an advantage, as it allows for a clear representation of reality and applicability across various situations in economics. Furthermore, the line between theoretical and applied game theory is often blurred, with many formal models serving as the foundation for the development of practical scenarios. Mathematics is frequently used as a tool in game theory to facilitate this adaptation, to precisely define concepts, evaluate the consistency of ideas, and explore the implications of assumptions (Osborne, 2004).

Historically, game theory began to develop after World War II, with the foundational work of John von Neumann and Oskar Morgenstern (1944, as cited in Osborne, 2004). In addition, it was further advanced by John Nash, who introduced the concept of equilibrium in non-

cooperative games (Nash, 1950). Nash equilibrium has become one of the most influential concepts in economics and the social sciences, enabling the modeling of strategic interactions without requiring direct coordination or communication among players. On the other hand, the development of cooperative game theory, especially through concepts such as the Shapley value (Shapley, 1953) – focuses on the fair distribution of gains among coalition members, considering their relative contributions.

Thus, game theory is divided into two main branches: cooperative and non-cooperative. While both originate from a common logic of rational behavior, they differ in terms of communication, the possibility of binding agreements, and the structure of interaction (Osborne, 2004). This distinction is highly relevant for the analysis of strategic situations in the real world, as it reflects the difference between situations in which players can communicate and make joint decisions and those in which they are forced to act independently, without influencing the choices of others (Leyton-Brown & Shoham, 2008).

Cooperative games are those in which players have the chance to form coalitions and agree on collective strategies that maximize joint payoff. In this context, players are motivated to coordinate and negotiate to achieve stable and optimal outcomes. For example, in economics, cooperative games are often used to model scenarios such as labor unions, cartels, or other forms of collaboration within industries (Tadelis, 2013), where players (e.g., companies) can agree on prices, output levels, or resource allocation.

In contrast, non-cooperative games refer to situations where players lack the ability or incentive to form coalitions and make joint decisions. Here, each player focuses solely on individual payoff and optimizes their strategy in a competitive environment. The outcomes of non-cooperative games—such as the Nash equilibrium—represent situations where each player chooses a strategy that maximizes their own payoff, given the strategies of the others. In this state, no player has an incentive to unilaterally change their strategy. However, these outcomes are often stable but not necessarily cooperative, especially in environments where communication and trust are limited (Grabisch, 2021).

Understanding this distinction is fundamental when choosing the most appropriate model for analyzing a specific situation, as each type of game is based on different assumptions, constraints, and outcomes. Through the analysis of cooperative and non-cooperative games, one can gain deeper insight into how decision-making varies depending on collaboration, competition, and the strategies of players.

The primary objective of this paper is to compare cooperative and non-cooperative games, focusing on their solutions, assumptions, and practical implications. By analyzing selected models and practical examples, the paper aims to highlight the key similarities and differences between the two approaches, as well as to identify the conditions under which one is more suitable than the other. From a methodological perspective, the research relies on theoretical analysis and formal modeling, employing fundamental game theory concepts such as Nash equilibrium and the Shapley value. (Nash, 1950; Shapley, 1953).

Additionally, this paper seeks to examine concrete scenarios relevant to economic and social contexts, to demonstrate the real-world implications of choosing one type of game over the other. Structurally, the paper is divided into seven sections. The first section introduces the basic theoretical concepts and definitions related to cooperative and non-cooperative game theory. The second section presents a comparative analysis through a specific theoretical

explanation of both cooperative and non-cooperative games. This is followed by the next two chapters, in which a broader practical comparison of both types of games is provided, aimed at providing a comprehensive understanding of the importance of choosing the appropriate game structure. The fifth section discusses side-by-side comparison and practical implications while at the same time suggesting directions for potential hybrid approaches. The paper concludes with a summary and proposals for further research.

