EKONOMSKI VJESNIK

ECONVIEWS

Review of contemporary business,
entrepreneurship and economic issues

EFOS

Josip Juraj Strossmayer
University of Osijek

**Faculty of Economics
and Business in Osijek**

EKONOMSKI VJESNIK

ECONVIEWS

Review of contemporary business,
entrepreneurship and economic issues

EKONOMSKI VJESNIK / ECONVIEWS

Review of contemporary business, entrepreneurship and economic issues

English Language Editor
Ivanka Ferčec

Indexed in:
- Emerging Sources Citation Index, Web of Science, Clarivate Analytics, New York, USA
- CAB Abstract, Wallingford, United Kingdom
- Econlit, Pittsburg, USA
- EBSCOhost, Ipswich, USA
- ABI/INFORM, ProQuest, London, United Kingdom
- DOAJ - Directory of Open Access Journals, United Kingdom
- CEEOL - Central and Eastern European Online Library, Frankfurt am Main, Germany
- Index Copernicus, Warszawa, Poland
- Hrčak – Portal znanstvenih časopisa Republike Hrvatske (MZOS, Srce & HIDD)

Graphic Design
Grafika, Osijek

Published biannually

Address of the Editorial Board
Josip Juraj Strossmayer University of Osijek,
Faculty of Economics and Business in Osijek
Trg Ljudevita Gaja 7
31000 Osijek
Croatia

hrcak.srce.hr/ojs/index.php/ekonomski-vjesnik/

https://www.efos.unios.hr/istrazivacki-rad/casopisi/ekonomski-vjesnik/

ekonomski.vjesnik@efos.hr

CONTENTS

Review paper

Conference review

Original scientific articles

Murat Baş
Erzincan Binali Yıldırım University
Faculty of Economics
and Administrative Sciences
24050 Erzincan, Türkiye
murat.bas@erzincan.edu.tr

Moh. Qasim Ayaz
Salam University
Management Sciences Department
1007 Kabul, Afghanistan
Qasim.ayaz@salam.edu.af

Kaya Ağin
Erzincan Binali Yildirim University
Vocational School of
Higher Education
24050 Erzincan, Türkiye
kagin@erzincan.edu.tr

THE MEDIATING ROLE OF KNOWLEDGE SHARING IN THE RELATIONSHIP BETWEEN LEADER-MEMBER EXCHANGE AND WORK INTEGRATION

Abstract

Purpose: The purpose of this study is to explain the mediating effect of knowledge sharing (KS) between leader-member exchange (LMX) and work integration (WI) on healthcare workers. The study also examines the direct effects of LMX on WI among selected healthcare workers in Türkiye.

Methodology: The study targeted a sample of N = 295 to provide relevant information. The data were collected using an adopted questionnaire, including LMX (Baş et al., 2010), KS (Chennamaneni et al., 2012), and WI (Schaufeli et al., 2002). A purposive stratified sampling method was used since the participants were considered based on predefined criteria. The study used a 5-point Likert scale ranging from 1= strongly disagree to 5= strongly agree. SPSS 22.0 and AMOS 17.0 packages were used for data analysis and interpretation of statistical outputs.

Results: The study revealed a significant positive relationship between variables. Goodness of fit values of the scales were also checked. First, the mediating variable (knowledge sharing) between LMX and WI was investigated. Second, the effect of LMX on WI was explored. Based on the findings, a significant positive relationship was found between the variables of the study.

Conclusion: According to the research results, significant relationships were found between LMX and WI, between LMX and KS, and between KS and WI. In addition, it was determined that knowledge sharing had a mediating role in the effect of LMX on WI, which was the main purpose of the research. According to the research results, high-quality LMX relationships enable employees to be more open to KS and to integrate more into their work. In this case, KS acts as a bridge that strengthens the link between LMX and WI. Employees find greater meaning in their work by sharing their knowledge, which in turn increases their commitment to it. As a result, leaders creating a culture that supports knowledge sharing can be seen as an effective way to increase work integration and improve organizational performance.

Keywords: Leader-member exchange, knowledge sharing, work integration, healthcare professionals

1. Introduction

Today, the number of studies on work integration and leader-member exchange is increasing. Theorists and organizational practitioners are highly motivated to find better ways to improve individual performance, enhance work integration, and increase job satisfaction (Podsakoff et al., 2007; Chan & Mak, 2012; Örtqvist & Wincent, 2006). Based on the studies in the literature, it was found that employee attitudes toward innovative participation are mainly triggered by their sense due to the relationships that different teams have within the working environment, departments, or with their leaders (Tierney et al., 1999). It was also concluded that LMX has an impact on interpersonal relationships by fostering constructive collaboration among employees to achieve greater outcomes (Bae et al., 1997; Gerstner & Day, 1997). It was shown in our study that LMX has the potential to reduce and limit or even improve employees' work integration. Furthermore, it causes a wide range of strategic and operational outcomes, and organizational short and long-term goals can be met on time (Adil & Awais, 2016).

It is believed that human exchanges and communication are the primary sources of knowledge sharing; knowledge sharing can be characterized as the interpersonal relation and interaction level that facilitates transmitting and receiving information from one employee to another (Argote & Ingram, 2000). According to Widen-Wulff and Suomi (2007), knowledge sharing occurs in individual minds and can only exist if individual minds agree to transform what they know. Du et al. (2007) believe in knowledge sharing via deliberate and resource-intensive efforts. Nooteboom (1999) suggests three different dimensions of knowledge: (1) breadth, (2) depth, and (3) tastiness.

Based on the above argument, it is perceived and identified that studies on the impact of LMX on business integration have not fully addressed the issue. Michel and Tews (2016) found that the investigation of LMX requires further elaboration. The majority of previous studies, for instance Hong et al., 2004; Kenney & Gudergan, 2006; Zahra et al., 2020, focused on either organizational integration or knowledge integration (knowledge sharing). However, many studies related to incorporation in the literature do not adequately address the complexities of today's globalized business environment. In the current study, we propose an integrated work frame that involves leader-member exchange, knowledge sharing, and work integration with experiential statistics from healthcare organizations, aiming to develop a more comprehensive viewpoint of the relationship between the selected scales. In line with this purpose, the study seeks answers to the following questions.

RQ1: To what extent does leader-member exchange affect work integration among healthcare workers?

RQ2: What is the relation between leader-member exchange and knowledge sharing?

RQ3: To what extent does knowledge sharing affect work integration among healthcare workers in Türkiye?

RQ4: Does knowledge sharing mediate the relation between leader-member exchange and work integration?

According to previous studies, the leader-member exchange affects employee performance and strengthens the relation between employees and leaders. This is proven by several experts, for example, according to Rugian et al. (2017), employee performance significantly contributes to leader-member exchange. Further, according to Arsintadiani and Harsono (2002), LMX positively affects employee work integration. Meanwhile, Sa'adah et al. (2022) suggest further investigation into LMX and work integration, particularly regarding service organizations. To support these arguments, we propose the following specific objectives for this study:

RO1: Examine the influence of leader-member exchange on the work integration of healthcare workers.

RO2: Investigate the relationship between leader-member exchange and knowledge sharing among healthcare workers in Türkiye.

RO3: Explore the effect of knowledge sharing on work integration among healthcare workers.

RO4: Understand the mediating effect of knowledge sharing between leader-member exchange and work integration.

2. Literature review and hypothesis development

2.1 Leader-member exchange and work integration

The concept of leadership theory has become a significant focus of scientific investigation. It has established many experimental research considerations in managerial and organizational contexts. LMX theory was revolutionary and for two main

reasons. First, leader-member exchange, or LMX, focuses on the unique dyadic relationship that exists between leaders and their followers. Second, leaders do not form the same type of affiliation or connection with each of their followers (Gerstner & Day, 1997). According to LMX theory, leaders vary their interactions with individuals and, as a result, start different relationships with their followers. Two issued meta-analyses on LMX supported various relations between LMX and a variety of attitudinal and behavioral outcomes. The initial studies on LMX theory were predominantly focused on individual relationships and behavioral outcomes within organizations (Gerstner & Day, 1997; Ilies et al., 2007). Workers looking for new opportunities and striving to advance their workplace are required since today's global businesses are constantly changing (Oldham & Cummings, 1996; Rank et al., 2004; Unsworth, 2001). There are various studies on leadership; Tierney et al., 1999; Zhou and George, 2003 explained the ideas of leadership, Jaussi and Dionne, 2003; Jung et al., 2003; Sosik et al., 1998 examined the character of transformational leadership, Wang and Noe, 2010 focused on motivational leadership, and Zhang and Bartol, 2010 studied how to empower leadership. Other researchers have begun to investigate the link between a relational concept of leadership, namely LMX, and innovation (Atwater & Carmeli, 2009; Scott & Bruce, 1994; Tierney et al., 1999). LMX theory is distinguished from other leadership approaches by its explicit emphasis on special, dyadic relationships and the notion that leaders and followers negotiate their relationship over time (Dansereau et al., 1973; Graen & Schiemann, 1978). Therefore, our study suggests the first hypothesis as follows:

H1*: Leader-member exchange has a positive relationship with work integration.*

2.2 Leader-member exchange and knowledge sharing

In today's knowledge-based economy, knowledge is the most important foundation for organizational existence (Nonaka & Takeuchi, 1995; Nonaka et al., 1994). Individual knowledge sharing is one of the most important practices for achieving organizational effectiveness since it is the basis for creativity and innovative knowledge creation (Nonaka & Takeuchi, 1995; Nonaka et al., 1994; Quigley et al., 2007). However, encouraging individuals to do so in a competitive workplace is difficult due to the unique characteristics of knowledge sharing. For

example, each individual's awareness is considered a valued resource (French & Raven, 1959; Jang et al., 2002; Cabrera & Cabrera, 2005). Furthermore, shared knowledge is accessible to all, and individuals can exacerbate free rider issues (Cabrera & Cabrera, 2002). It is indicated in some studies that individuals may lose their competitive advantage because of their knowledge sharing, which comes at a high cost and carries significant risk.

As a result, employees cannot share their knowledge unless there is a clear motivator, such as reciprocity (Chiu et al., 2006; Ipe, 2003). The current study aims to use social exchange theory to better understand how employees share their knowledge. Researchers (Cropanzano & Mitchell, 2005; Huang et al., 2008) argue that social exchange theory emphasizes reciprocity prospects or standards and can provide valuable insight into why people choose to share or not to share their information and knowledge with others. For instance, social exchange theory is utilized to analyze how trust, fairness, and management support are all connected to information or knowledge sharing (Mooradain et al., 2006; Wu et al., 2007).

Individuals with poor exchange ideology, on the other hand, do not easily change their attitudes or actions based on how other workers perceive them (Pazy & Ganzach, 2010). Many researchers (Chiaburu & Byrne, 2009; Coyle-Shapiro & Neuman, 2004; Pazy & Ganzach, 2010; Redman & Snape, 2005; Scott & Colquitt, 2007; Takeuchi et al., 2011) claim that exchange philosophies have a negative direct effect on attitudes and behaviors of employees. Individuals with a durable exchange philosophy prefer to keep score and are open to social exchange. In the absence of favorable care or rewards, they are more likely to exhibit negative attitudes and poor performance (Chiaburu & Byrne, 2009). Thus, based on previous literature, the current study suggests the following hypothesis:

H2*: LMX has a positive relationship with knowledge sharing.*

2.3 Knowledge sharing and work integration

Knowledge sharing is a common theme in knowledge integration, and it will be used to illustrate the implications that arise from the path-dependent nature of knowledge by farming the task of knowledge integration as a cycle (Carlile & Rebentisch, 2003). In line with the belief that individuals' com-

munications are the primary source of knowledge sharing, the term can be broadly defined as interpersonal exchanges that involve transmitting and receiving knowledge, information, or ideas from others (Argote & Ingram, 2000). A link between tacit and explicit knowledge and performance efficiency, especially in hostile environments, was discovered (Akgün et al., 2005). Since tacit knowledge is hidden, unspoken, and explicit, it can be expressed through individuals' social networks (Horváth, 2007). Knowledge must circulate and move continuously throughout the organization because it is considered valuable for work integration.

Furthermore, there is work integration and organizational performance as long as there is a flow of information, knowledge, and ideas (Stewart et al., 2000). Knowledge sharing is the transmission or distribution of knowledge from one person or community to another (Chieu Hsu, 2008). Knowledge sharing is a force that promotes knowledge development, exchange, and high performance, as well as job-related intellectual capital efficiency, which leads to work integration (Liebowitz, 2001). Considering the earlier literature, our third hypothesis is proposed as follows.

H3: *There is a positive relationship between knowledge sharing and work integration.*

2.4 Leader-member exchange, knowledge sharing and work integration

According to Dinh et al. (2014), LMX is the archetypal social exchange approach in leader-follower dyadic relationships. LMX is concerned with several characteristics that reflect the relationship between the leader and the members, such as re-

spect, constancy, and fairness (Liden & Maslyn, 1998). The importance of LMX in an organization is ultimately enhanced by a leader's fairness toward employees and other members of the organization (Colquitt et al., 2001). Wang & Noe (2010), on the other hand, discovered that the function of LMX affects people's job performance both directly and indirectly. Individuals provide a wide range of services in exchange for a desired outcome, such as pay and self-esteem (Wang & Noe, 2010). Employees with a high level of LMX affiliation are more operative in achieving high work integration (Gerstner & Day, 1997).

Why do employees not share their knowledge? Why do they share knowledge both on and off the internet? Why do they share knowledge through knowledge networks, teams, programs, and divisions? What exactly do they share, and how does it work in the real world? Many research scholars have been captivated by all these questions recently (Foss et al., 2010). Knowledge sharing can be defined as interpersonal communication linking the transmission and reception of knowledge from one person to another (Argote & Ingram, 2000). According to social exchange theory and the norms of mutuality, individuals who expect advantages or favorable treatment feel compelled to respond positively, even though the timing and form of their return may be uncertain (Blau, 1964; Gouldner, 1960). Moreover, negative returns could also be possible when exchange partners are out of control (Shore & Barksdale, 1998; Cropanzano & Mitchell, 2005). Based on these arguments, the fourth hypothesis is suggested as follows:

H4: *Knowledge sharing has a mediating effect on the relationship between LMX and WI.*

Figure 1 Research model
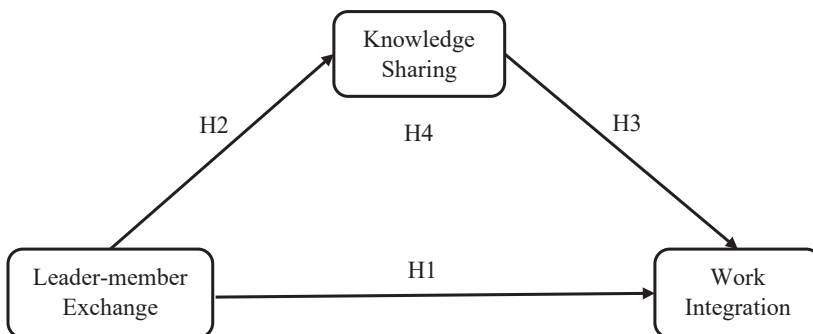
Source: Authors

3. Methodology

In the current study, we examined the relationship between LMX and work integration in the presence of an intervening variable (Knowledge Sharing). The data of this study was collected using the survey method from healthcare workers working in the hospital in Erzincan province in August and November during the COVID-19 period in 2021. The population consists of 895 healthcare workers working in public hospitals in Erzincan province. Using the convenience sampling method, data was obtained through a survey from the healthcare professionals participating in the study. The sample size required to represent the population in the study was determined to be 269 at a 95% confidence level (Ural & Kılıç, 2005, p. 43). Five hundred surveys were distributed to healthcare workers, and after removing incomplete and incorrectly filled surveys, 295 completed surveys were analyzed. The study was conducted during the COVID-19 pandemic to explore the perceptions of dedicated healthcare workers who risked their lives for their profession and see how they view the challenges posed by the pandemic. The aim was to examine leader-member interaction, knowledge sharing, and interactions of healthcare workers during the pandemic, who serve people especially during this difficult period and under hard conditions, and their interactions in the work integration process. Especially during the COVID-19 pandemic, effective communication between employees and their managers, along with knowledge sharing and work integration, was crucial for controlling the spread of the disease and preventing infections.

A purposive stratified sampling technique was used for data collection through the distribution of questionnaires designed on a Likert scale, ranging from (1) strongly disagree to (5) strongly agree. Additionally, the research questionnaire consisted of two different parts. First, the participants' sociodemographic characteristics such as gender, age, education, and experience were investigated, while in the second part, scales related to the variables of the current study, such as leader-member exchange, work integration and knowledge sharing, were included. Reliability analysis, factor analysis, regression, correlation, and mediation analysis were also employed. Moreover, AMOS version 17 and SPSS version 22 were used. Specifically, the PROCESS macro in SPSS was employed to examine the role of the mediation variable between the independent variable (LMX) and the dependent variable (Work Integration).

To provide further details about the measurement, our study utilized 12 items related to leader-member exchange, initially developed by Liden and Maslyn (1998) and later adopted by Turkish researcher Baş et al. (2010). The study included 4 items on knowledge sharing, developed by Chennamaneni (2012), and 17 items on work integration, developed by Schaufeli et al. (2002). Moreover, Cronbach's alpha value for the selected items was .8, which was reliable for data collection. The participants were instructed to respond using the designed Likert scale, ranging from (1) strongly disagree to (5) strongly agree, for each item in this study.

4. Results

4.1 Sociodemographic aspects of the study

Participants' characteristics were categorized by age, gender, educational level, experience, and marital status to understand the sociodemographic aspects of the study. Out of a total of 295 healthcare workers, 54.9% were female and 45.1% were male. Among them, 77.3% were married and 22.7% were single. The age distribution of the participants was as follows: 10.2% were aged 18-26, 28.5% were aged 27-35, 28.8% were aged 36-44, and 32.3% were aged 45 and above. It was essential to know the qualification level of the participants. The statistical results explained that 12.2% were high school graduates, 36.6% were undergraduates, and 51.2% were graduates. Furthermore, the participants' tenure was also explored in the study; 26.4% reported having 0-10 years of experience, 38.6% reported 11-21 years of experience, and 34.9% reported 22 or more years of experience in the healthcare sector.