## **2. THEORETICAL FRAMEWORK**

### **Key concepts**

#### *Games with two players*

Games with two players are basic analytical models when we talk about game theory. They serve as for deeper understanding of strategic interaction between both parties (Osborne, 2004). Additionally, they define theoretical borders for formalizing behavior in situations with confrontation or potential cooperation. This type of game is classified as cooperative or non-cooperative, depending on whether the players can enter binding contracts (Tadelis, 2013). In a cooperative context, the involved parties decide together for maximization of the common outcome, while in the non-cooperative context, each actor independently chooses a strategy in the direction of optimizing individual gain (Osborne, 2004; Leyton-Brown & Shoham, 2008).

#### *N-players games*

Games involving more than two players introduce a higher degree of complexity due to the exponential growth in possible interactions, coalitions, and strategic configurations (Tadelis, 2013). These games are especially relevant for analyzing collective decision-making, coalition formation, and the fair allocation of outcomes. In cooperative game theory, analytical tools such as the Shapley value, the core, and various coalition stability concepts are employed to evaluate each player's contribution and determine equitable distributions (Shapley, 1953). The Shapley value assigns a value to each player based on their marginal contribution to every possible coalition they could be part of, reflecting how much value a player adds on average to different combinations of players. In the non-cooperative domain, strategic behavior is typically modeled through generalizations of Nash equilibrium in multi-agent settings, where each participant seeks to optimize their own utility while accounting for the strategies of others (Nash, 1950; Osborne, 2004). As the number of players increases, concerns such as stability, trust, and coordination become increasingly important (Tadelis, 2013).

#### *Games with a variable number of players*

In certain contexts, the number of participants is not fixed; players may enter or exit the game or coalition at different stages (Osborne, 2004). These variable-player games demand a distinct theoretical approach, as they introduce dynamic and often unpredictable coalition structures. Such games require consideration of entry and exit costs, dynamic stability, resource reallocation, and coalition flexibility (Myerson, 1991). The analysis of these environments often involves extended concepts such as the dynamic core, time-dependent Shapley value, and adaptive strategic frameworks. The underlying challenge is to ensure fair and stable outcomes even as the composition of participants changes over time.

#### *Strategies*

A strategy is defined as a complete plan of action that a player follows throughout the game (Osborne, 2004). Depending on the nature of the game, the strategy may be either simple or complex, specifying the player's action in every possible contingency. Strategies are often categorized as either pure (deterministic), where a player selects a single specific action, or

mixed (probabilistic), where choices are made according to a probability distribution over available actions (Tadelis, 2013).

### *Payoffs*

The payoff denotes the value a player receives because of the strategies chosen by all participants (Myerson, 1991). It typically reflects the player's utility, profit, or another quantifiable benefit, and serves as the core measure of individual rationality and motivation. Payoff functions underpin the evaluation of strategic behavior and are essential to determining optimal responses and outcomes in any game-theoretic context (Tadelis, 2013).

### *Equilibrium*

The concept of equilibrium is central to game-theoretic analysis. It identifies outcome configurations in which no player has an incentive to unilaterally deviate from their chosen strategy (Osborne, 2004). The most prominent notion is the Nash equilibrium, defined as a set of strategies - one for each player - such that no player can improve their payoff by independently changing their strategy, given the strategies of the others (Nash, 1950). Equilibria may arise in pure or mixed strategies and serve as a predictive and normative tool for analyzing stability and rational outcomes in strategic environments (Tadelis, 2013).

## **3. COOPERATIVE GAMES**

### **Definition**

Cooperative games are a form of strategic interaction in which the players have opportunities to conclude binding agreements and to form coalitions. In this context, instead of analyzing the individual strategies of each participant, the focus is placed on groups of players (coalitions) and their ability to achieve a joint profit (Osborne, 2004; Myerson, 1991). The analysis assumes that the agreements between the players are fully binding and enforceable (Tadelis, 2013; Grabisch, 2021).

### **Characteristics**

Cooperative games are characterized by several essential features:

- Formation of coalitions: Players can group together and form coalitions, which allows for collective action and maximization of common interests.
- Binding agreements: It is assumed that once agreements are reached, they are fully enforced, which makes the model suitable for analyzing stable long-term coalitions (Osborne & Rubinstein, 2016).
- Characteristic function: For each possible coalition, a value is defined that represents the potential gain that coalition can achieve on its own, independent of the rest of the system (Myerson, 1991).
- Distribution of gains: A central issue in cooperative theory is how the gains from cooperation are distributed among the members of the coalition. To that end, various concepts are used, such as core, Shapley value, nucleolus, and bargaining set (Shapley, 1953; Tadelis, 2013).

### **Example**

As an example, for the cooperation game, the author's interpretation of the glove game is presented "the glove game" (Osborne & Rubinstein, 2016).

Let us consider a game in which there are 3 players. Players 1 and 2 possess a right glove, and player 3 a left glove. Possessing a pair of gloves is worth 1. The players must cooperate to generate profit. Thus, we have

$$v(S) = \begin{cases} 1 & \text{if } S \text{ includes at least one right – glove player (1 or 2) and player 3} \\ 0 & \text{otherwise} \end{cases}$$

Table 1: The values of  $v(S)$  for the glove game

S	{1}	{2}	{3}	{1,2}	{1,3}	{2,3}	N
v(S)	0	0	0	0	1	1	1

(Source: Made by the author)

Now we compute the Shapley value for the glove game presented above. In theory, the Shapley value is a vector in  $\mathbb{R}^N$  given by

$$\Phi(N, v) = \frac{1}{n!} \sum_{\sigma \in G(N)} m^\sigma$$

In this formula, it is important to explain the key symbols where:

In the tables above, the following symbols are used:

- $N$  refers to the set of all players participating in the game.
- $G(N)$  denotes the set of all possible permutations, or different possible orders, in which the players can form a coalition.
- $\sigma$  (*sigma*) represents one specific permutation or order of players from the set  $G(N)$ .
- $v$  stands for the characteristic function, which assigns a value or payoff to each possible coalition of players.
- $m^\sigma$  indicates the marginal contribution of the players according to the specific permutation  $\sigma$ , showing how much value each player adds when joining the coalition in that order.

Then, for our example, we obtain:

Table 2: The values of  $v(S)$  for the glove game

order $\sigma$	$m_1^\sigma$	$m_2^\sigma$	$m_3^\sigma$
{1, 2, 3}	0	0	1
{1, 3, 2}	0	0	1
{2, 1, 3}	0	0	1
{2, 3, 1}	0	0	1
{3, 1, 2}	1	0	0
{3, 2, 1}	0	1	0
$\Phi(v)$	1/6	1/6	4/6 = 2/3

(Source: Created by the author, theoretical basis in Shapley, 1953)

Note that in the glove game, the core is the point  $\{0,0,1\}$ . Hence, in general, the Shapley value does not lie in the core.

#### 4. NON-COOPERATIVE GAMES

##### Definition

A non-cooperative game is a formal representation of strategic interaction where players make individual decisions without the possibility of forming binding agreements or coalitions (Fudenberg & Tirole, 1991). Each player acts independently, seeking to maximize their own payoff, considering the actions of others. The focus lies on predicting individual behavior in situations where outcomes depend not only on one's own decisions but also on the strategies chosen by others. In contrast to cooperative games, non-cooperative game theory analyzes outcomes that arise purely from strategic choices, without external enforcement of cooperation (Myerson, 1991).

##### Characteristics

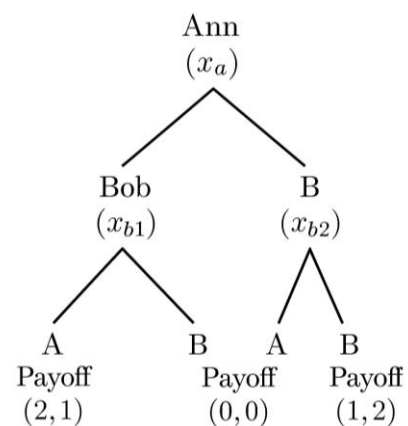
Non-cooperative games possess several defining features:

- Individual Decision-Making: Each player selects their strategy independently, aiming to maximize their own profitability (Binmore, 2007).
- Strategic Interdependence: A player's payoff is influenced not only by their own decision but also by the decisions of all other players (Tadelis, 2013).
- Game Representation: These games are typically represented in normal (strategic) form, through payoff matrices, or in extensive form, which models sequential moves via game trees. (see pictures below)
- No Binding Agreements: Players cannot make enforceable contracts or promises; coordination, if it exists, must emerge endogenously from equilibrium behavior.
- Equilibrium Concept: The central solution concept is the Nash Equilibrium, in which no player can gain by unilaterally changing their strategy given the strategies of others (Nash, 1950).