4.2 Reliability and factor analysis results of the scales

Table 1 Mean for variables, Cronbach's alpha, Kaiser-Meyer-Olkin, standard deviation and correlation coefficients between variables

Variables	α	KMO	Mean.	SD.	1	2	3
Leader-Member Exchange	.933	.817	3.92	0.86	-		
Work Integration	.937	.842	3.95	0.75	.555**	-	
Knowledge Sharing	.920	.761	4.15	1.03	.213**	.484**	-

** = p<0.01 * = p<0.05
Source: Authors

Table 1 presents the mean, standard deviation, Cronbach's alpha, Kaiser-Meyer-Olkin, and correlation coefficients between the study variables. The statistics revealed that the Cronbach's alpha value for LMX, WI and KS is .70. The results also showed a positive correlation between the variables. LMX and WI were found at a 99% significant level (r = .555; P = .000). The study also revealed a positive relationship between LMX and KS (r = .213; P = .000) at a 99% significant level. Meanwhile, a positive correlation (r = .484; p = .000) at a 99% significance level was also found. Moreover, the KMO values and sphericity values of the scales met the reference values (KMO > 0.60 and sphericity value < 0.05) (Büyüköztürk, 2006).

Table 2 Goodness of fit values of the scales

Variables	χ2/df	RMSEA	CFI	GFI	NFI	TLI
Leader-Member Exchange	3.945	0.077	0.986	0.968	0.978	0.974
Work Integration	4.986	0.079	0.993	0.992	0.991	0.957
Knowledge Sharing	3.829	0.072	0.958	0.922	0.956	0.973

Source: Authors

The goodness of fit values of the variables in Table 2 show that they provide the referenced goodness of fit values (Hooper et al., 2008).

Table 3 Regression analysis results related to the mediation test (N=295)

Model summary	R	R²	F	P	df1	df2
	0.431	0.186	48.93	0.000	2.00	427.00
Knowledge Sharing						
Variables	B	SH	t	P	LLCI	ULCI
Constant	3.12	0.275	11.34	0.000	2.582	3.666
LMX	0.256	0.068	3.733	0.000	0.121	0.391
Work integration						
Variables	B	SH	t	P	LLCI	ULCI
Constant	1.196	0.182	6.572	0.000	0.837	1.554
Knowledge Sharing	0.382	0.032	8.599	0.000	0.213	0.339
LMX **(Direct Effect)**	0.412	0.038	10.638	0.000	0.335	0.488
LMX **(Total Effect)**	0.483	0.042	11.421	0.000	0.399	0.566
		Effect	SH	LLCI	ULCI	
Mediator **(Knowledge Sharing)** Effect		0.071	0.031	0.024	0.144	

Note: Standardized regression coefficients are specified.
Source: Authors

It was decided that regression analysis based on the bootstrap method should be applied to analyze whether knowledge sharing mediates the relationship between healthcare professionals' leader-member interaction and work integration. It was suggested that the bootstrap method gives more consistent results than the method used by Baron and Kenny (1986) and the Sobel test (Gürbüz, 2019; Hayes, 2018). The PROCESS macro application developed by Hayes (2018) was used to analyze the scales in Table 3. Model 4 was chosen in the analysis application, and the 5000 resampling option was

selected by the bootstrap technique. In mediation effect analysis performed by bootstrap, CI (confidence interval) values were at 95%. The confidence interval should not be zero (0) to support the hypotheses (Gürbüz, 2019). Table 3 shows the coefficients of the multiple regression model. According to this table, it was questioned whether there is a significant and positive relationship between leader-member interaction and work integration in the H1 hypothesis.

The statistical results showed a significant positive relationship between LMX and work integration (b = 0.412; p = 0.000). Therefore, the H1 hypothesis was accepted. In the H2 hypothesis, a significant positive relationship exists between leader-member interaction and knowledge sharing. Accordingly, when examining the H2 hypothesis, it is shown that leader-member interaction accounts for 25.6%

of knowledge sharing. According to this finding, the H2 hypothesis was accepted. The H3 hypothesis questioned a significant positive relationship between knowledge sharing and work integration. The statistical results revealed that knowledge sharing explains 38.2% of work integration. Accordingly, the H3 hypothesis was accepted. The significance of the relationships between variables allowed us to question whether there is a mediating effect. The fourth hypothesis examined whether knowledge sharing had a partial mediating effect on the relationship between leader-member interaction and work integration. When it is added to the model as a means of knowledge sharing, it is seen that knowledge sharing has a mediating effect on the relationship between LMX and work integration (b = .071, 95% BCA CI [.024, .144]). In this context, the H4 hypothesis was also accepted.

Figure 2 Research model
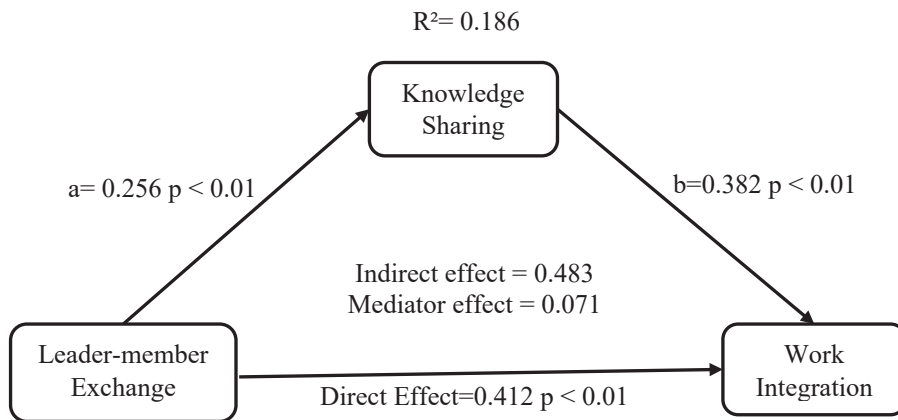
Note: Standardized beta coefficients are reported. The R² value shows the variance explained.
Source: Authors

5. Discussion

The current study explored the mediating role of knowledge sharing between LMX and WI of the selected healthcare employees in Türkiye. Our findings indicate that LMX as a leadership theory focuses on the relationship between leaders and employees within the organization and has a positive effect on work integration. Importantly, our study revealed that knowledge sharing plays a significant mediating role between LMX and WI. A close examination of the data showed that LMX is positive-

ly related to knowledge sharing, particularly at high levels of WI. As a result, leaders' interactions with employees can have a significant impact on factors such as cooperation, trust, and commitment. In this context, knowledge sharing can strengthen the impact of LMX on WI. Leaders can increase the flow of information by encouraging open communication with employees. This can enable employees to better understand each other, exchange ideas and find solutions together. Sharing knowledge can also build trust among employees and foster team spirit.

Leaders can also create appropriate environments to encourage knowledge sharing among employees. For example, tools such as regular meetings, team workshops or digital platforms for knowledge sharing can be used. This increases the flow of information between team members and strengthens work integration.

To extend our knowledge, this is the first study that has verified the mediating role of knowledge sharing between LMX and WI in Türkiye's healthcare sector. For this reason, we integrated WI related literature with LMX theory. The study's findings further suggest that implementing better and higher-quality LMX can enhance creativity and WI. Our study also revisited earlier studies that focused on affirmative interpersonal relationships at workplaces to empower organizations (Dutton, 2003; Heaphy & Dutton, 2008), which may add new insight and importance to organizational studies. Further, this study suggests that individuals who possess positive social connections with their line managers and supervisors, with mutual attentiveness, trust, and high work integration, are more innovatively involved in their assigned tasks and responsibilities (Macey & Schneider, 2008; Kanungo, 1982) and lead toward job commitment (Golden & Veiga, 2008; Allen & Meyer, 1990). Finally, the practical implication of our findings suggests that LMX positively affects WI. Knowledge sharing also has a positive role in LMX and WI.

6. Limitation and future direction

This study has some limitations that must be addressed in future studies. We adopted a subordinate-centered perspective to explain the employees' views on the quality of the relationship between leaders and members. Future studies should incorporate the leaders' perspectives as well (Graen & Uhl-Bien, 1995; Gerstner & Day, 1997; Scandura & Schriesheim, 1994; Howard-Schwind, 2010) to provide a more detailed understanding of the relationship between LMX and WI. Second, our findings were based on self-reported data, so standard method bias cannot be completely ruled out (Podsakoff et al., 2003). Third, we were interested in evaluating the role of knowledge sharing between LMX and WI; our findings captured employees' subjective perception of WI (Atwater & Carmeli, 2009). Future researchers should examine the relationship between LMX and other variables, such

as work involvement, work autonomy, and creative work behavior. Finally, the fourth direction toward future studies might be exploring the moderating relationship between variables, as we only explored the mediating role of knowledge sharing with LMX and WI.

7. Conclusion

To sum up, this study has explored the crucial problem of promoting WI among Turkish healthcare professionals, particularly elucidating the critical function that knowledge sharing plays as a mediator in the context of LMX. The study's conclusions provide valuable insights into how interpersonal connections, leadership, and knowledge sharing function in the healthcare industry. The need to encourage seamless work integration is becoming more apparent as healthcare organizations adapt to the demands of a constantly changing environment. The results highlight the beneficial effects of LMX on WI and the significant role of leaders in fostering a diverse and cooperative work environment. Furthermore, a mediating effect of knowledge sharing highlights its function as a catalyst for converting the constructive exchanges between leaders and healthcare professionals into concrete results, promoting solidarity, cooperation, and shared knowledge.

Moreover, our study provided critical answers to the following questions: (1) It was discovered that LMX significantly improved the extent to which Turkish healthcare personnel integrated into their workplaces. The positive dynamics between leaders and members are crucial, as evidenced by the correlation between higher-quality relationships and greater levels of work integration. (2) The study found a substantial correlation between knowledge sharing and LMX among Turkish healthcare professionals. Health professionals were more likely to participate in knowledge sharing activities and promote a cooperative knowledge sharing culture if they reported higher-quality interactions with their bosses. (3) Knowledge sharing was a significant predictor of job integration among Turkish healthcare workers. A culture of knowledge sharing favors the overall integration of healthcare teams, as evidenced by the positive correlation between increased knowledge sharing and improved work integration. (4) The study verified that knowledge sharing is a mediator between WI and LMX among

healthcare workers. The positive relationship between WI and LMX was mediated by knowledge sharing, indicating the mediating role that knowledge sharing plays in promoting cohesive and integrated healthcare staff.

References

1. Adil, M. S. & Awais, A. (2016). Effects of leader-member exchange, interpersonal relationship, individual feeling of energy and creative work involvement towards turnover intention: a path analysis using structural equation modeling. *Asian Academy of Management Journal*, *21*(2), 99-133. https://doi.org/10.21315/aamj2016.21.2.5

2. Akgün, A. E., Byrne, J., Keskin, H., Lynn, G. S. & Imamoglu, S. Z. (2005). Knowledge networks in new product development projects: A transactive memory perspective. *Information & Management*, *42*(8), 1105-1120. https://doi.org/10.1016/j.im.2005.01.001

3. Allen, N. J. & Meyer, J. P. (1990). Organizational socialization tactics: A longitudinal analysis of links to newcomers' commitment and role orientation. *Academy of Management Journal*, *33*(4), 847-858. https://doi.org/10.2307/256294

4. Argote, L. & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes*, *82*(1), 150-169. https://doi.org/10.1006/obhd.2000.2893

5. Arsintadiani, D. & Harsono, M. (2002). Pengaruh tingkat LMX terhadap penilaian kinerja dan kepuasan kerja dengan kesamaan gender dan locus of control sebagai variabel moderator. *Jurnal Perspektif*, *7*(2), 113-122.

6. Atwater, L. & Carmeli, A. (2009). Leader–member exchange, feelings of energy, and involvement in creative work. *The Leadership Quarterly*, *20*(3), 264-275. https://doi.org/10.1016/j.leaqua.2007.07.009

7. Bae, Y. S., Kang, S. W., Seo, M. S., Baines, I. C., Tekle, E., Chock, P. B. & Rhee, S. G. (1997). Epidermal growth factor (EGF)-induced generation of hydrogen peroxide: role in EGF receptor-mediated tyrosine phosphorylation. *Journal of Biological Chemistry*, *272*(1), 217-221. https://doi.org/10.1074/jbc.272.1.217

8. Baron, M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173-1182. https://doi.org/10.1037/0022-3514.51.6.1173

9. Baş, T., Keskin, N. & Mert, İ. S. (2010). Lider Üye Etkileşimi (Lüe) Modeli Ve Ölçme Aracının Türkçe'de Geçerlik Ve Güvenilirlik Analizi. *Ege Academic Review*, *10*(3), 1013-1039. https://doi.org/10.21121/eab.2010319626

10. Blau, P. M. (1964). Justice in social exchange. *Sociological Inquiry*, *34*(2), 193-206. https://doi.org/10.1111/j.1475-682X.1964.tb00583.x

11. Büyüköztürk, Ş. (2006). *Sosyal bilimler için veri analizi el kitabı* (6th ed). Pegem A Yayıncılık.

12. Cabrera, A. & Cabrera, E. F. (2002). Knowledge-sharing dilemmas. *Organization Studies*, *23*(5), 687-710. https://doi.org/10.1177/0170840602235001

13. Cabrera, E. F. & Cabrera, A. (2005). Fostering knowledge sharing through people management practices. *The International Journal of Human Resource Management*, *16*(5), 720-735. https://doi.org/10.1080/09585190500083020

14. Carlile, P. R. & Rebentisch, E. S. (2003). Into the black box: The knowledge transformation cycle. *Management Science*, *49*(9), 1180-1195. https://doi.org/10.1287/mnsc.49.9.1180.16564

15. Chan, S. C. & Mak, W. M. (2012). Benevolent leadership and follower performance: The mediating role of leader–member exchange (LMX). *Asia Pacific Journal of Management*, *29*(2), 285-301. https://doi.org/10.1007/s10490-011-9275-3

16. Chennamaneni, A., Teng, J. T. & Raja, M. K. (2012). A unified model of knowledge sharing behaviours: theoretical development and empirical test. *Behaviour & Information Technology*, *31*(11), 1097-1115. https://doi.org/10.1080/0144929X.2011.624637

17. Chiaburu, D. S. & Byrne, Z. S. (2009). Predicting OCB role definitions: Exchanges with the organization and psychological attachment. *Journal of Business and Psychology*, *24*(2), 201-214. https://doi.org/10.1007/s10869-009-9100-x

18. Chieu Hsu, I. (2008). Knowledge sharing practices as a facilitating factor for improve organizational performance through human capital: A preliminary test. *Expert Systems with Applications*, *35*(1), 1316-1326. https://doi.org/10.1016/j.eswa.2007.08.012

19. Chiu, C. M., Hsu, M. H. & Wang, E. T. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems*, *42*(3), 1872-1888. https://doi.org/10.1016/j.dss.2006.04.001

20. Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. & Ng, K. Y. (2001). Justice at the millennium: a meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, *86*(3), 425-445. https://doi.org/10.1037/0021-9010.86.3.425

21. Coyle-Shapiro, J. A. & Neuman, J. H. (2004). The psychological contract and individual differences: The role of exchange and creditor ideologies. *Journal of Vocational Behavior*, *64*(1), 150-164. https://doi.org/10.1016/S0001-8791(03)00031-9

22. Cropanzano, R. & Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of Management*, *31*(6), 874-900. https://doi.org/10.1177/0149206305279602

23. Dansereau Jr, F., Cashman, J. & Graen, G. (1973). Instrumentality theory and equity theory as complementary approaches in predicting the relationship of leadership and turnover among managers. *Organizational Behavior and Human Performance*, *10*(2), 184-200. https://doi.org/10.1016/0030-5073(73)90012-3

24. Dinh, J. E., Lord, R. G., Gardner, W. L., Meuser, J. D., Liden, R. C. & Hu, J. (2014). Leadership theory and research in the new millennium: Current theoretical trends and changing perspectives. *The Leadership Quarterly*, *25*(1), 36-62. https://doi.org/10.1016/j.leaqua.2013.11.005

25. Dutton, J. E. (2003). *Energize Your Workplace: How to Create and Sustain High-quality Connections at Work*. John Wiley & Sons.

26. Du, R., Ai, S. & Ren, Y. (2007). Relationship between knowledge sharing and performance: A survey in Xi'an, China. *Expert Systems with Applications*, *32*(1), 38-46. https://doi.org/10.1016/j.eswa.2005.11.001

27. Foss, N. J., Husted, K. & Michailova, S. (2010). Governing knowledge sharing in organizations: Levels of analysis, governance mechanisms, and research directions. *Journal of Management Studies*, *47*(3), 455-482. https://doi.org/10.1111/j.1467-6486.2009.00870.x

28. French, J. R., Raven, B. & Cartwright, D. (1959). The bases of social power. In Shafritz, J. M. et al. (Eds.), *Classics of Organization Theory* (7th ed) (pp. 311-320). Wadsworth Cengage Learning.

29. Gerstner, C. R. & Day, D. V. (1997). Meta-analytic review of leader–member exchange theory: Correlates and construct issues. *Journal of Applied Psychology*, *82*(6), 827-844. https://doi.org/10.1037/0021-9010.82.6.827

30. Golden, T. D. & Veiga, J. F. (2008). The impact of superior–subordinate relationships on the commitment, job satisfaction, and performance of virtual workers. *The Leadership Quarterly*, *19*(1), 77-88. https://doi.org/10.1016/j.leaqua.2007.12.009

31. Gouldner, A. W. (1960). The norm of reciprocity. *American Sociological Review*, *25*(2), 165-167. https://doi.org/10.2307/2092623

32. Graen, G. B. & Uhl-Bien, M. (1995). Relationship-based approach to leadership: Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective. *The Leadership Quarterly*, *6*(2), 219-247. https://doi.org/10.1016/1048-9843(95)90036-5

33. Graen, G. & Schiemann, W. (1978). Leader–member agreement: A vertical dyad linkage approach. *Journal of Applied Psychology*, *63*(2), 206-212. https://doi.org/10.1037/0021-9010.63.2.206

34. Gürbüz, S. (2019). *Amos ile yapısal eşitlik modellemesi.* Seçkin Yayıncılık.

35. Hayes, A. F. (2018). Partial, conditional, and moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, *85*(1), 4-40. https://doi.org/10.1080/03637751.2017.1352100

36. Heaphy, E. D., & Dutton, J. E. (2008). Positive social interactions and the human body at work: Linking organizations and physiology. *Academy of Management Review*, *33*(1), 137-162. https://doi.org/10.5465/amr.2008.27749365

37. Hong, P., Doll, W. J., Nahm, A. Y. & Li, X. (2004). Knowledge sharing in integrated product development. *European Journal of Innovation Management*, *7*(2), 102-112. https://doi.org/10.1108/14601060410534393

38. Hooper, D., Coughlan, J. & Mullen, M. (2008). Evaluating model fit: a synthesis of the structural equation modelling literature. In Brown, A. (Ed.). *Proceedings of the 7th European Conference on Research Methodology for Business and Management Studies* (pp. 195-200). London: Regent's College.

39. Horváth, I. (2007). Comparison of three methodological approaches of design research. In *DS 42: Proceedings of ICED 2007, the 16th International Conference on Engineering Design* (pp. 361-362). Paris: Design Society.

40. Howard-Schwind, M. (2010). *Instructional Leadership Responsibilities of Assistant Principals in Large Texas High Schools.* University of North Texas.

41. Huang, Q., Davison, R. M. & Gu, J. (2008). Impact of personal and cultural factors on knowledge sharing in China. *Asia Pacific Journal of Management*, *25*(3), 451-471. https://doi.org/10.1007/s10490-008-9095-2

42. Ilies, R., Nahrgang, J. D. & Morgeson, F. P. (2007). Leader-member exchange and citizenship behaviors: A meta-analysis. *Journal of Applied Psychology*, *92*(1), 269-277. https://doi.org/10.1037/0021-9010.92.1.269

43. Ipe, M. (2003). Knowledge sharing in organizations: A conceptual framework. *Human Resource Development Review*, *2*(4), 337-359. https://doi.org/10.1177/1534484303257985

44. Jang, S., Hong, K., Bock, G. W. & Kim, I. (2002). Knowledge management and process innovation: the knowledge transformation path in Samsung SDI. *Journal of Knowledge Management*, *6*(5), 479-485. https://doi.org/10.1108/13673270210450582

45. Jaussi, K. S. & Dionne, S. D. (2003). Leading for creativity: The role of unconventional leader behavior. *The Leadership Quarterly*, *14*(4-5), 475-498. https://doi.org/10.1016/S1048-9843(03)00048-1

46. Jung, D. I., Chow, C. & Wu, A. (2003). The role of transformational leadership in enhancing organizational innovation: Hypotheses and some preliminary findings. *The Leadership Quarterly*, *14*(4-5), 525-544. https://doi.org/10.1016/S1048-9843(03)00050-X

47. Kanungo, R. N. (1982). Measurement of job and work involvement. *Journal of Applied Psychology*, *67*(3), 341. https://doi.org/10.1037/0021-9010.67.3.341

48. Kenney, J. L. & Gudergan, S. P. (2006). Knowledge integration in organizations: an empirical assessment. *Journal of Knowledge Management*, *10*(4), 43-58. https://doi.org/10.1108/13673270610679354

49. Liden, R. C. & Maslyn, J. M. (1998). Multidimensionality of leader-member exchange: An empirical assessment through scale development. *Journal of Management*, *24*(1), 43-72. https://doi.org/10.1016/S0149-2063(99)80053-1

50. Liebowitz, J. (2001). Knowledge management and its link to artificial intelligence. *Expert Systems with Applications*, *20*(1), 1-6. https://doi.org/10.1016/S0957-4174(00)00044-0

51. Macey, W. H. & Schneider, B. (2008). The meaning of employee engagement. *Industrial and Organizational Psychology*, *1*(1), 3-30. https://doi.org/10.1111/j.1754-9434.2007.0002.x

52. Michel, J. W. & Tews, M. J. (2016). Does leader–member exchange accentuate the relationship between leader behaviors and organizational citizenship behaviors? *Journal of Leadership & Organizational Studies*, *23*(1), 13-26. https://doi.org/10.1177/1548051815606429

53. Mooradian, T., Renzl, B. & Matzler, K. (2006). Who trusts? Personality, trust and knowledge sharing. *Management Learning*, *37*(4), 523-540. https://doi.org/10.1177/1350507606073424

54. Nonaka, I. & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press. https://doi.org/10.1093/oso/9780195092691.001.0001

55. Nonaka, I., Byosiere, P., Borucki, C. C. & Konno, N. (1994). Organizational knowledge creation theory: a first comprehensive test. *International Business Review*, *3*(4), 337-351. https://doi.org/10.1016/0969-5931(94)90027-2

56. Nooteboom, B. (1999). Innovation, learning and industrial organisation. *Cambridge Journal of Economics*, *23*(2), 127-150. https://doi.org/10.1093/cje/23.2.127

57. Oldham, G. R. & Cummings, A. (1996). Employee creativity: Personal and contextual factors at work. *Academy of Management Journal*, *39*(3), 607-634. https://doi.org/10.2307/256657

58. Örtqvist, D. & Wincent, J. (2006). Prominent consequences of role stress: A meta-analytic review. *International Journal of Stress Management*, *13*(4), 399-422. https://doi.org/10.1037/1072-5245.13.4.399

59. Pazy, A. & Ganzach, Y. (2010). Predicting committed behavior: Exchange ideology and pre-entry perceived organisational support. *Applied Psychology*, *59*(2), 339-359. https://doi.org/10.1111/j.1464-0597.2009.00393.x

60. Podsakoff, N. P., LePine, J. A. & LePine, M. A. (2007). Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: A meta-analysis. *Journal of Applied Psychology*, *92*(2), 438-454. https://doi.org/10.1037/0021-9010.92.2.438

61. Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y. & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903. https://doi.org/10.1037/0021-9010.88.5.879

62. Quigley, N. R., Tesluk, P. E., Locke, E. A. & Bartol, K. M. (2007). A multilevel investigation of the motivational mechanisms underlying knowledge sharing and performance. *Organization Science*, *18*(1), 71-88. https://doi.org/10.1287/orsc.1060.0223

63. Rank, J., Pace, V. L. & Frese, M. (2004). Three avenues for future research on creativity, innovation, and initiative. *Applied Psychology*, *53*(4), 518-528. https://doi.org/10.1111/j.1464-0597.2004.00185.x

64. Redman, T. & Snape, E. (2005). Exchange ideology and member-union relationships: An evaluation of moderation effects. *Journal of Applied Psychology*, *90*(4), 765-773. https://doi.org/10.1037/0021-9010.90.4.765

65. Rugian, M. S., Saerang, I. & Lengkong, V. P. (2017). Pengaruh disiplin kerja, pelatihan, kualitas kehidupan kerja dan konflik pekerjaan-keluarga terhadap kinerja karyawan (studi kasus pada PT Bank BTPN Tbk Cabang Utama Manado). *Jurnal EMBA: Jurnal Riset Ekonomi, Manajemen, Bisnis Dan Akuntansi*, *5*(2), 487-611.

66. Sa'adah, N. & Rijanti, T. (2022). The Role of Knowledge Sharing, Leader-Member Exchange (LMX) on Organizational Citizenship Behavior and Employee Performance: An Empirical Study on Public Health Center of Pati 1, Pati 2 and Trangkil in Central Java. *International Journal of Social and Management Studies*, *3*(1), 112-131.

67. Scandura, T. A. & Schriesheim, C. A. (1994). Leader-member exchange and supervisor career mentoring as complementary constructs in leadership research. *Academy of Management Journal*, *37*(6), 1588-1602. https://doi.org/10.2307/256800

68. Schaufeli, W. B., Salanova, M., González-Romá, V. & Bakker, A. B. (2002). The Measurement of Engagement and Burnout: A Two Sample Confirmatory Factor Analytic Approach. *Journal of Happiness Studies*, *3*(1), 71-92. https://doi.org/10.1023/A:1015630930326

69. Scott, B. A. & Colquitt, J. A. (2007). Are organizational justice effects bounded by individual differences? An examination of equity sensitivity, exchange ideology, and the Big Five. *Group & Organization Management*, *32*(3), 290-325. https://doi.org/10.1177/1059601106286877

70. Scott, S. G. & Bruce, R. A. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. *Academy of Management Journal*, *37*(3), 580-607. https://doi.org/10.2307/256701

71. Shore, L. M. & Barksdale, K. (1998). Examining degree of balance and level of obligation in the employment relationship: A social exchange approach. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *19*(S1), 731-744. https://doi.org/10.1002/(SICI)1099-1379(1998)19:1+<731::AID-JOB969>3.0.CO;2-P

72. Sosik, J. J., Kahai, S. S. & Avolio, B. J. (1998). Transformational leadership and dimensions of creativity: Motivating idea generation in computer-mediated groups. *Creativity Research Journal*, *11*(2), 111-121. https://doi.org/10.1207/s15326934crj1102_3

73. Stewart, K. A., Baskerville, R., Storey, V. C., Senn, J. A., Raven, A. & Long, C. (2000). Confronting the assumptions underlying the management of knowledge: an agenda for understanding and investigating knowledge management. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, *31*(4), 41-53. https://doi.org/10.1145/506760.506764

74. Takeuchi, R., Yun, S. & Wong, K. F. E. (2011). Social influence of a coworker: A test of the effect of employee and coworker exchange ideologies on employees' exchange qualities. *Organizational Behavior and Human Decision Processes*, *115*(2), 226-237. https://doi.org/10.1016/j.obhdp.2011.02.004

75. Tierney, P., Farmer, S. M. & Graen, G. B. (1999). An examination of leadership and employee creativity: The relevance of traits and relationships. *Personnel Psychology*, *52*(3), 591-620. https://doi.org/10.1111/j.1744-6570.1999.tb00173.x

76. Unsworth, K. (2001). Unpacking creativity. *Academy of Management Review*, *26*(2), 289-297. https://doi.org/10.2307/259123

77. Ural, A. & Kılıç, İ. (2005). *Bilimsel Araştırma Süreci ve SPSS ile Veri Analizi*. Detay Yayıncılık Ankara.

78. Wang, S. & Noe, R. A. (2010). Knowledge sharing: A review and directions for future research. *Human Resource Management Review*, *20*(2), 115-131. https://doi.org/10.1016/j.hrmr.2009.10.001

79. Widen-Wulff, G. & Suomi, R. (2007). Utilization of information resources for business success: The knowledge sharing model. *Information Resources Management Journal*, *20*(1), 46-67. https://doi.org/10.4018/irmj.2007010104

80. Wu, W. L., Hsu, B. F. & Yeh, R. S. (2007). Fostering the determinants of knowledge transfer: a team-level analysis. *Journal of Information Science*, *33*(3), 326-339. https://doi.org/10.1177/0165551506070733

81. Zahra, S. A., Neubaum, D. O. & Hayton, J. (2020). What do we know about knowledge integration: Fusing micro- and macro-organizational perspectives. *Academy of Management Annals*, *14*(1), 160-194. https://doi.org/10.5465/annals.2017.0093

82. Zhang, X. & Bartol, K. M. (2010). Linking empowering leadership and employee creativity: The influence of psychological empowerment, intrinsic motivation, and creative process engagement. *Academy of Management Journal*, *53*(1), 107-128. https://doi.org/10.5465/amj.2010.48037118

83. Zhou, J. & George, J. M. (2003). Awakening employee creativity: The role of leader emotional intelligence. *The Leadership Quarterly*, *14*(4-5), 545-568. https://doi.org/10.1016/S1048-9843(03)00051-1

Anita Peša
University of Zadar
Department of Economics
23000 Zadar, Croatia
apesa@unizd.hr

Marijana Jerić
University of Zadar
Department of Economics
23000 Zadar, Croatia
mjeric@unizd.hr

Miguel Ángel Latorre Guillem
University Catholic of
Valencia "San Vicente Martir"
Faculty of Law, Economics
and Social Sciences
Department of Accounting,
Finance and Management Control
Headquarters Valencia-Marques
de Campo
46003 Valencia, Spain
mangel.latorre@ucv.es

THE IMPACT OF MACROECONOMIC FACTORS ON REAL ESTATE PRICES: EVIDENCE FROM SPAIN AND CROATIA

Abstract

Purpose: The paper deals with the comparison of real estate price trends in two European markets such as Spain and Croatia, well-known tourist destinations. Trends in macroeconomic variables in the period from 2013 to 2022 were considered, such as the house price index, overdue mortgage or rent payments, direct investment, production in construction, real GDP per capita, average adjusted wage per employee and unemployment rate. The aim of this study is to compare real estate prices in these two countries with regard to macroeconomic variables.

Methodology: The applied methodology is based on the model averaging technique, which has been used in a few previous similar research studies.

Results: The obtained results point to the fact that real estate price movements are strongly negatively influenced by the unemployment rate in both observed countries.

Conclusion: The real estate prices in Croatia are strongly negatively influenced by arrears on mortgage or rent payments, while in Spain, they are strongly positively influenced by production in construction. Recommendations for further research refer to research that would cover a wider range of Southeast European countries with common characteristics of strongly developed tourism.

Keywords: Determinants of real estate prices, macroeconomic factors, real estate market, model averaging technique

1. Introduction

The growth of real estate prices has been a trend in the last few years in all tourist destinations in Europe. Many studies have been conducted on the subject of the real estate market in order to determine what factors most influence price changes. While Spain and Croatia demonstrate a similar orientation towards tourism, additional research is important to determine which variables influence real estate prices in each country and whether these variables align or differ between the two countries.

Concerning the examination of price dynamics within the Spanish real estate market, one of the focal countries in this analysis, inflation is currently experiencing a notable increase, aligning closely with trends observed across the Eurozone. In response to this scenario, European regulatory bodies have decided to increase the nominal interest rate, or Euribor, as a measure to manage inflationary pressures. These macroeconomic policy decisions exert multifaceted impacts, influencing, firstly, the cost associated with mortgage loans in Spain. Additionally, they affect broader economic indicators such as growth rates and job creation. Consequently, the rate of escalation of property market prices in Spain is predicted to slow down, which can be attributed to the limited availability of mortgage financing. However, the trajectory of property prices in Spain has exhibited an unconventional pattern since the start of the pandemic, showing increased activity amid international economic uncertainty.

The real estate market in Croatia has experienced significant growth in recent years, with property prices reaching unprecedented levels. Apart from the entry into the Eurozone, what strongly influences the growth of real estate prices in Croatia is the developed tourism industry, which is now opening its doors to foreigners who want to engage in tourism in the Republic of Croatia. In addition, COVID-19 and remote work have prompted many EU citizens to buy real estate in Mediterranean countries. Similarly to Spain, there is a strong influence of inflation growth on the prices of goods and services and thus on the real estate market.

The subject of this research analyzes the real estate markets in Spain and Croatia, focusing on real estate prices and macroeconomic variables typically used in similar research, namely: direct mortgage or rent payments, investments, production in construction, real GDP per capita, average adjusted wage per employee and unemployment rate. Therefore, the null and alternative hypotheses are established as follows: The null hypothesis posits that none of the selected variables significantly influence housing prices, while the alternative hypothesis asserts that at least one variable has a non-zero coefficient that impacts housing prices.

Given that the real estate market is renowned for its high volatility within the economy, the relationship between real estate prices and macroeconomic factors has garnered substantial attention from both researchers and policymakers in recent decades. For economists who subscribe to the notion that government policies can effectively manage the economy, the real estate market emerges as a crucial instru-

ment in this management process. Consequently, this analysis seeks to shed light on whether the price trends in real estate can be elucidated by macroeconomic fundamentals. Considering the plethora of studies conducted in transitional economies, we can discern how the findings in this country compare to those in nations of similar development levels.

Conversely, the econometric methods employed to scrutinize the determinants of real estate can be categorized into several groups: vector autoregressive models (VAR models), error correction models (ECM) approaches, panel regression, and simple and multiple linear regression. This study will employ a methodology founded on the model averaging technique; to the best of our knowledge, this method has been infrequently utilized in prior research addressing this subject in other countries.

The applied methodology is based on the model averaging technique, which has not been used often in previous research on real estate prices. The obtained results point to the fact that the movement of real estate prices best describes and is most closely related to unemployment and late mortgage or rent payments. The first part provides an overview of the results of empirical studies which examined real estate price determinants. The second part describes the data and defines the variables which will be included in the model. The third part describes the methodology that was used for data modeling, and subsequently moves on to the analysis of the empirical results. The last part of the study summarizes the most important results of the analysis which was carried out.

2. Theory review

In the last few decades, numerous authors have dealt with the real estate industry and researched what affects the movement of real estate prices. Different macroeconomic variables were examined in different countries, such as gross domestic product, GDP (Adams & Fuss, 2010), interest rates, budget ratio, mortgage rate, industrial production, consumption of households foreign direct investment, real shares price, income and inflation rate, (Lovrinčević & Vižek, 2008, Grum & Govekar, 2016), trade, unemployment rate (Kotseva and Yanchev, 2017), construction costs and number of households (Pashardes & Savva, 2009; Sivitanies, 2015), the current account balance (Allen et al., 2016), the price of business real estate (Giussani et al., 1993), and others. The influence of these macroeconomic variables was examined with the help of different methodologies such as Ordinary Least Squares (OLS), the multiple linear regression model (Grum and Govekar, 2016; Gas-

pareniene et el., 2017), the Analytic Hierarchy Process (APH) (Tupenaite et al., 2017), the ECM model (Lovrinčević & Vižek, 2008; Weiner & Fuers, 2015), Structural Vector Auto Regression (SVAR) (Dumi**čić** et al., 2012), and the model averaging technique (Radonjić et al., 2019). Most of the authors agreed that many of these macroeconomic variables influence the movement of real estate prices, but there are also authors that prove the opposite relationship. The movement of real estate prices affects the movement of macroeconomic variables. The authors such as Mavrodiy (2005), Hon-Chung (2009) and Grum and Govekar (2016) claimed that the prices of real estate and activity are different in different countries due to a specific environment. GDP and unemployment are the most important factors linked to the price of business real estate rental (Chin, 2003).

In reference to Spain, there are several authors that indicate previous evidence. Authors Álvarez-Román and Garcia-Posada (2021) evaluate whether real house prices in Spain are justified by their long-run fundamentals, such as per capita real income, unemployment rate and population density. They assess house price overvaluation across the 50 Spanish regions in the period from 1985 to 2018. As overvaluation is mostly explained by high household leverage and the business cycle. Zhenyu and Taltavull (2020) examine the macroeconomic determinants that affect International Capital Flows towards the Spanish real estate market over the period from 1995 to 2017. This paper notes that another reason for the increase in Spanish real estate prices is the association with Spanish economic growth. It suggests that as GDP growth rates increase, the flow of capital into the real estate sector will rise, potentially increasing the willingness to purchase Spanish real estate and driving up prices. Lourenco and Rodrigues (2014) present

a significant expansion of residential investment since 2000, well above GDP dynamics. An indicator of excess housing stock suggests that around 30% of dwellings in Spain were excess in 2006. From 2007 onward, there was a sharp contraction of residential investment in Spain, which is distinct from the less pronounced evolution of GDP. Ghirelli et al. (2023) provide a metric to measure the synchronization of housing price cycles across Spanish cities and study changes in city-level price synchronization over time. They focus on the period from the first quarter of 1989 to the first quarter of 2018. In this study, the authors show that differences in population growth and economic structure are key factors in explaining the evolution of house price synchronization across Spanish cities.