Picture 1: Game matrix representation

$A \setminus B$	A	B
A	2,1	0,0
B	0,0	1,2

Picture 2: Game tree representation



(Source: Meeting game (Fujiwara-Greve, 2015)

(Source: Tree for the sequential decision meeting game (Fujiwara-Greve, 2015)

**Example**

An example of this type of game is the classical prisoner’s dilemma. There is one representation of the prisoner’s dilemma.

Namely two guys have been arrested by the police. Since there is not enough evidence to convict them, the police interrogator offered the following deal the following deal to each of them:

If you testify that the other guy is guilty (defect), you will be set free and the other will be imprisoned for 3 years; unless he also defects, in which case you both will be imprisoned for 2 years. If you both keep your mouth shut (cooperate), each of you will serve 1 year in prison.

Each prisoner knows that the same offer has been made to the other. Moreover, each one knows that the other knows that he knows, and so forth. This assumption about the game is called common knowledge and is a necessary prerequisite to have a game of complete information. Without common knowledge, the behavior of prisoners could be different.

The situation is formalized as a strategic game  $G = (N; A; u)$ , where the two players in  $N = \{p_1; p_2\}$  can play actions from  $A = \{C, D\}$ , where C stands for Cooperate with the other player (don't speak to the police) and D stands for Defect (make a deal with the police to testify against the other prisoner). After playing the players will get payoffs denoted by the utility function  $u = (u_1, u_2)$ , where  $u_i: A \mapsto \mathfrak{R}$  is depicted in Table 3. Since players are awarded years in prison (they get punishment rather than reward), the payoffs take a negative sign, to denote that more years are worse (non-preferable) to less years.

Table 3: The values of prisoner’s dilemma game

$u_1, u_2$		$p_2$ player	
		Cooperate	Defect
$p_1$ player	Cooperate	(-1, -1)	(-3, 0)
	Defect	(0, -3)	(-2, -2)

(Source: Created by the author)

The codomain of  $u_i$  is the set of all potential payoff values that can result for player i from the different combinations of strategies chosen by all players in the game.

**5. COMPARATIVE ANALYSIS: COOPERATIVE VS NON-COOPERATIVE GAMES**

**Overview of Main Similarities and Differences**

Cooperative and non-cooperative game theory represent two major strands within the broader field of game theory, both offering analytical tools for understanding strategic interactions

among rational players (Osborne & Rubinstein, 2016; Myerson 1991). While they share a foundational reliance on the rationality of agents and formal mathematical models to describe and predict behavior, they differ significantly in their treatment of communication, enforceability, and the structure of interactions.

In both frameworks, players are assumed to act strategically with the aim of maximizing their individual payoffs. Strategic interdependence is central, as outcomes depend not only on a player's own decisions but also on the choices made by others. Moreover, each framework seeks to identify stable outcomes through rigorous solution concepts such as the Nash equilibrium in non-cooperative games, or the core and the Shapley value in cooperative games (Nash, 1950; Shapley, 1953).

The primary distinction lies in the extent to which cooperation is permitted or enforceable. Cooperative games allow for the formation of binding agreements between players who can join forces and coordinate their actions to maximize collective benefits or in other words, to cooperate with each other. These games are typically represented in characteristic function form, which assigns a value to every possible coalition (cooperation) of players. Non-cooperative games, in contrast, assume that players act independently and cannot form enforceable coalitions. Interactions are modeled using strategic (or normal) form representations, in which each player's strategy is set, and corresponding payoffs are explicitly defined (Osborne & Rubinstein, 2016).