3. Data description

This paper uses annual data from 2013 to 2022. The data were collected from the Eurostat database. The dependent variable is the House price index that captures price changes of all residential properties purchased by households (flats, detached houses, terraced houses, etc.), both new and existing, independently of their final use and their previous owners. Only market prices are considered, self-build dwellings are therefore excluded. The land component is included. The data is expressed as an annual average index 2015=100, showing the three-year percentage change and the annual average rate of change. The independent variables are Arrears on mortgage or rent payments, Direct investment[1], Production in construction, Real GDP per capita[2], Average full time adjusted salary per employee, and Total unemployment rate[3], as shown in Table 1. All variables are collected for two observed countries – Croatia and Spain.

Table 1 Macroeconomics variables and symbols

Variable	Symbol
House Price Index (dependent variable)	HPI
Arrears on mortgage or rent payments	AOM
Direct investment	DI
Production in construction	PIC
Real GDP per capita	GDP
Average full time adjusted salary per employee	SALA
Total unemployment rate	UNEM

Source: Created by the authors

---

[1] Data for the independent variable Direct investment is used as a value in million of units regarding the national currency.

[2] The indicator is calculated as the ratio of real GDP to the average population of a specific year.

[3] Data for the independent variable Total unemployment rate refers to unemployed people between the ages of 15 and 74.

The first step in creating a model is to see the trends for the observed periods. Figures 1 and 2 show the trends for Croatia and Spain, respectively.

Figure 1 Variable trends for Croatia
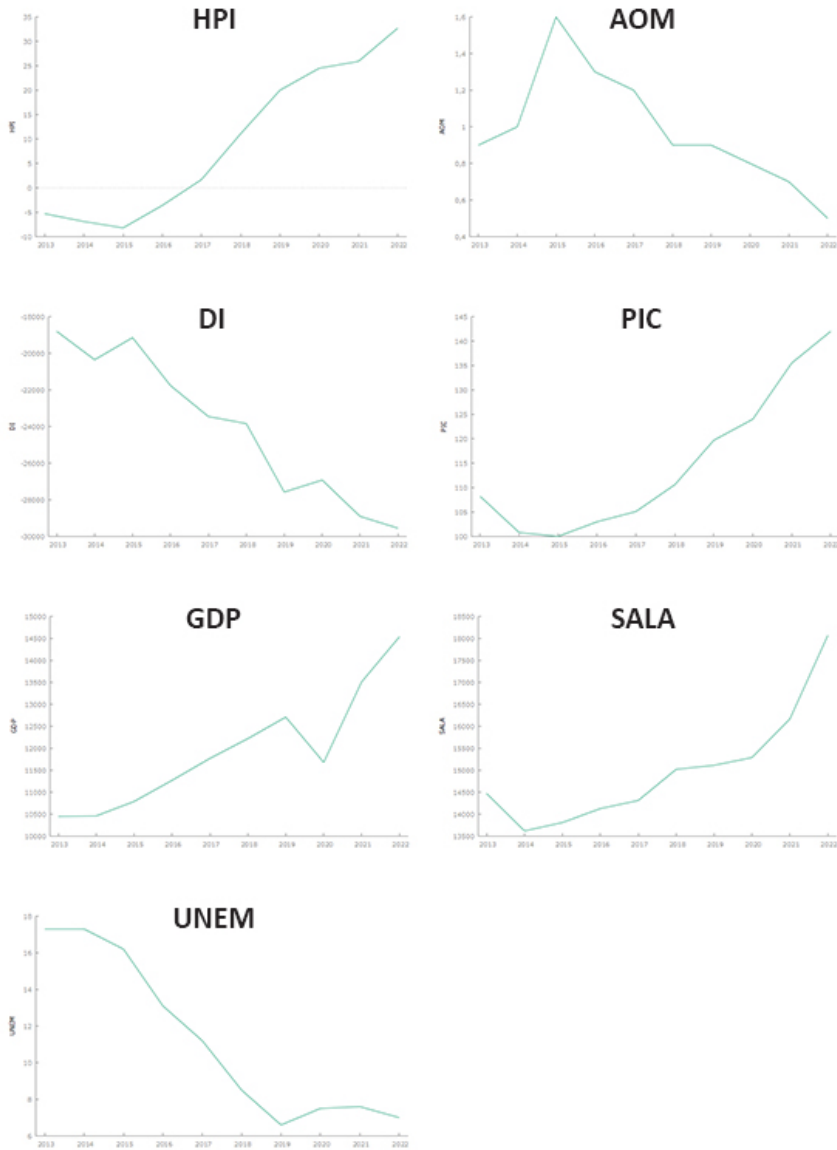
Source: Created by the authors

Figure 2 Variable trends for Spain
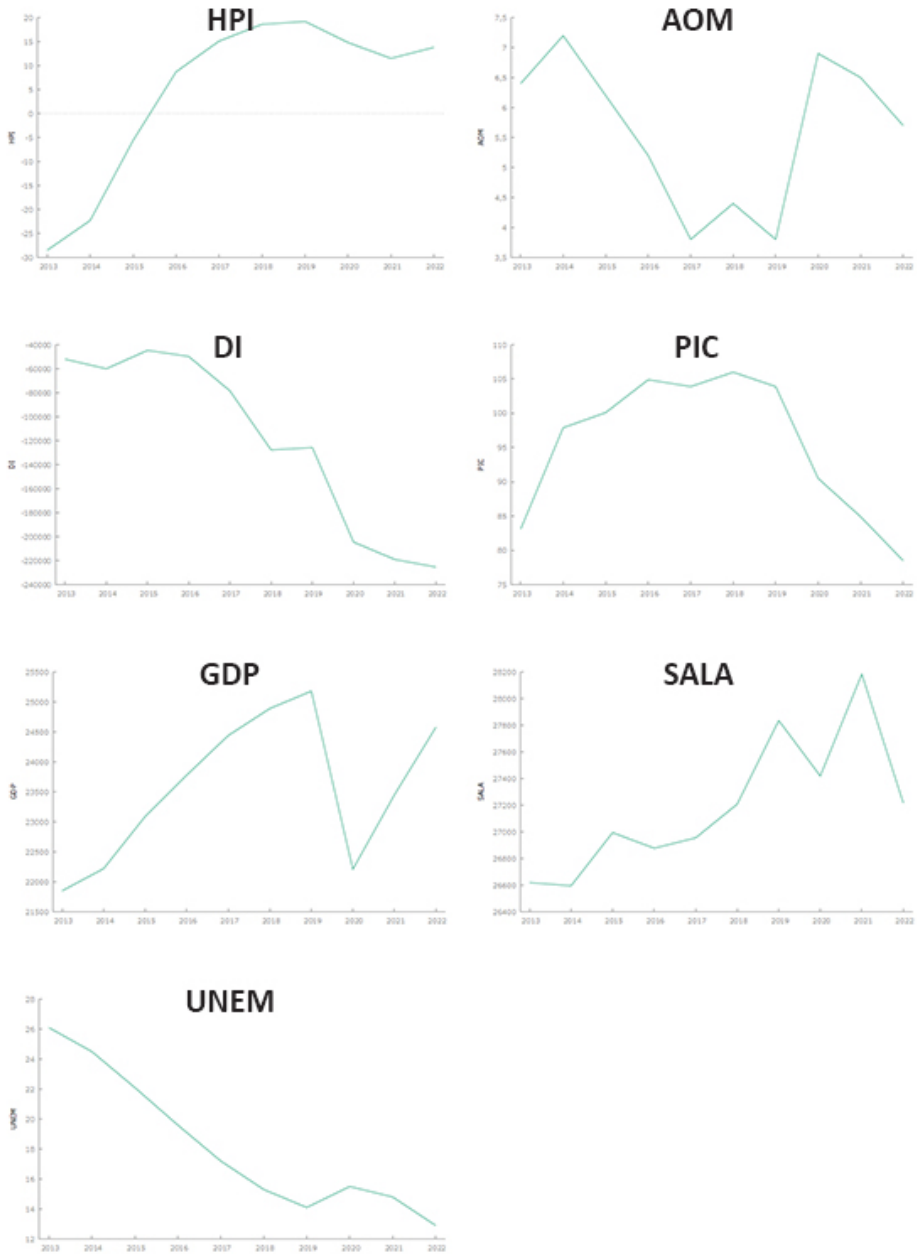
Given that both Croatia and Spain have a significant share of GDP in tourism, for the purposes of this paper we use annual instead of quarterly data due to seasonal fluctuations. From the attached trends, we can see that Croatia and Spain have very similar trends in individual variables. The dependent variable, the house price index, has followed a positive trend since 2015 for Croatia, or since 2014 for Spain. Since 2015, the variable Arrears on mortgage or rent payments has been decreasing in Croatia year by year, while in Spain, the decline started in 2014. The lowest values in Spain were recorded in 2017 and 2019, but there has been a sudden increase since 2020. The Direct investment variable is very similar for Croatia and Spain and it follows a negative trend. Production in construction has followed a positive trend in Croatia since 2015, while in Spain, there has been a sharp decline since 2020. GDP had a positive trend until 2020 when, due to the pandemic, GDP decreased for both countries. Average full-time adjusted salary per employee follows a positive trend for both countries, while the Total unemployment rate is negative, i.e. decreasing.

Table 2 provides a summary of the descriptive statistics for the defined variables in both countries.

Table 2 Descriptive statistics for the defined variables

Variable	CROATIA				SPAIN			
	Mean	Std_dev	Min	Max	Mean	Std_dev	Min	Max
HPI	9.19	15.56	-8.2	32.7	4.53	17.31	-28.5	19.2
AOM	0.98	0.32	0.5	1.6	5.61	1.25	3.8	7.2
DI	-24030.1	4020.42	-29551	-18799	-118722	73563.16	-225376	-44657
PIC	114.89	14.86	100	142	95.37	10.25	78.5	106
GDP	11940	1340.59	10450	14540	23569	1203.27	21850	25180
SALA	15000.4	1321.36	13620	18069	27190.3	508.86	26595	28184
UNEM	11.23	4.42	6.6	17.3	18.21	4.62	12.9	26.1

Source: Created by the authors

For the analyzed period, the average House price index is 9.19 in Croatia and 4.53 in Spain. The maximum achievement in Croatia was 32.7 in 2022 and 19.2 in Spain in 2019, and the minimum achievement of -8.2 was recorded in Croatia in 2015 and -28.5 in Spain in 2013. In the same period, an average for Arrears on mortgage was 0.98 in Croatia and 5.61 in Spain, with a minimum of 0.5 and a maximum of 1.6 in Croatia, and a minimum of 3.8 and a maximum of 7.2 in Spain. Direct investment has a negative mean od -2,4030.1 and -118,722 with a minimum of -29,551 and -225,376 and a maximum of -18,799 and -44,657 for Croatia and Spain, respectively. Production in construction has a higher average for Croatia than for Spain. For Croatia, an average is 114.89, with a minimum of 100 and a maximum of 142, while for Spain, an average is 95.37, with a minimum of 78.5 and a maximum of 106. On the other hand, Spain has a higher GDP and Average full time adjusted salary per employee. GDP for Spain averages 23,569, and for Croatia, the average GDP is 11,940. Average full time adjusted salary per employee is 15,000.4 for Croatia and 27,190.3 for Spain. The total unemployment rate for Croatia is 11.23, and for Spain, it is 18.21.

4. Methodology

The specification of the model examining the impact of macroeconomic variables on the House Price Index in Spain and Croatia is based on economic theory, particularly macroeconomic principles. It employs a multiple linear regression model, using the model averaging technique (Radonjić et al., 2019) based on Freedman's paradox (Freedman, 1983). This model incorporates one or two possible combinations of macroeconomic indicators, with time lags ranging from zero to three. We used a longer lag because the effects of housing price determinants are present over a longer time (Mallick and Mahalik, 2015). Steel (2020) interprets the rapid growth in the use of model averaging methods in economics as an acknowledgment of the significant benefits that can be achieved by adopting

principled approaches as a powerful methodology for addressing uncertainty.

The data used in this research was collected from the Eurostat. The collected data is annual, and the observed time series refer to the period from 2013 to 2022.

After the variables have been defined, it is necessary to set up a mathematical equation of the considered multiple regression model for Croatia (1) and Spain (2), using the least squares method in the following form:

$$Y \text{ (CRO)} = C + \beta1X1 + \beta2X2 + \beta3X3 + \beta4X4 + \beta5X5 + \beta6X6 + \varepsilon i, \quad (1)$$

$$Y \text{ (ESP)} = C + \beta1X1 + \beta2X2 + \beta3X3 + \beta4X4 + \beta5X5 + \beta6X6 + \varepsilon i, \quad (2)$$

where:

Y – the House Price Index,
C – a constant,
X1, X2, X3, X4, X5, X6 – annual values of indicators of independent variables,
β1-6 – values of the associated reactivity coefficients,
εi – residual value.

In addition to setting the mathematical equation of the model, the null and alternative hypotheses have also been formulated:

I. Definition of the null hypothesis according to which no regressor variable in the model is significant: H0: β1= β2= β3=…= βn=0.

II. Alternative hypothesis: There is at least one non-zero variable: HA: βn≠0.

The tests performed on the variables are the multicollinearity and stationarity test, the extended Dickey-Fuller unit root test and OLS regression analysis.

5. Results

The first test performed is the multicollinearity test, that is, the determination of variables whose mutual collinearity is greater than 80%. Based on the multicollinearity test, it was determined that the relevant independent variables for the Croatian model are Arrears on mortgage or rent payments and Total unemployment rate. Arrears on mortgage or rent payments, Total unemployment rate and Production in construction are relevant independent variables for the Spanish model. We examined possible combinations with one and two independent variables for Croatia, and one, two and three independent variables for Spain, with a time lag ranging from zero to four. We assessed 70 models, and out of 70 models, we chose models whose coefficient of determination is over 0.9, so we have five models for Croatia and four models for Spain, as can be seen in Tables 3 and 4.

Table 3 Assessed models for Croatia HPI_CRO

Dependent	Model	# of variables	Variable	Coefficient	StdError	t-ratio	p-value	R²	Adj. R²
HPI	CRO_120	2	AOM	-19.0822	5.61623	-3.398	0.0115	0.9381	0.9205
HPI	CRO_120	2	UNEM	-2.40603	0.23502	-10.24	1.83e-0.5	0.9381	0.9205
HPI	CRO_220	2	log (AOM)	-18.0838	3.30468	-5.472	0.0009	0.9679	0.9588
HPI	CRO_220	2	log (UNEM)	-27.3861	2.56737	-10.67	1.4e-05	0.9679	0.9588
HPI	CRO_420	2	lag (AOM), -1	-6.5805	4.43332	-1.484	0.1883	0.9570	0.9427
HPI	CRO_420	2	lag (UNEM), -1	-3.2105	0.19299	-16.64	3.01e-06	0.9570	0.9427
HPI	CRO_412	1	lag (UNEM), -1	-3.4368	0.26608	-12.92	3.87e-06	0.9468	0.9393
HPI	CRO_512	1	lag (UNEM), -2	-3.31514	0.27568	-12.03	2.01e-05	0.9455	0.9364

Source: Created by the authors

The impact of macroeconomic factors on real estate prices modeled on the House Price Index in Croatia is explained by two single-factor and three double-factor models. The most significant model appears to be the double-factor model with log AOM and

log UNEM, as shown in Table 3. The coefficient for AOM is -18.0838, and for UNEM it is -27.3861, with a p-value of 0.0009 for AOM and 1.4e-05 for UNEM. The coefficient of determination is 96.79%, which makes it a statistically significant model.

Table 4 Assessed models for Spain HPI_ESP

Dependent	Model	# of variables	Variable	Coefficient	StdError	t-ratio	p-value	R²	Adj. R²
HPI	ESP_130	3	AOM	-0.207457	1.42727	-0.1454	0.8892	0.9655	0.9483
HPI	ESP_130	3	PIC	0.477688	0.111138	4.298	0.051	0.9655	0.9483
HPI	ESP_130	3	UNEM	-3.52220	0.154974	-22.73	4.75e-07	0.9655	0.9483
HPI	ESP_121	2	AOM	-3.16420	1.33952	-2.362	0.0502	0.9219	0.8995
HPI	ESP_121	2	UNEM	-3.12114	0.444727	-7.018	0.0002	0.9219	0.8995
HPI	ESP_123	2	PIC	0.492927	0.087944	5.605	0.0008	0.9654	0.9555
HPI	ESP_123	2	UNEM	-3.54895	0.188683	-18.81	2.98e-07	0.9654	0.9555
HPI	ESP_223	2	log (PIC)	54.9915	10.0940	5.448	0.0010	0.9459	0.9304
HPI	ESP_223	2	log (UNEM)	-66.9303	4.75363	-14.08	2.16e-06	0.9459	0.9304

Source: Created by the authors

The impact of macroeconomic factors on real estate prices modeled on the House Price Index in Spain is explained by four double-factor models and one triple-factor model. The most significant model appears to be the double-factor model with PIC and UNEM, as shown in Table 4. The coefficient for PIC is 0.492927, and for UNEM it is -3.54895, with a p-value of 0.0008 for PIC and 2.98e-07 for UNEM. The coefficient of determination is 96.54%, which makes it a statistically significant model.

6. Discussion

The aim of this study is to compare real estate prices between Croatia and Spain with regard to macroeconomic variables. In terms of the Spanish real estate market, in this study, we estimate that the dependent variable, the House Price Index, has followed a positive trend since 2014. The variable Arrears on mortgage or rent payment was declining until 2019, but from 2020 onwards, a sudden increase has been observed. Production in construction in Spain shows a sharp decline since 2020. GDP in Spain showed a positive behavior until 2019, but decreased very sharply in COVID-19 pandemic, in 2020. Currently, Spanish GDP stood at 6.4% and 5.8% in 2021 and 2022, respectively.

The average full-time adjusted salary per employee follows a positive trend, while the total unemployment rate is negative. Thus, our results obtained from Spain in our OLS model that includes HPI= AOM+UNEM+PIC (in Table 1, see macroeconomics variables and symbols) for Spain are relevant. However, if we exclude the Arrears on mortgage or rent payments variable from the Spanish model (HPI = PIC + UNEM), we get better results with the coefficient of determination. The obtained results in this second model point to the fact that real estate price movements are strongly influenced by the behavior of unemployment and Production in construction in Spain. Thus, a downward trend in unemployment would reactivate domestic demand for housing and, together with an increase in construction production costs due to higher raw materials, transportation, and financing costs, would lead to higher prices in the Spanish real estate market.

Croatia and Spain are famous touristic destinations that have a significant share of GDP and show very similar trends of individual variables in our research. The dependent variable, the House Price Index in Croatia has followed a positive trend since 2015, and the variable Arrears on mortgage or rent payments has been decreasing year by year.

The Direct investment variable is very similar for Croatia and Spain and it follows a negative trend. Production in construction has followed a positive trend in Croatia since 2015. GDP recorded a positive trend until 2020 when, due to the pandemic, GDP decreased for both countries. The average full-time adjusted salary per employee follows a positive trend for both countries, while the Total unemployment rate is negative, i.e. decreasing. Thus, our results obtained from Croatia in our OLS model that includes HPI = log AOM + log UNEM (in Table 2, see macroeconomics variables and symbols) for Croatia are relevant. The obtained results in the model point to the fact that real estate price movements are strongly negatively influenced by the behavior of unemployment and the Arrears on mortgage or rent payments, which reduce income for commercial banks, can threaten financial stability, damage a household's future creditworthiness, and reflect the extent of financial difficulties faced by households and the distributional effects within an economy. They may also indicate that households are unable (or unwilling) to move to more affordable dwellings. Arrears are thus associated with diminished labor mobility, which may reduce the speed at which an economy recovers from a recession (for more details, see Gerlach-Kristen & Lyons, 2015). Buiter (2010) also argues that changes in house price impact consumption, as real estate can be used as collateral for consumer loans. Housing prices are determined by income per capita real interest rates and credit availability (Egert & Mihaljek, 2007). Grum and Govekar (2016) show that higher unemployment is associated with lower residential property prices in Greece, France, Poland and Norway. This in the line with the results of our research in Croatia and Spain. Both countries are EU countries that struggled with the 2008 crisis. Croatia, like other transitional countries still has many economic challenges (see Radonjić et al. (2019) for Montenegro, Stanković (2022) for Bosnia and Herzegovina). We proved that production in construction has a positive impact on real estate prices in Spain, as expected and claimed before (see Ma et al., 2017).

7. Conclusion

The study analyzes Croatia and Spain in the light of real estate prices and measures the impact of various macroeconomic variables on the volatility of those prices. Our research applied the model aver-

aging technique to time-lagged variables in various combinations, following similar studies referenced in the literature. In the first step, we included variables such as Arrears on mortgage or rent payments, Direct investment, Production in construction, Real GDP per capita, Average full-time adjusted salary per employee and Total unemployment rate. After the multicollinearity test, we conclude that the independent variables relevant for the Croatian model are Arrears on mortgage or rent payments and Total unemployment rate. Independent variables Arrears on mortgage or rent payments, Total unemployment rate and Production in construction are relevant for the Spanish model. After applying more than 70 models, we found that the unemployment rate has a strong negative impact on house prices in Croatia and Spain, a finding that is corroborated by numerous other researchers and across various countries.

Our results are in the line with other research studies and most of them confirmed that when unemployment is high in a country, the real estate industry experiences a downturn, resulting in less construction activity, fewer loans granted to businesses and households, and an overall stifling of economic development. Production in construction in Spain shows a sharp decline since 2020 and confirms our assumptions that growth of the real estate industry is strongly influenced by the expansion of the construction industry. Arrears on mortgage or rent payments should have a negative influence on real estate prices, which is confirmed for Croatia in our research. Delays in payments lead to non-performing loans, prompting financial institutions to exercise greater caution when approving loans. This, in turn, reduces demand for real estate, causing real estate prices to decline.

At the end of 2023, real estate prices in both Croatia and Spain increased significantly. Both countries are strongly tourist-oriented. Spain had a strong rise in prices when the Euro was introduced back in 2002. Since then, there has been significant growth in real estate prices, excluding the years of global crisis in 2008 and 2009. Croatia, on the other hand, recorded a substantial increase in real estate prices, which was further intensified by its entry into the Eurozone in 2023. Inflationary pressures, present in all EU countries, also contributed to this trend.

However, several limitations warrant consideration. Firstly, the reliance on historical data may restrict the ability of the study to capture sudden

market shifts or structural changes. Additionally, the focus on macroeconomic variables may overlook nuanced local dynamics or regional variations within Croatia and Spain, potentially limiting the generalizability of findings. Recommendations for further research suggest exploring a broader range of Southeast European countries with similar characteristics of strong tourism sectors. Such studies could offer deeper insights into the interplay between tourism, macroeconomic variables, and real estate dynamics, thus facilitating more comprehensive policymaking and market forecasting in the region. Additionally, exploring the long-term implications of economic integration, such as Croatia's entry into the Eurozone, on real estate markets warrants further investigation to inform future decision-making processes. Further research could analyze some additional factors, such as geopolitical tensions, global economic trends, cultural differences, regulatory measures, digital platform industry, environmental sustainability, consumer behavior and risk propensity in investment that could influence the observed dynamics of real estate prices in Croatia and Spain.

References

1. Adams, Z. & Fuss, R. (2010). Macroeconomic Determinants of International Housing Markets. *Journal of Housing Markets*, *19*(1), 38-50. https://doi.org/10.1016/j.jhe.2009.10.005

2. Álvarez-Román, L. & Garcia-Posada, M. (2021). Are house prices overvalued in Spain? A regional approach. *Economic Modelling*, *99*, 105499. https://doi.org/10.1016/j.econmod.2021.105499

3. Allen, L., Barnes, C. & Lynch, H. (2016). The Role of Macroeconomic Fundamentals in the Determination of House Prices. *Journal of Applied Business and Economics*, *18*(5), 42-48.

4. Buiter, W. H. (2010). Housing wealth isn't wealth. *Economics*, *4*(1), 1-29. https://doi.org/10.5018/economics-ejournal.ja.2010-22

5. Chin, W. H. (2003). *Macro-economic Factors Affecting Office Rental Values in Southeast Asian Cities: Case of Singapore, Hong Kong, Taipei, K. Lumpur and Bangkok.* PRESS Brisbane.

6. Dumičić, K., Čeh Časni, A. & Šprajaček, P. (2012). The linkage between real estate prices and the macro-economy in Croatia: VAR approach. *Croatian Operational Research Review*, *3*(1), 289-299.

7. Egert, B. & Mihaljek, D. (2007). Determinants of house prices in Central and Eastern Europe. *Comparative Economic Studies*, *49*(3), 337-489. https://doi.org/10.1057/palgrave.ces.8100221

8. Freedman, D. A. (1983). A note on screening regression equations. *The American Statistician*, *37*(2), 152-155. https://doi.org/10.1080/00031305.1983.10482729

9. Gaspareniene, L., Remeikiene, R. & Skuka, A. (2017). Assessment of the impact of macroeconomic factors on housing price level: Lithuanian case. *Intellectual Economics*, *10*(2), 122-127. https://doi.org/10.1016/j.intele.2017.03.005

10. Gerlach-Kristen P. & Lyons S. (2015). *Mortgage arrears in Europe: The impact of monetary and macroprudential policies* (SNB Working Paper No. 2015-05). Zurich: Swiss National Bank.

11. Ghirelli, C., Leiva-León, D. & Urtasun, A. (2023). Housing prices in Spain: convergence or decoupling?. *SERIEs*, *14*(2), 165-187. https://doi.org/10.1007/s13209-023-00275-1

12. Giussani, B., Hsia, M. & Tsalocos, S. (1993). A Comparative Analysis of the Major Determinants of Office Rental Values in Europe. *Journal of Property Valuation and Investment*, *2*(11), 157-173. https://doi.org/10.1108/14635789310031487

13. Grum, B. & Govekar, D. K. (2016). Influence of Macroeconomic Factors on Prices of Real Estate in Various Cultural Environments: Case of Slovenia, Greece, France, Poland and Norway. *Procedia Economics and Finance*, *39*, 597-604. https://doi.org/10.1016/S2212-5671(16)30304-5

14. Hon-Chung, H. (2009). The Impact of Property Market Developments on the Real Economy of Malaysia. *International Research Journal of Finance and Economics*, *30*, 66-86.

15. Kotseva, P. & Yanchev, M. (2017). *Analysis of the Housing Market Developments and the Underlying Macroeconomic Fundamentals in Bulgaria* (BNB Discussion Paper No. 103). Sofia: Bulgarian National Bank.

16. Lourenco, R. F. & Rodrigues, P. M. (2014). *The dynamics and contrast of house prices in Portugal and Spain*. Economic Bulletin, Banco de Portugal. https://www.bportugal.pt/node/43818

17. Lovrinčević, Ž. & Vižek, M. (2008). House price determinants in the Republic of Croatia and potential effects of housing market liberalization. *Economic Review*, *59*(12) 723-740.

18. Ma, L., Liu, C. & Reed, R. (2017). The impacts of residential construction and property prices on residential construction outputs: an inter-market equilibrium approach. *International Journal of Strategic Property Management*, *21*(3), 296-306. https://doi.org/10.3846/1648715X.2016.1255675

19. Mallick, H. & Mahalik, M. K. (2015). Factors determining regional housing prices: evidence from major cities in India. *Journal of Property Research*, *32*(2), 123-146. https://doi.org/10.1080/09599916.2014.963642

20. Mavrodiy, A. (2005). *Factor Analysis of Real Estate Prices*. [Master thesis, National University "Kyiv-Mohyla Academy"]. National University "Kyiv-Mohyla Academy".

21. Pashardes, P. & Savva, C. S. (2009). Factors Affecting House Prices in Cyprus: 1988-2008. *Cyprus Economic Policy Review*, *3*(1), 3-25.

22. Radonjić, M., Đurišić, V., Rogić, S. & Đurović, A. (2019). The impact of macroeconomic factors on real estate prices: Evidence from Montenegro. *Ekonomski pregled*, *70*(4), 603-626. https://doi.org/10.32910/ep.70.4.2

23. Sivitanides, P. (2015). Macroeconomic Influences on Cyprus House Prices: 2006Q1-2014Q2. *Cyprus Economic Policy Review*, *9*(1), 3-22.

24. Stanković, D. (2022). Analysis of the impact of macroeconomic factors on real estate prices in Bosnia and Herzegovina. *Economics*, *10*(2),147-160. https://doi.org/10.2478/eoik-2022-0015

25. Steel, M. F. (2020). Model averaging and its use in economics. *Journal of Economic Literature*, *58*(3), 644-719. https://doi.org/10.1257/jel.20191385

26. Tupenaite, L., Kanapeckiene, L. & Naimaviciene, J. (2017). Determinants of Housing Market Fluctuations: Case Study of Lithuania. *Procedia Engineering*, *172*, 1169-1175. https://doi.org/10.1016/j.proeng.2017.02.136

27. Weiner, D. & Fuerst, F. (2015). *The Dynamics of House Prices in Israel* (University of Cambridge Working Paper No. 2015-12). Cambridge: University of Cambridge.

28. Zhenyu, S. & Taltavull, P. (2020). International capital movement towards the Spanish real estate sector. *Journal of Property Investment & Finance*, *38*(2), 107-127. https://doi.org/10.1108/JPIF-05-2019-0067

Josip Arnerić
University of Zagreb
Faculty of Economics and Business
10000 Zagreb, Croatia
jarneric@net.efzg.hr

Luka Osojnik
University of Zagreb
Faculty of Economics and Business
10000 Zagreb, Croatia
losojnik@net.efzg.hr

EXPLANATORY FACTORS OF EUROZONE EQUITY ETF TRACKING ERROR

Abstract

Purpose: According to the existing literature, it remains unclear whether a specific ETF outperforms or underperforms its benchmark index in terms of tracking error, especially during crisis periods. Therefore, this study concentrates on the largest and most liquid Eurozone equity ETF, iShares Euro Stoxx 50, which tracks the Euro STOXX 50 index, with the fundamental objective of identifying explanatory factors of tracking errors during crisis periods, encompassing the COVID-19 pandemic and the onset of the Ukrainian war.

Methodology: The added value of current research lies in the utilization of Markov regime switching regression with two-state variables. This approach is supported by the idea that the influence of explanatory factors on tracking error may vary between bearish and bullish regimes, which typically align with non-crisis and crisis periods, respectively.

Results: Empirical findings indicate that an increase in volatility led to a stronger decrease in tracking error during periods of market stress than in a bullish regime, while a negative impact of illiquidity on tracking error is similar for both regimes. Unlike a single-regime model, Markov switching exhibits a negative relationship between the net flows and tracking error, as expected. The effect of premium/discount seems to be both positive and negative, but a weaker influence was found during a bearish regime due to herding behavior of investors or higher trading costs.

Conclusion: This study relies on an ex-ante approach with its main advantage of providing a forward-looking estimate of tracking error that takes into account changes in market conditions and the ETF's underlying holdings, unlike historical or ex-post tracking error.

Keywords: ETF tracking error, Eurozone equity market, crisis, Markov switching regression

1. Introduction

Active investing in a highly efficient market requires a complex understanding of available information, which can be challenging and time-consuming for an average investor. Fortunately, the evolution of technology has revolutionized the investment landscape, enabling investors to efficiently allocate their assets into the stock market without the constant need to interpret current market information and react accordingly. This paradigm shift has given rise to the concept of passive investing, where investors seek to replicate the performance of an index or a specific set of assets. Thus, exchange-traded funds (ETFs) have become increasingly popular passive investments in Europe in recent years (Le Sourd & Safaee, 2021). Beginning with a modest $100 billion in assets under management in 2000, the ETF market skyrocketed to $1 trillion by 2010, and almost achieved a $10 trillion market cap in 2020. In

contrast, US-based mutual funds, which have been present in the market for nearly a century, held approximately $27 trillion in assets as of 2020. One of the reasons for such an extraordinary increase in popularity are the benefits it offered. Some of these are minimal fees associated with owning and trading securities, high liquidity while managing a large basket of stocks, diversification benefits and simple tradability (Mussavian & Hirsch, 2002; Madhavan, 2014). As a result, measuring the performance of ETFs with respect to tracking error has attracted great attention among scientists and practitioners (Johnson, 2009; Dorocáková, 2017; Tsalikis & Papadopoulos, 2019; Feder-Sempach & Miziołek, 2023). It measures how closely an ETF tracks its benchmark index. A low tracking error indicates the ability of an ETF to replicate its target portfolio or the target index almost perfectly, while a high tracking error suggests that the ETF deviates from its benchmark. Several studies attempted to examine the performance of European ETFs with respect to tracking error and identify the most influenced factors. Feder-Sempach & Miziołek (2023) concluded that tracking error of eurozone ETFs over a period of ten years was relatively low, on average 0.3%, although ETFs with distributing income had a higher tracking error against accumulation income ETFs.

However, according to the existing literature, it is still unclear if specific ETF outperforms or underperforms its benchmark index in terms of tracking error, particularly in a crisis period. In general, during a crisis period, market volatility tends to increase, which can make it more difficult for an ETF to perfectly track its underlying benchmark (Johnson et al., 2013). This is because the prices of the ETF's constituent securities may fluctuate more widely than usual, and the ETF's portfolio manager may not be able to rebalance the portfolio efficiently due to market conditions such as liquidity constraints and transaction costs. Additionally, during a crisis period, there may be significant changes in the composition of the underlying benchmark as some securities may become more or less important to the benchmark compared to the ETF's holdings, which can also contribute to tracking error (Vardharaj et al., 2004). By contrast, in some cases, the tracking error of an ETF may decrease during a crisis period if the ETF's portfolio manager is able to identify undervalued securities and add them to the portfolio. A decreased tracking error implies that ETF outperforms the benchmark index and *vice versa*.

In that context, the fundamental objective of the current study is to find out if the tracking error of the Eurozone ETF relative to its benchmark index decreased or increased during crisis periods, specifically covering COVID-19 pandemic and the Ukrainian war. Additionally, it offers empirical evidence for widely used market-based measures that may influence tracking error, such as market volatility, illiquidity proxy (bid-ask spread), net flow, premium or discount and trading volume. For the same reason, daily observations from May 27, 2019 to May 26, 2023 are provided by Refinitiv Eikon service[1]. According to different market regimes, the influence of the aforementioned factors on tracking error might be distinctive, and thus this paper tries not only to provide empirical evidence and comprehensive explanations of those distinctive influences, but also to fill the gap in eurozone ETF performance analysis with respect to regime switching methodology.

This study relies on an ex-ante approach with its main advantage of providing a forward-looking estimate of tracking error that takes into account changes in market conditions and the ETF's underlying holdings, unlike historical or ex-post tracking error. It uses the root of the squared residual from a simple regression of net asset value (NAV) returns on the benchmark returns as an indicator of a daily tracking error. Entire research focuses on the Eurozone equity ETF, which tracks the Euro STOXX 50 index. The STOXX 50 index is a widely followed benchmark for the Eurozone equity market, representing the performance of fifty blue-chip companies from 18 Eurozone countries. There are several exchange-traded funds that track the STOXX 50 index, but the most popular one is the iShares Euro Stoxx 50 ETF as the largest and most liquid one, with over 10 billion USD in assets under management as of May 2023. This ETF is particularly attractive to investors who want to reinvest the dividends into the fund, and not to pay them out, which is unique to accumulating income ETFs. Unlike distributing income ETFs, the accumulating ones maximize future returns. This was yet another reason, except the size and its liquidity, for selecting iShares Euro Stoxx 50 ETF.

Added value of the research consists of the employed methodology considering regime switching regression with two-state variables that follow the Markov chain. The major reason for this is that the

---

influence of tracking error explaining factors may deviate over time depending on whether bearish or bullish state is in the market. Those two market regimes usually coincide with non-crisis and crisis periods, and according to the span of time-series data, it captures the COVID-19 pandemic and the Ukrainian war. This is particularly important as the prices of securities tend to fluctuate more during times of stress and market turmoil, which can hinder the ability of the portfolio manager to efficiently manage portfolios of the constituent ETF. The inability of the portfolio manager to balance portfolios strays the price of the ETF away from its benchmark index and hence decreases the ETF tracking error. Conversely, actively managed ETFs aiming to outperform their benchmark index may experience a decrease in tracking error during stressful periods. This occurs when the portfolio manager successfully identifies undervalued securities and incorporates them into the portfolio, leading to superior performance compared to the index. In addition, a Markov switching approach is appropriate for nonlinear time-series models with regimes determined by unobserved states which must be inferred from the data and the parameters are expanded to include the transition probabilities. Unlike current literature, which mainly revolves around the American market and the ETFs, this paper shifts attention and contributes to the literature which focuses on the European market.