Another key difference concerns the treatment of trust, communication, and information. In cooperative games, players are assumed to be able to communicate freely, share information, and form binding agreements that ensure commitment to joint strategies. Trust is externally sustained through institutions or contracts, which legitimizes long-term cooperation and stability. On the other hand, in non-cooperative games, the lack of enforceable commitments implies that communication may be limited or strategically manipulated, and information asymmetries play a more central role. Players cannot rely on trust enforced from outside the game, and instead, must base their actions on the rational expectations of others' behavior. This makes the study of incentives, beliefs, and strategic uncertainty essential within non-cooperative models.

### **Trust, communication, and information problems**

While the general contrasts between cooperative and non-cooperative games have been outlined, the dynamics of trust, communication, and information merit closer attention due to their central role in shaping strategic environments. In these terms, cooperative games explicitly assume that players can communicate, share private information, and enforce agreements - often through institutional or contractual mechanisms. These assumptions enable coordination, reduce uncertainty, and promote stable, mutually beneficial outcomes.

In contrast, non-cooperative games, however, model environments where such mechanisms are absent, unreliable, or strategically constrained. Players must operate under conditions of incomplete or asymmetric information, and communication - if it occurs - is often limited or used manipulatively. Trust is not externally enforced but must instead be incentivized internally, emerging (if at all) from the structure of the game itself. This endogenous development of trust and the absence of guarantees underscore the need for precise mechanism design, where systems must be crafted to align individual incentives with socially desirable outcomes despite informational and strategic barriers. As a result, the study of non-cooperative games places particular emphasis on the role of beliefs, signaling, and equilibrium behavior under uncertainty.

### **Implications for Real-World Applications**

Understanding the distinctions between cooperative and non-cooperative games carries significant implications for decision-making in real-world domains such as economics, politics, and negotiations. The applicability of one framework over the other often depends on the institutional context, the nature of interaction among agents (players), and the availability of enforcement mechanisms.

In economic contexts, cooperative game theory is especially relevant in situations where binding agreements can be established (Myerson, 1991). Examples include labor unions negotiating with employers, joint ventures between firms, or the formation of cartels where competitors coordinate prices or output. These settings allow for coalition formation and collective bargaining, making cooperative models suitable for analyzing stability, profit distribution, and incentives to remain within the coalition. In contrast, non-cooperative models are fundamental for understanding competitive markets where firms act independently, such as in pricing strategies, auctions, or oligopolistic rivalry. Here, the lack of enforceable contracts and the strategic behavior of agents necessitate a non-cooperative approach.

In political science, cooperative games are used to analyze coalition governments, party alliances, and international treaties, all of which rely on negotiated agreements and shared commitments. These situations typically involve explicit coordination among actors who align their interests toward common goals. Conversely, many aspects of political behavior, such as voting dynamics, strategic lobbying, or legislative bargaining, are better captured through non-cooperative frameworks. In scenarios like these, individual actors pursue their objectives independently, often within institutions that do not guarantee binding outcomes.

Negotiation scenarios also vary in terms of their cooperative or non-cooperative nature. When parties can make enforceable commitments - such as in trade agreements, diplomatic accords, or environmental treaties - a cooperative model is appropriate. Nonetheless, many real-world negotiations take place without enforceable commitments and in settings characterized by asymmetric information and low levels of trust. These are more accurately modeled as non-cooperative games, where outcomes depend on the strategic choices of each party and their expectations about others' behavior.

These distinctions help explain the growing interest in hybrid approaches that incorporate both cooperative and non-cooperative elements. In many strategic environments, actors may have partial communication, informal trust, or weak enforcement mechanisms. Hybrid models offer a more nuanced framework, reflecting the layered realities of trust, information, and institutional structure that shape outcomes across economic and political domains.

## **6. CONCLUSION**

The theoretical distinction between cooperative and non-cooperative games lies at the core of game theory and reflects two fundamentally different approaches to modeling strategic interaction (Myerson, 1991; Osborne & Rubinstein, 2016). Throughout this study, we have examined how each framework conceptualizes the behavior of rational agents, the structure of their interactions, and the mechanisms through which outcomes are achieved. Cooperative games rely on the assumption that players can negotiate, form binding agreements, and share information transparently. This allows for joint strategy formation and the possibility of fair profit distribution through well-established concepts such as the Shapley value (Shapley,

1953). Non-cooperative games, by comparison, depict environments where players act on their own, frequently under conditions of limited or uneven information, and without enforceable agreements (Nash, 1950). This creates a fundamentally different environment in which equilibrium concepts such as the Nash equilibrium become central for analyzing behavior.

By providing a detailed comparison of cooperative and non-cooperative games, the study highlights that each approach offers unique strengths and is applicable under different sets of assumptions and institutional conditions (Myerson, 1991). Cooperative games are particularly well suited for analyzing long-term partnerships, labor unions, cartel formations, and other scenarios where sustained collaboration and enforceable commitments are possible. Non-cooperative games, on the other hand, excel in modeling competitive markets, political bargaining, and negotiation settings where actors pursue their self-interest without the possibility of forming binding coalitions.

Furthermore, this paper has shown that the presence or absence of trust, communication, and institutional enforcement critically influences which game-theoretic framework is appropriate for a given situation. In cooperative settings, trust and communication are external assumptions, supported by contracts or shared norms. While in non-cooperative games, trust must emerge endogenously from the structure of incentives, and communication is often strategic and non-binding. This distinction has significant implications for fields such as economics, where the design of institutions and market rules can be seen as efforts to shift environments from non-cooperative to cooperative dynamics (Myerson, 1991).

In addition, the growing complexity of real-world interactions has given rise to hybrid models that incorporate elements of both cooperative and non-cooperative game theory. These approaches recognize that in many settings, actors may partially cooperate while still retaining individual agency, or that communication and trust may exist in some dimensions but not in others. As a result, hybrid modeling frameworks allow for a more realistic representation of strategic environments in domains such as international negotiations, joint ventures, political alliances, and platform-based digital markets.

The implications of this research extend beyond theoretical interest. Understanding when cooperation is possible or likely, how incentives can be aligned in competitive environments, and what institutional mechanisms can foster stability and fairness is essential for policymakers, business leaders, and negotiators alike (Myerson, 1991). Whether analyzing mergers, labor disputes, environmental agreements, or legislative bargaining, the tools of game theory provide powerful insights into the dynamics of human interaction.

In conclusion, cooperative and non-cooperative game theory should not be viewed as competing paradigms but rather as complementary tools within a broader analytical framework. Their combined application offers a richer and more flexible approach to modeling complex strategic behavior, guiding decision-making processes in a variety of social, economic, and political contexts. Future research in game theory should continue to explore the boundaries between these models and refine hybrid approaches that better reflect the realities of strategic interdependence in an increasingly interconnected world (Myerson, 1991).

## **REFERENCES**

Binmore, K. (2007), *Playing for Real: A Text on Game Theory*, Oxford: Oxford University Press.

- Fudenberg, D. and Tirole, J. (1991), *Game Theory*, Cambridge: MIT Press.
- Fujiwara-Greve, T. (2015), *Non-Cooperative Game Theory*, Tokyo: Springer.
- Grabisch, M. (2021), *Cooperative Game Theory*, Paris: Paris School of Economics.
- Leyton-Brown, K. and Shoham, Y. (2008), *Essentials of Game Theory*, Morgan & Claypool Publ..
- Myerson, R. (1991), *Game Theory: Analysis of Conflict*, Cambridge, MA: Harvard University Press.
- Nash, J. (1950), *Equilibrium points in n-person games*, Proceedings of the National Academy of Sciences.
- Osborne, M. J. (2004), *Introduction to Game Theory*, Oxford: Oxford University Press.
- Osborne, M. J. and Rubinstein, A. (2016), *A course of Game Theory*, Cambridge: The MIT Press.
- Shapley, L. S. (1953), "A value for n-person games", In: Contributions to the Theory of Games II, Princeton University Press, pp. 307-317.
- Tadelis, S. (2013), *Game Theory: An Introduction*, Princeton University Press.