The rest of the paper is organized as follows. Section 2 explains the theoretical framework of ETFs, including a review of previous studies. Section 3 presents data and methodology. Section 4 provides empirical findings, while Section 5 offers a comprehensive discussion of the results. Finally, Section 6 provides a conclusion.

2. Theoretical framework and a review of previous studies

Both mutual funds and ETFs serve to reduce risk through diversification. However, there are key distinctions between the two. ETFs are pooled investment vehicles that track specific indexes, mostly passively managed. With a focus on risk reduction rather than returns, ETFs have seen a significant increase in total market cap in the last two decades. This simplicity makes them attractive to retail investors. By contrast, mutual funds lack the ability for investors to sell shares at any time and are actively managed. The research conducted by Kaminsky (2001) suggests that mutual funds play a significant role among institutional investors as the primary channel for financial flows into emerging markets. On the other hand, the findings of Sy and Ong (2004) indicate that this phenomenon is more pronounced in the European market compared to the United States. This active management of the fund comes with greater initial and management costs as well as higher transaction costs compared to ETFs which drive retail investors toward cheaper alternatives. Broman (2016) says that higher liquidity ETF shares attract investors who are not willing to invest directly into illiquid assets such as commodities, emerging markets, etc. When buying ETFs from the broker, one will find that there are many different options of an ETF that they want to buy from different markets. For example, an ETF from the London stock exchange will show prices in pounds, whereas on some exchanges there will be a limited number of ETFs available to be bought. Investors also want to avoid buying from multiple exchanges as there are annual charging fees. One important factor to consider is the volume that certain instruments have on different exchanges. This is particularly important when buying ETFs because of their tracking error. If the volume and trading frequency of the ETF is low, it can exhibit higher spreads and therefore increase tracking error. Therefore, it is wise to choose a stock exchange with higher trading volume compared to other exchanges.

There are several approaches to obtaining ETF tracking error (a historical approach, an ex-post or an ex-ante approach), each with its own pros and cons (De Rossi, 2015). Most studies have used regression analysis after obtaining tracking error, indicating that European ETFs generally exhibit good performance in terms of tracking their benchmarks. However, there is significant variation depending on the specific ETF under analysis, the measurement of tracking error, the observed period, and the approaches employed to examine the factors influencing tracking error. Understanding these factors can assist investors in making informed decisions when selecting ETFs for their portfolios. Because of the nature of ETFs and their mechanics, the movements of tracking error can be expected to move in directions which align with the theory behind the variables. The deviation of ETF prices from their NAV is primarily maintained through the arbitrage process. Theoretically, an increase in pre-

mium/discount should invite arbitrageurs, which in turn should align ETF prices better with its NAV. Moreover, ETFs that include international stocks are expected to exhibit greater deviation due to the continued trading of their shares on the domestic exchange, while the market for the underlying securities in the creation basket is closed. Similarly, in theory, ETFs containing illiquid securities should experience higher deviations as the arbitrage process would require a larger deviation to compensate for the higher transaction costs associated with trading those less liquid securities. Increased trading volume positively impacts liquidity and bid-ask spread, while market volatility increases bid-ask spreads. Net inflows and their effect on the tracking error highly depend on the state of the market as market participants have different behaviors during each regime. To expand the theory, the examination also includes a review of what the literature suggests about variable relationships. ETFs which invest in less liquid assets may experience difficulty in replicating their benchmark index leading to higher tracking errors. Bae & Kim (2020) have documented a positive relation between illiquidity and tracking error.

Hillard & Le (2022) found that emerging European markets have higher tracking error in comparison to developed Europe, and it was around 0.67% and 0.33%, respectively. Rompotis (2011) states that tracking error for ETFs with higher expense ratios was higher. Additionally, Tsalikis & Papadopoulos (2019) confirmed that tracking error for European ETFs was, on average, higher than that of US ETFs, while a possible explanation for the aforementioned could lie in the economies of scale and thus lower costs. Chu (2011) also found that economies of scale will improve tracking ability, while their research suggests that expense, delay in receiving dividends, the trading cost and the market risk increase tracking error. Additionally, Elton et al. (2019), as well as Chu & Xu (2021), suggested that tracking error is significantly influenced by delayed reinvestments of dividend. Regardless of the tracking error measurement, higher assets under management (AUM) positively affect tracking ability. The study also found that higher expense ratios are associated with higher tracking errors, although statistical significance is observed only for one measurement. Another study by Frino & Gallagher (2001) presented evidence that tracking error is positively and significantly correlated with dividend payments, and also

that there were seasonal patterns with higher error rates in January and May, and a lower error rate in the quarters ending in March, June, September, and December. Aber et al. (2009) stated that the range of daily price fluctuations was very large, which indicated that active traders or arbitrageurs were more likely to profit than passive traders. Blitz et al. (2012) revealed in their study that index funds and ETFs in Europe underperform their benchmarks by larger amounts than their reported expenses, with dividend taxes explaining a significant portion of underperformance. This highlights the need to account for dividend taxes in evaluating fund performance and measuring fund costs accurately.

Other well-known factors explaining ETFs tracking error are market volatility, trading volume, the net flow as well as premium or discount. Higher market volatility and trading volumes can lead to wider bid-ask spreads, which can increase the cost of trading and result in higher tracking errors, as observed in several studies, including Ben-David et al. (2019). In a study on Hong-Kong ETFs, Chu (2011) demonstrated that trading volume increases tracking error; however, it is not significant, while Yiannaki (2015) suggested that there is a weak correlation between tracking error and trading volumes. Dorocáková (2017) found that fluctuations in the underlying index can have a relative influence on tracking error. In the case of bid-ask spreads, Meinhardt et al. (2015) indicated a positive relation to tracking error for the German ETF market.

On the demand side, the net flow ETF may affect its tracking error. The ETF net flow tends to increase during a bulish period when investors are more optimistic and confident about the future of financial markets. Conversely, during a bearish period in the market, the ETF net flow tends to decrease as investors become more risk-averse and seek to reduce their exposure to equities. According to research of Ben-David et al. (2017), tracking error is negatively related to the ETF net flow. Another study by Osterhoff, & Kaserer (2016) confirmed that the net flow had a significant negative effect on tracking error for small ETFs.

Divergence of ETF market prices from their net asset value, reported as premium (or discount), is yet another explaining factor of ETF tracking error. A study by Wong & Shum (2010) found that tracking error of the examined ETFs is consistently positive in both bullish and bearish markets. This suggests that investors are willing to pay a premium for ETF

investments, as ETFs provide positive returns that can cover transaction costs and potentially yield returns in different market conditions. Rompotis (2010) found tracking error to be positively affected by premium/discount, while Li and Zhao (2014) found that premiums can lead to increased tracking error in ETFs that hold illiquid securities.

Aber et al. (2009) suggested that ETFs traded more at a premium than at a discount, indicating that the market tended to overvalue ETFs compared to their NAV. Additionally, premiums have shown to be higher for newly created ETFs, as documented in a study by Piccotti (2018), which indicates that investors are willing to pay a premium in order to access the liquidity benefits provided by ETFs, which allow indirect availability to less accessible underlying securities.

3. Data and methodology

In previous studies, researchers have used both NAV returns and closing market price returns to evaluate the tracking error of ETFs compared to their benchmark index returns (Zawadzki, 2020). However, due to its advantages, the NAV-based measurement of tracking error is widely recognized as the preferred approach. NAV returns consider dividends or any income generated by the underlying assets, providing a more accurate and reliable measure of the ETF's performance in accordance with GIPS - Global Investment Performance Stand-

ards (CFA, 2020). Additionally, changes in the net asset value reflect what an investor would actually receive from holding the ETF (Osojnik, 2023). In contrast, closing price ETF returns may be influenced by short-term price fluctuations that do not necessarily reflect the underlying performance of the ETF. Therefore, using closing price returns to assess tracking error can be misleading. Furthermore, the difference between ETF market prices and their respective net asset values introduces another variable known as the premium or discount. This variable will be utilized to explain ETF tracking error. Mispricing of an ETF in relation to its net asset value creates arbitrage opportunities through the creation and redemption mechanism, which can be advantageous for investors (Osojnik, 2023).

The first impression of tracking error can be made by visual inspection of iShares Euro Stoxx 50 ETF net asset values and market closing prices of a benchmark Euro STOXX 50. Figure 1 uses a dual scale axis for comparison and highlights the shaded area covering turbulent periods of the COVID-19 crisis and the Ukrainian war. Both net asset values and closing prices are expressed in the same currency (EUR), but with different scales. Although Figure 1 clearly indicates that the ETF tracks its benchmark quite well with few disparities during a bullish regime, commenting on the ETF performance solely based on price differences is not possible; instead, log returns are considered.

Figure 1 ETF net asset values vs. benchmark index prices

Source: Authors' construction using data provided by Refinitiv Eikon

Before analysis continues, all variables of interest are derived from the raw data. Firstly, tracking error is estimated following the ex-ante approach by regressing NAV returns of the ETF (Ret_t^{NAV}) on benchmark returns (Ret_t^{BEN}). The root of squared regression residual for each trading day resulted in tracking error (Osojnik, 2023):

$$track\_error_t = \sqrt{(Ret_t^{NAV} - 0.0109 - 0.9949 Ret_t^{BEN})^2} . \ (1)$$

In the above expression, -0.0109 and -0.9949 are the constant term and the slope coefficient, respectively.

Daily NAV returns of ETF and benchmark returns, used in the regression, are obtained following the same formulation:

$$Ret_t^{NAV} = \frac{NAV_t - NAV_{t-1}}{NAV_{t-1}} 100\%$$

$$Ret_t^{BEN} = \frac{C_t^{BEN} - C_{t-1}^{BEN}}{C_{t-1}^{BEN}} 100\%, \quad (2)$$

where NAV_t and NAV_{t-1} are ETFs net asset values on the current and previous trading day, while C_t^{BEN} and C_{t-1}^{BEN} are closing prices of a benchmark index.

Next, an illiquidity proxy measure is obtained as bid-ask spread with end-of-day ETF quotes toward its mid-quote, by the following expression:

$$illiquidity_t = \frac{2(A_t^{ETF} - B_t^{ETF})}{A_t^{ETF} + B_t^{ETF}} 100\%. \quad (3)$$

ETF daily premium is also expressed as a percentage like all other variables according to:

$$premium_t = \frac{C_t^{ETF} - NAV_{t-1}}{NAV_{t-1}} 100\%, \quad (4)$$

where C_t^{ETF} are closing (market) prices of the ETF on day t. The same indicator (4) exhibits a discount (negative values) when the ETF market price is lower than its NAV. This means that investors buy the ETF at a price cheaper than the underlying value of its assets. Conversely, an ETF is trading at a premium when its market price is higher than its NAV (positive values of the aforementioned indicator).

The daily ETF net flow, which represents the inflow and outflow of ETF, is given by the formula:

$$net\_flow_t = \frac{TNAV_t - \left(1 + \frac{Ret_{t-1}^{ETF}}{100}\right) TNAV_{t-1}}{TNAV_{t-1}} 100\%, \quad (5)$$

where the total net asset value $TNAV_t$ represents a product of NAV per share and the number of outstanding shares on the current day. The previous day total net asset value $TNAV_{t-1}$ is adjusted with the respective ETF daily return to account for the performance effect of the change, which is independent of capital flows.

Market volatility is measured by the official Euro Stoxx 50 volatility index (VSTOXX), which is the European version of VIX, reflecting investor's sentiment as expectations of future volatility.

Summary statistics of variables of interest are reported in Table 1. All values of variables are expressed in percentages, except the volatility index and trading volume. Only trading volume is transformed into logs due to a large scale and extreme variations of trading across days.

Table 1 Descriptive statistics of iShares Euro Stoxx 50 ETF tracking error and its explanatory predictors along with ADF unit root test

Variable	Min	Max	Mean	SD	Median	ADF test
Tracking error	0.00	1.08	0.07	0.09	0.05	-15.1558***
Illiquidity proxy	0.01	0.48	0.06	0.04	0.05	-11.5989***
Volatility index	10.69	85.62	23.21	8.85	21.36	-13.8413***
Net flow	-5.74	3.59	-0.03	0.48	-0.01	-20.9258***
Premium/discount	-3.55	2.19	0.04	0.21	0.05	-18.7641***
Logs of volume	7.37	13.56	9.84	0.85	9.82	-13.4245***

Note: significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Source: Authors' calculation using data provided by Refinitiv Eikon

It can be noticed in Table 1 that the mean and median tracking errors are 0.07% and 0.05%, respectively. The maximum value of 1.08% can be expected during high market volatility and stress, when assets in the portfolio become less liquid and more difficult to allocate. The null hypothesis of the Augmented Dickey-Fuller (ADF) unit root test is rejected at the significance level of 1%, indicating that all considered variables are stationary. ADF in the levels is performed without trend and without drift, except for the net flow, and the premium/discount as their mean is approximately zero, and thus a drift term is not omitted for those two variables (Osojnik, 2023). Stationarity of all variables is preferred to eliminate possible suspicion of the results in the post-estimation phase caused by the non-stationarity issue.

Figure 2 shows the clustering of tracking error, volatility and illiquidity, particularly in crisis periods which can be identified as bearish states of the ETF regime. Therefore, it is not surprising that these three variables are more correlated than other variables, indicating that illiquidity and volatility contribute positively to tracking error (Figure 3).

Figure 2 Time-series of variables observed from May 27, 2019 to May 26, 2023
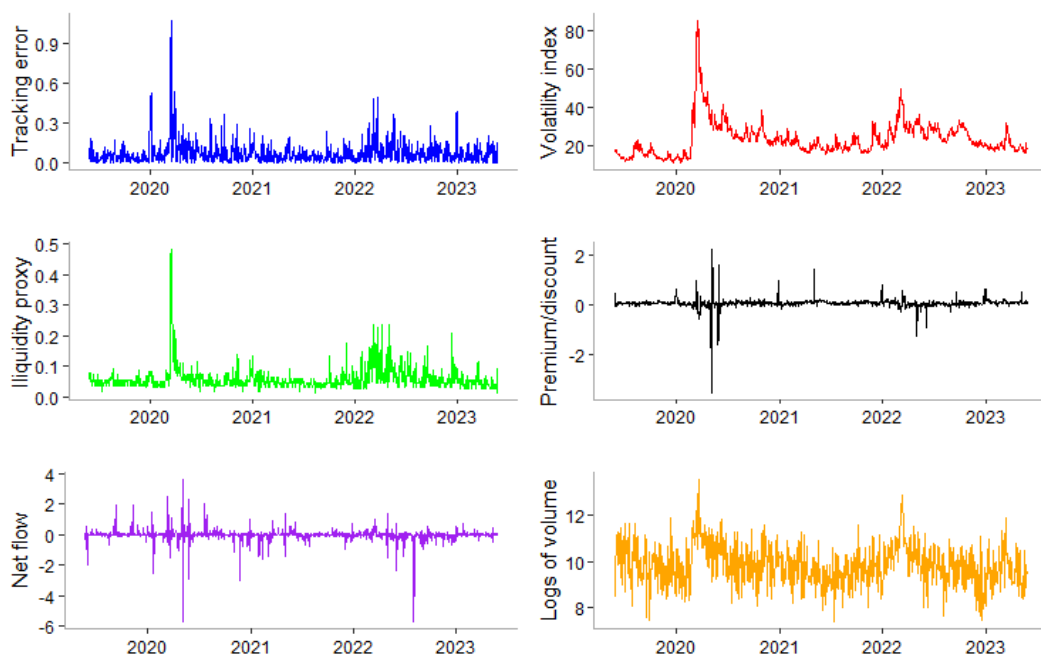
Source: Authors' construction using data provided by Refinitiv Eikon

Figure 3 Correlation matrix between the observed variables
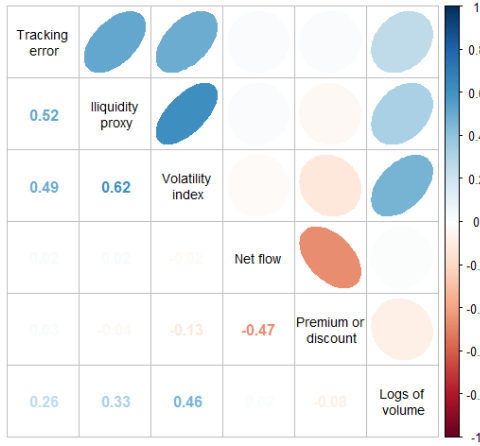
Source: Authors' construction using data provided by Refinitiv Eikon

An application of Markov regime-switching (MRS) models has attracted great interest in capturing dynamics of financial time-series, primarily due to the nonlinear dependence between considered variables as well as their nonstationary property (time-varying moments). In these circumstances, the main advantage of MRS is that it allows regression parameters to switch across multiple states or regimes, with the probabilities of switching between these states being dependent on the current state (Peovski et al., 2022). For example, the MRS model can capture changes of the dependence between two or more variables during different economic cycles or market regimes, such as high and low volatility regimes or bearish and bullish regimes, which usually coincides with crisis and non-crisis periods (Osojnik, 2023).

A simple Markov switching bivariate regression model, which considers two states of regime, can be formalized as follows:

$$y_t = \alpha_{St} + \beta_{St} \cdot x_{St} + u_{St}$$

$$u_{St} \sim WN(0, \sigma_{St}^2)$$

$$\alpha_{St} = \alpha_1(2 - S_t) + \alpha_2(S_t - 1) \qquad (6)$$

$$\beta_{St} = \beta_1(2 - S_t) + \beta_2(S_t - 1)$$

$$\sigma_{St}^2 = \sigma_1^2(2 - S_t) + \sigma_2^2(S_t - 1),$$

where $S_t = j$ is a discrete state variable that indicates in which regime the Markov process is $j = 1, 2, ..., k$. Consequently, if the process is in the first regime state, then $S_t = 1$, with parameters α_1, β_1 and σ_1^2, but if the process is in the second regime state, then $S_t = 2$, with parameters α_2, β_2 and σ_2^2. Assuming that the conditional probability density function is Gaussian:

$$f(S_t) = \frac{1}{\sqrt{2\pi\sigma_{St}^2}} exp\left\{-\frac{(y_t - \alpha_{St} - \beta_{St} \cdot x_{St})^2}{2\sigma_{St}^2}\right\}, \qquad (7)$$

then a log-likelihood function $lnL = \sum_{t=1}^{T} ln\{f(S_t)\}$ can be maximized with respect to parameters α_1, α_2, β_1, β_2, σ_1^2 and σ_2^2. However, the state variable is usually unobserved in practical applications, but it is commonly assumed that it follows a Markov chain process with a k-dimensional state space (Hamilton, 1989). The specificity of a Markov chain process is the first-order dependence, implying that a state variable at the moment t depends only on the previous state of the process at the moment t-1 (Goldfeld & Quandt, 1973). Thus, for k=2, the log-likelihood function takes the form:

$$lnL = \sum_{t=1}^{T} ln\left[\left(\sum_{j=1}^{2} \frac{1}{\sqrt{2\pi\sigma_{St}^2}} exp\left\{-\frac{(y_t - \alpha_{St} - \beta_{St} \cdot x_{St})^2}{2\sigma_{St}^2}\right\}\right) Pr(S_t = j | I_{t-1})\right] \cdot (8)$$

The probability density function (8) for each observation $t = 1, 2, ..., T$ is presented as a weighted sum of conditional probability density functions for both regime states $j = 1, 2$. The associated weights $Pr(S_t = j | I_{t-1})$ are interpreted as conditional probabilities that the process is in the state j at the moment t, conditioned on all information from previous periods up to and including the moment $t-1$. These conditional probabilities are called ex ante probabilities (Kim & Nelson, 2017). In order to maximize the log-likelihood function it is necessary to assume a priori the behavior of a discrete state variable S_t. It is assumed that the state variable is generated by a first-order Markov process:

$$Pr(S_{t-1}, S_{t-2}, ..., S_1, I_{t-1}) = Pr(S_t | S_{t-1}). \qquad (9)$$

Ex ante probabilities $Pr(S_t = j | I_{t-1})$ are generated by a matrix of transitional probabilities, the so-called stochastic matrix:

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} p & (1-p) \\ (1-q) & q \end{bmatrix}. \qquad (10)$$

The matrix of transition probabilities P is an irreducible and primitive matrix (Hamilton, 1989). This

means that all states of the Markov chain communicate with each other, i.e. that there is a probability of transition from state *i* to state *j*, as well as a probability of transition from state *j* to state *i*. Therefore, it is assumed that all elements of the stochastic matrix are greater than zero (a primitive matrix). In the matrix P the probability $p_{ij} = Pr(S_t = j | S_{t-1} = i)$ is the conditional probability that the process is in the state *j* at the moment *t* if it was in the state *i* at the previous moment *t-1*. For example, p_{12} is interpreted as the probability of transition from the first state to the second state of the regime, and p_{22} as the probability that the process will remain in the second state. The probabilities p_{11} and p_{12} are complementary, just like the probabilities in p_{21} and p_{22}. The transition probabilities *p* and *q* i1 the stochastic matrix *P* are commonly parameterized using inverse logit transformation:

$$p = \frac{e^{p_0}}{1 + e^{p_0}} \; ; \quad q = \frac{e^{q_0}}{1 + e^{q_0}}. \tag{11}$$

Upon transitional probabilities $Pr(S_t = j | S_{t-1} = i)$, conditional probabilities $Pr(S_t = j | I_{t-1})$ can be generated and then the log-likelihood function can be maximized by the parameters α_1, α_2, β_1, β_2, σ_1^2, σ_2^2 and q_0. Since the process of maximizing the log-likelihood function is iterative, in each new iteration conditional probabilities are updated using Kim's smoothing algorithm, the so-called Kim's filter. Kim's smoothing algorithm can be described in two steps (Kim & Nelson, 2017). In the first step, at the beginning of iteration, *ex ante* probabilities are calculated as follows:

$$Pr(I_{t-1}) = \sum_{j=1}^{2} Pr(S_t = j | S_{t-1} = i) Pr(S_{t-1} = i | I_{t-1})$$

$$j = 1, 2. \tag{12}$$

In the second step, according to the Bayes rule, for the observed values of response variable y_t, the so-called filtered probabilities are obtained:

$$Pr(I_{t-1}, y_t) = \frac{f(y_t | S_t = j, I_{t-1}) Pr(I_{t-1})}{\sum_{j=1}^{2} f(y_t | S_t = j, I_{t-1}) Pr \, Pr(I_{t-1})} \cdot \tag{13}$$

However, initial probabilities need to be determined before the iterative procedure of maximizing the likelihood function can begin. For the initial probabilities, Hamilton (1989) proposed unconditional probabilities of the state of the regime, i.e. steady state probabilities:

$$\mu_1 = \frac{1 - p}{2 - p - q} \; ; \quad \mu_2 = \frac{1 - q}{2 - p - q}. \tag{14}$$

Based on the transitional probabilities of the regime state, the expected duration of the process in the *j*- th regime state can be calculated as:

$$d_1 = \frac{1}{1 - p} \; ; \quad d_2 = \frac{1}{1 - q}. \tag{15}$$

In the two-state regime model, the first regime state is assumed to be a low-volatility state and the second regime state is a high-volatility state. Then the parameters μ_1 and μ_2 can be interpreted as the expected probabilities that the process is in the regime of low (high) volatility in the long term, while the parameters d_1 and d_2 show the duration of the process in low and high volatility regimes in terms of days (Osojnik, 2023). Furthermore, it is worthwhile to analyze the time it takes for the process to switch from low to high volatility states, and *vice versa*.

4. Empirical results

In accordance with the previously described methodology and research objectives, assuming two states of regime *k=2* , the following MRS regression model is estimated:

$$\begin{aligned} track\_error_t &= \alpha_{St} + \beta_{1,St} \cdot illiquidity_{St} \\ &+ \beta_{2,St} \cdot volatility_{St} + \beta_{3,St} \\ &\cdot net\_flow_{St} + \beta_{4,St} \cdot premium_{St} \\ &+ \beta_{5,St} \cdot ln(volume)_{St} + u_{St}. \end{aligned} \tag{16}$$

For two states of regime, 16 parameters (constant term, five coefficients, and error standard deviation for each state along with two transitional probabilities) are estimated by the approximate maximum likelihood method using the expectation–maximization (EM) algorithm due to its convenience (Perlin, 2010). For comparison purposes, a single-regime regression model is also estimated to verify the switching property of regression coefficients (Table 2). In the post-estimation phase, appropriateness of the MRS approach is supported by diagnostic checking of unobserved error term u_{st} that should follow a white noise process with zero mean and constant variance for each state of regime (Table 3).

Table 2 Estimates of a single-regime model and a two-states regime model

	Single	Two-states regime-switching model	
Variable	regime	Regime 1	Regime 2
(Intercept)	0.0172	0.0165***	-0.0321***
	(0.0134)	(0.0043)	(0.0006)
Volatility index	0.0012***	0.0008***	0.0029***
	(0.0001)	(0.0002)	(0.0001)
Illiquidity proxy	1.1595***	1.0975***	1.0049***
	(0.0404)	(0.0666)	(0.0096)
Net flow	0.0522***	-0.1482***	-0.0155***
	(0.0044)	(0.0067)	(0.0014)
Premium/discount	-0.0424***	-0.3891***	0.0507***
	(0.0117)	(0.0277)	(0.0030)
Logs of volume	-0.0051***	0.0033***	-0.0514***
	(0.0015)	(0.0004)	(0.0031)
Error standard deviation	0.0101	0.0089	0.0024
Transitional probability	-	0.8509	0.5762
Observations	1019	1019	-
R2	0.897	0.9318	0.9925
AIC	-6536.0	-8507.4	-
BIC	-6501.5	-8365.2	-
Log.Lik.	3274.995	4265.739	-
RMSE	0.011	0.008	0.002

Note: significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; standard errors in parenthesis
Source: Authors' estimation using data provided by Refinitiv Eikon

For 1019 observations, the results of a single-regime model and a two-states regime-switching model are compared in Table 2. All variables are statistically significant at a 0.1% significance level. An increase in volatility and illiquidity increases tracking error. The results are the same for single-regime and two-states regime models. However, tracking error increases more in the second regime (0.29%) than in the first one (0.08%) with respect to a 1% increase in volatility, while 1% change of illiquidity has approximately the same impact on tracking error in both regimes (increases by 1.09% and 1%, respectively).

For the case of net flows, it shows a positive relationship with tracking error only in a single-regime model. The expected negative relationship was present for the two-states switching model, also indicating a steeper coefficient for a bullish period, meaning that the effect of net flows is stronger during bull periods, i.e. tracking error reduces by 0.14% with respect to a 1% increase of net flow. Furthermore, the results indicate that a premium/discount affects tracking error negatively in a single-regime model; however, their effect is both negative and positive for the bulish and bearish periods, respectively, for the two-states switching model.

Furthermore, the transition probability matrix provides information about probability transitions between two regime states. The probabilities $p_{11} = 0.8509$ and $p_{22} = 0.5762$ indicate the likelihood of remaining in the first and second states of the regime, respectively. Conversely, $p_{12} =$

0.1491 represents the probability of transitioning from the first state to the second state, while $p_{21} = 0.4237$ denotes the probability of transitioning from the second state to the first state. Accordingly, it is more likely to remain in the bullish state regime once the market gets to that state and approximately stays in that state for 7 trading days (a week and a half). In addition, the transition from a bearish to a bullish state of regime is 2.8 times more likely than the reverse, with a probability of 0.4237 compared to 0.1491.

Goodness-of-fit measures confirm the appropriateness and superiority of the two-states regime model over the single-regime model in terms of R2, information criteria, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), respectively, as well as the root mean square error (RMSE). In both regimes, R2 is substantially greater than in a single-regime case. Likewise, RMSE indicates lower regression standard errors in both states against a single state. Smaller AIC and BIC are observed within the Markov regime-switching model, supporting its preference. An R2 value close

to 1 does not imply that the results from regression of time-series are spurious. Spurious regression occurs when regressing two or more independent time-series, resulting in false relationships due to nonstationary properties. To address this issue, the Augmented Dickey-Fuller (ADF) unit root test was conducted on each variable (Table 1), confirming their stationarity. Furthermore, the absence of clear trending behavior in the variables (Figure 2) eliminates the possibility of spurious results stemming from common trends. Once the Markov switching model parameters are estimated, the filtered probabilities of the regime states are obtained by Kim's filtering algorithm, which is a byproduct of the iterative log-likelihood maximization procedure (Perlin, 2010). Inspecting both the filtered and smoothed probabilities is useful for interpreting the switching regression coefficients associated with different time periods (Osojnik, 2023). In Figure 4, it is clear that regime 1 corresponds to the bullish state of the market, while regime 2 corresponds to the bearish state, and more importantly, it covers crisis periods including the COVID-19 pandemic and the onset of the Ukrainian war.

Figure 4 Filtered and smoothed probabilities of the Markov two-states switching model
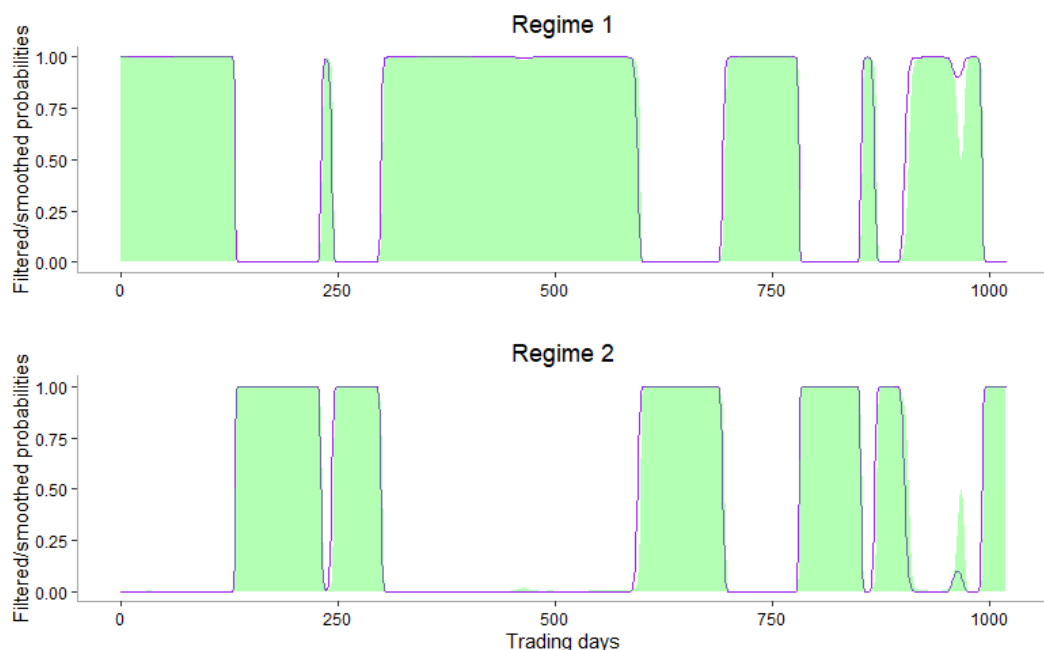
Source: Authors' construction using data provided by Refinitiv Eikon

The validity of the Markov switching model requires diagnostic checking of residuals. Two series of residuals are generated in total, one for each state of regime. Three diagnostic plots are constructed for each residual series in Figure 5 (the correlogram of the residuals, the correlogram of the squared residuals, and the normal quantile-quantile plot), whereas formal diagnostic tests (the Ljung-Box test, the ARCH test and the Jarque-Bera test) are performed on the weighted residuals, i.e. a linear combination of two residual series using smoothed probabilities as the weights. In Figure 5, correlograms indicate no autocorrelation of residuals in both regimes and no autocorrelation of squared residuals, confirming the serial independence of error terms as well homoscedasticity (error terms have constant variance). The same conclusion is supported by non-rejection of the Ljung-Box test null hypothesis with 5 and 10 time lags, and by non-rejection of the ARCH test null hypothesis for autoregressive conditional heteroscedasticity at all significance levels (Table 3). According to the Jarque-Bera test, the normality assumption of weighted residuals is met.

Figure 5 Diagnostic plots of two regime residuals
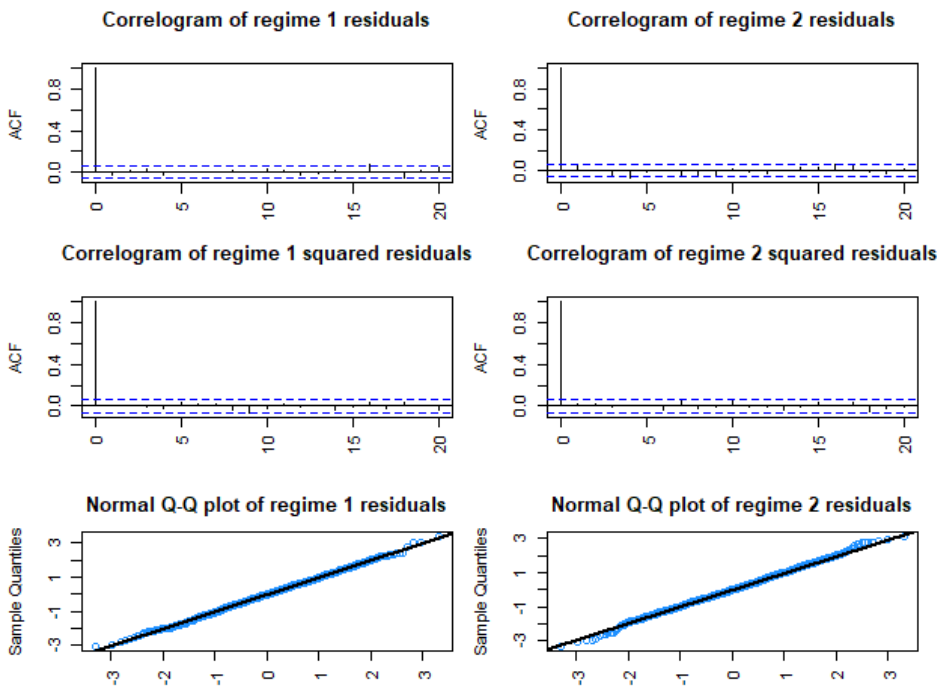
Source: Authors' construction using data provided by Refinitiv Eikon

Table 3 Diagnostic checking of weighted residuals

Test	Statistic
Ljung-Box (5)	1.8469
Ljung-Box (10)	3.7356
ARCH	20.7821
Jarque-Bera	1.3569

Note: significance levels * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Source: Authors' calculation using data provided by Refinitiv Eikon

5. Discussion

The primary objective of this study was to address a critical gap in the analysis of Eurozone ETF performance by focusing specifically on the European market and investigating the impact of market-related variables on ETF tracking error. Both models, the single regime, and the two-states switching regime, confirm that volatility and illiquidity increase tracking error. This is likely because the increase is associated with higher trading costs for arbitrageurs, reducing their ability to create and redeem ETF shares and underlying assets.

When comparing results obtained by both models, it becomes evident that parameters in the single-regime case are overestimated or exhibit unexpected signs. For instance, the impact of illiquidity is overestimated, and the net flow shows a misleading direction of influence. According to Ben-David et al. (2019), increases in trading volume can lead to wider bid-ask spreads, which should, in turn, increase tracking error. Our results confirm this relationship, but only for the bullish period in the market (the first state of the regime). In contrast, during the bearish regime (the second state), trading volume reduces tracking error more than it does in the bullish regime.

Theoretically, an increase in premium should attract authorized participants, hedge funds, and arbitrageurs, thereby decreasing tracking error. This effect is confirmed only for bullish periods (-0.38%). The results regarding the effect of premium/discount on tracking error align with those documented by Rompotis (2012). Interestingly, our study found similar results for bearish periods, contrary to previous findings. One possible explanation for this positive relationship during bearish periods could be attributed to herding behavior among investors, as documented by Ferreruela & Mallor (2021). Shum & Kang (2013) also noted higher premiums/discounts in ETFs during crisis periods, suggesting reduced arbitrage activity due to heightened trading costs. Our analysis supports these findings, indicating consistency with the economic literature for most variables. Notably, in the two-states switching-regime model, the coefficients for net flow, premium, and volume change vary, providing a more comprehensive understanding of the variables influencing tracking error.

These findings underscore the importance of utilizing the two-states switching methodology for researchers. The introduction of the two-states switching methodology clearly demonstrates how the effects of variables such as net flow, premium, and volume can change. In addition to the current variables, supplementary variables like ETF provider's rebalancing frequency, benchmark index composition, and expense ratios could help explain tracking error. However, obtaining this information from publicly available data is not straightforward, and the significance of these variables may be questionable due to their daily time-invariance. For example, expense ratios are typically reported as annual fees, not on a daily basis.

6. Conclusion

The research explored how the tracking error of a Eurozone ETF, concerning its benchmark index, is influenced during crisis periods, including the COVID-19 pandemic and the Ukrainian war. By carefully examining market periods or regimes, the paper offers empirical evidence on several market-based measures, such as market volatility, liquidity proxy, net flow, premium or discount, and trading volume, to comprehensively understand their influence on tracking error. In addition to presenting in-depth explanations and empirical evidence, this study contributes to the literature by employing a regime-switching methodology. The findings support existing economic literature to some extent and highlight the importance of considering different market regimes. While all variables are statistically significant, it was found that an increase in volatility and illiquidity led to a decrease in tracking error. However, the method of switching regimes has shown that the influence of volatility on tracking error is stronger during periods of market stress than in bullish periods (0.29% and 0.08%, respectively), while in terms of illiquidity, the influence is the same for both regimes. When considering trading volume, the results confirm the findings of Ben-David et al. (2019) that an increase in volume does increase tracking error, but only slightly (0.003%). The relationship holds true only for the first regime (the bullish period). On top of that, the study found the influence of volume to be both negative and stronger for periods of market stress (-0.05%). Regarding net flows, the results surprisingly show a positive relationship with tracking error in a single-regime model. However, the use of switching regimes yields the excepted negative relationship between net flows and tracking error. To be

more precise, 1% of net flow reduces tracking error for 0.14% during bullish periods and 0.01% during bearish periods. Lastly, the results show that premium/discount negatively affects tracking error using a single-regime model. With a two-states switching model, the results yield interesting findings. The effect of premium/discount seems to be both positive and negative. A negative influence was expected, and it was found only during the bullish period of the market (-0.38%). A positive but weaker (0.05%) influence of premium/discount on tracking error was found during periods of stress and it was also documented by Rompotis (2012). One of the explanations for a positive influence could lie in herding behavior exhibited by investors during periods of market stress. Another reason could be higher trad-

ing costs which arise in the periods of market stress. Higher trading costs tend to make arbitrage more expensive. Hence, it keeps authorized participants waiting for the price between ETF and its NAV to be further and further away, explaining the positive relationship between premium/discount and tracking error. One limitation of the current research is the fact that only a single ETF is used in the analysis. It would be of great value if future researchers take into account ETFs with different liquidities and sizes.

Acknowledgements

References

1. Aber, J. W., Li, D. & Can, L. (2009). Price volatility and tracking ability of ETFs. *Journal of Asset Management*, *10*(4), 210-221. https://doi.org/10.1057/jam.2009.13

2. Bae, K. & Kim, D. (2020). Liquidity risk and exchange-traded fund returns, variances, and tracking errors. *Journal of Financial Economics*, *138*(1), 222-253. https://doi.org/10.1016/j.jfineco.2019.02.012

3. Ben-David, I., Franzoni, F. & Moussawi, R. (2017). Exchange-Traded Funds. *Annual Review of Financial Economics*, *9*(1), 169-189. https://doi.org/10.1146/annurev-financial-110716-032538

4. Ben-David, I., Franzoni, F. & Moussawi, R. (2019). Do ETFs increase volatility? *Journal of Finance*, *73*(6), 2471-2535. https://doi.org/10.1111/jofi.12727

5. Blitz, D., Huij, J. & Swinkels, L. (2012). The Performance of European Index Funds and Exchange-Traded Funds. *European Financial Management*, *18*(4), 649-662. https://doi.org/10.1111/j.1468-036x.2010.00550.x

6. Broman, M. S. (2016). Liquidity, Style Investing, and Excess Comovement of Exchange-Traded Fund Returns. *SSRN Electronic Journal*, 1-53. http://dx.doi.org/10.2139/ssrn.2307229

7. CFA (2020). *Global Investment Performance Standards (GIBS)*. Chartered Financial Analyst Institute. https://www.cfainstitute.org/-/media/documents/code/gips/2020-gips-standards-firms.ashx

8. Chu, P. K. K. (2011). Study on the tracking errors and their determinants: Evidence from Hong Kong exchange traded funds. *Applied Financial Economics*, *21*(5), 309-315. https://doi.org/10.1080/09603107.2010.530215

9. Chu, P. K. K. & Xu, D. (2021). Tracking Errors and Their Determinants. *Journal of Prediction Markets*, *15*(1), 67-96. https://doi.org/10.5750/jpm.v15i1.1851

10. De Rossi, G. (2015). Measuring the tracking error of exchange traded funds. In Koopman, S. J. & Shephard, N. (Eds.), *Unobserved Components and Time Series Econometrics* (pp. 30-44). Oxford Academic. https://doi.org/10.1093/acprof:oso/9780199683666.003.0004

11. Dorocáková, M. (2017). Comparison of ETF's performance related to the tracking error. *Journal of International Studies*, *10*(4), 154-165. https://doi.org/10.14254/2071-8330.2017/10-4/12

12. Elton, E. J., Gruber, M. J. & de Souza, A. (2019). Are Passive Funds Really Superior Investments? An Investor Perspective. *Financial Analysts Journal*, *75*(3), 7-19. https://doi.org/10.1080/0015198x.2019.1618097

13. Feder-Sempach, E. & Miziołek, T. (2023). How precisely European equity ETFs mirror their flagship benchmarks? Evidence from funds replicating performance of Euro Stoxx 50 Index. *Journal of Asset Management*, *24*(2), 121-135. https://doi.org/10.1057/s41260-022-00287-9

14. Ferreruela, S. & Mallor T. (2021). Herding in the bad times: The 2008 and COVID-19 crises. *The North American Journal of Economics and Finance*, *58*, 101531. https://doi.org/10.1016/j.najef.2021.101531

15. Frino, A. & Gallagher, D. R. (2001). Tracking S&P 500 Index Funds. *Journal of Portfolio Management*, *28*(1), 44-55. https://doi.org/10.3905/jpm.2001.319822

16. Goldfeld, S. M. & Quandt, R. E. (1973). A Markov Model for Switching Regressions. *Journal of Econometrics*, *1*(1), 3-15. https://doi.org/10.1016/0304-4076(73)90002-X

17. Hamilton, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, *57*(2), 357-384. https://doi.org/10.2307/1912559

18. Hillard, J. & Le, T. D. (2022). Exchange-Traded Funds Investing in the European Emerging Markets. *Journal of Eastern European and Central Asian Research*, 9(2), 260-270. http://dx.doi.org/10.15549/jeecar.v9i2.804

19. Johnson, W. F. (2009). Tracking errors of exchange traded funds. *Journal of Asset Management*, *10*(4), 253-262. https://doi.org/10.1057/jam.2009.10

20. Johnson, B., Bioy, H., Kellet, A. & Davidson, L. (2013). On The Right Track: Measuring Tracking Efficiency in ETFs. *The Journal of Index Investing*, *4*(3), 35-41. https://doi.org/10.3905/jii.2013.4.3.035

21. Kaminsky, G. L. (2001). Mutual Fund Investment in Emerging Markets: An Overview. *The World Bank Economic Review*, *15*(2), 315-340. https://doi.org/10.1093/wber/15.2.315

22. Kim, C. J. & Nelson, C. R. (2017). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.* The MIT Press. https://doi.org/10.7551/mitpress/6444.001.0001

23. Le Sourd, V. & Safaee, S. (2021). The European ETF Market: Growth, Trends, and Impact on Underlying Instruments. *Journal of Portfolio Management*, *47*(7), 95-111. https://doi.org/10.3905/jpm.2021.1.252

24. Li, M. & Zhao, X. (2014). Impact of leveraged ETF trading on the market quality of component stocks. *The North American Journal of Economics and Finance*, *28*, 90-108. https://doi.org/10.1016/j.najef.2014.02.001

25. Madhavan, A. (2014), Exchange-Traded Funds: An Overview of Institutions, Trading, and Impacts. *Annual Review of Financial Economics*, *6*(1), 311-341. https://doi.org/10.1146/annurev-financial-110613-034316

26. Meinhardt, C., Mueller, S. & Schoene, S. (2015). Physical and Synthetic Exchange-Traded Funds: The Good, the Bad, or the Ugly?. *The Journal of Investing*, *24*(2), 35-44. https://doi.org/10.3905/joi.2015.24.2.035

27. Mussavian, M. & Hirsch, J. (2002). European Exchange-Traded Funds. *Journal of Alternative Investments*, *5*(2), 63-77. https://doi.org/10.3905/jai.2002.319055

28. Osojnik, L. (2023). *Market predictors of European equity ETF performance* [Undergraduate thesis, University of Zagreb]. University of Zagreb. https://urn.nsk.hr/urn:nbn:hr:148:360888

29. Osterhoff, F. & Kaserer, C. (2016). Determinants of tracking error in German ETFs - the role of market liquidity. *Managerial Finance*, *42*(5), 417-437. https://doi.org/10.1108/MF-04-2015-0105

30. Peovski, F., Cvetkoska, V., Trpeski, P. & Ivanovski, I. (2022). Monitoring stock market returns: A stochastic approach. *Croatian Operational Research Review*, *13*(1), 65-76. https://doi.org/10.17535/crorr.2022.0005

31. Perlin, M. (2010). MS_Regress - The MATLAB Package for Markov Regime Switching Models. *SSRN Electronic Journal*, 1-40. https://doi.org/10.2139/ssrn.1714016

32. Piccotti, L. R. (2018), ETF Premiums and Liquidity Segmentation. *Financial Review*, *53*(1), 117-152. https://doi.org/10.1111/fire.12148

33. Rompotis, G. G. (2010). Does premium impact Exchange-Traded Funds' returns?. *Journal of Asset Management*, *11*(4), 298-308. https://doi.org/10.1057/jam.2009.23

34. Rompotis, G. G. (2011). Predictable patterns in ETFs' return and tracking error. *Studies in Economics and Finance*, *28*(1), 14-35. https://doi.org/10.1108/10867371111110534

35. Rompotis, G. G. (2012). The German Exchange Traded Funds. *The IUP Journal of Applied Finance*, *18*(4), 62-82. https://ssrn.com/abstract=2184748

36. Shum, P. M. & Kang J. (2013). Leveraged and inverse ETF performance during the financial crisis. *Managerial Finance*, *39*(5), 476-508. https://doi.org/10.1108/03074351311313825

37. Sy, A. N. & Ong, L. L. (2004). The Role of Mature Market Mutual Funds in Emerging Markets: Myth or Mayhem? *IMF Working Papers*, *4*(133), 1-24. https://doi.org/10.5089/9781451855524.001

38. Tsalikis, G. & Papadopoulos, S. (2019). ETFS – performance, tracking errors and their determinants in Europe and the USA. *Risk Governance and Control: Financial Markets and Institutions*, *9*(4), 67-76. https://doi.org/10.22495/rgcv9i4p6

39. Vardharaj, R., Fabozzi, F. & Jones, J. (2004). Determinants of Tracking Error for Equity Portfolios. *Journal of Investing*, *13*(2), 37-47. https://doi.org/10.3905/joi.2004.412305

40. Wong, K. H. Y. & Shum, W. C. (2010). Exchange-traded funds in bullish and bearish markets. *Applied Economics Letters*, *17*(16), 1615-1624. https://doi.org/10.1080/13504850903085035

41. Yiannaki, S. M. (2015). ETFs Performance Europe - A Good Start or Not? *Procedia Economics and Finance*, *30*, 955-966. https://doi.org/10.1016/S2212-5671(15)01346-5

42. Zawadzki, K. (2020). The performance of ETFs on developed and emerging markets with consideration of regional diversity. *Quantitative Finance and Economics*, *4*(3), 515-525. https://doi.org/10.3934/qfe.2020024

Ana Skledar Ćorluka
University of Applied Sciences Baltazar Zaprešić
10290 Zaprešić, Croatia
askledarcorluka@bak.hr

THE SIGNIFICANCE OF SOCIALLY ADEPT BEHAVIOR FOR ADVANCING CAREERS AND ENSURING LONG-TERM EMPLOYABILITY WITHIN ORGANIZATIONS

Abstract

Purpose: The aim of this paper is to highlight the increasing importance of social competencies in the business environment, with a particular focus on employee career advancement. The paper seeks to illustrate the significance of specific socially competent behaviors, evaluate their relevance across various industries, and assess how HR managers and top executives perceive these behaviors.

Methodology: The study included a sample of 113 employers. Statistical analyses were conducted using t-tests for dependent and independent samples, ANOVA, and Cronbach's alpha to ensure reliability.

Results: No significant differences were observed between the tested socially competent behaviors in relation to career advancement. These behaviors are equally valued across industries and by both HR managers and top executives in the context of long-term employability.

Conclusion: The findings underscore the importance of empathetic behavior as a key social competency for career progression and long-term employability, placing it on par with communication and conflict management skills. The study also indicates that empathy, communication, and conflict management are equally critical across all economic sectors, encouraging employers to invest in the development of these competencies regardless of their industry. Contrary to the assumption that HR managers prioritize empathy, communication, and conflict management more than top executives do, the research found no significant differences in their perspectives.

Keywords: Social competencies, career advancement, long-term employability, employees, employers

1. Introduction

In today's evolving business environment, competency management systems have become integral to human resource management, shaping organizational strategies and operations (Ivaldi et al., 2022). Social competencies, rooted in diverse disciplines including psychology, sociology, education, and management, are essential for understanding human interactions and behaviors (Wang et al., 2022; Bargsted et al., 2021). Psychology explores interpersonal dynamics and individual behavior, sociology analyzes social structures and norms, and education focuses on fostering social skills and emotional intelligence in students. In business, social competencies are critical to effective leadership, communication, and conflict resolution, pivotal in achieving organizational goals and fostering a productive workplace (Herrera Granda et al., 2023). The study focuses on several key social competencies:

1. **Empathy**: Initially conceptualized by Lipps and Titchener, empathy is defined today as a multidimensional construct requiring the ability to perceive, understand, and share the emotional states of others (Hirn, 2018).

2. **Communication skills**: Vital for organizational success, communication skills are essential behavioral abilities that include comforting, conflict management, persuasion, conversational skills, storytelling, regulatory skills, and referencing skills (Ivaldi et al., 2022)

3. **Conflict management**: Social conflict is inevitable and can have positive or negative consequences depending on how it is managed (Figueiredo et al., 2023).

The research highlights the increasing importance of social competencies in business, especially regarding employee career advancement. It aims to clarify the significance of specific socially competent behaviors, their relevance across industries, and how they are perceived by HR managers and executives. This study addresses a timely issue in today's business environment, examining how social competencies influence career opportunities. As organizations prioritize soft skills like communication and conflict resolution, this research offers a fresh perspective by linking these behaviors directly to workplace outcomes. It suggests that recognizing and developing socially competent behaviors can enhance employee development and retention strategies. Policymakers are encouraged to integrate social competency training into educational curricula and workforce initiatives to foster a more adaptable workforce.

2. Theoretical background

2.1 On social competencies

Interpersonal interactions play a crucial role both professionally and personally, influencing meaningful experiences across various aspects of life, from family and work to leisure activities (Selimović et al., 2018). Socially competent individuals exhibit behaviors such as socially responsible conduct, interpersonal skills, empathy, assertiveness, and cultural cooperation, among others (Maras, 2023). Gargie et al. (1981) define social competencies as goal-directed behaviors that individuals can learn and control. These competencies are essential for navigating social dynamics effectively in diverse contexts.

Social competencies are highly prized by employers for several compelling reasons. Strong communication skills are essential for effective collaboration with colleagues, clients, and partners, facilitating information flow and relationship-building, both internally and externally (Figueiredo et al., 2023). Empathetic employees who listen well and adapt to diverse team dynamics foster a productive work environment conducive to collaboration and mutual support (Galleli & Hourneaux Junior, 2021). In navigating the complexities of the business environment, adept conflict resolution skills are crucial; individuals who can understand other people's perspectives, show empathy, and negotiate win-win solutions contribute significantly to maintaining harmony and productivity (Collie, 2021). Especially in client-centric industries, such as services, employees with refined social competencies can cultivate trust and loyalty by communicating effectively, demonstrating empathy, and adapting to client needs (Collie, 2021). Leadership effectiveness is also amplified by social competencies; leaders who possess these skills are better equipped to motivate and inspire their teams, understand their needs, and foster a positive work culture that enhances overall company performance (Konrad et al., 2021). Ultimately, employers recognize that employees with

developed social competencies contribute to enhanced teamwork, client satisfaction, conflict resolution, and a positive organizational atmosphere, all of which are integral to achieving business success. Therefore, these competencies are often pivotal in recruitment and employee advancement decisions (Samat et al., 2020).

Social competencies are crucial for ensuring employees' long-term employability, encompassing their ability to maintain jobs, enhance career skills, and navigate market changes effectively (Fadhila & Sulistiyani, 2022). This involves continuous learning, adapting to technological advancements, and fostering career flexibility. Employees with strong social competencies demonstrate greater adaptability to diverse workplace dynamics, quickly adjusting to new challenges and team structures, thereby boosting their competitiveness (Fadhila & Sulistiyani, 2022). Additionally, they excel in building professional networks that open doors to business opportunities, mentorship, and career advancement (Nserko, 2020). By cultivating positive relationships with colleagues, superiors, and clients, socially skilled employees foster a productive work environment and contribute to organizational success, reinforcing their long-term employability (Mezhoudi et al., 2023). Their ability to collaborate effectively within teams enhances productivity and supports continuous career growth in dynamic business environments (Potts, 2022). Ultimately, developed social competencies empower employees not only to thrive amidst workplace changes but also to stay competitive and achieve enduring employability in today's dynamic job market.

There is extensive research highlighting the importance of social competencies for achieving success in life. A joint Harvard and Stanford study emphasizes the importance of emotional intelligence and its factors, placing it ahead of logical intelligence (Arghode et al., 2023). Goleman (1998) concludes that emotional intelligence is even twice as important as logical intelligence for achieving success. Morehouse (2007) underscores a strong connection between emotional intelligence, excellent results, and a positive organizational climate. McGee (2010) demonstrates in his research that a pleasant style of managing interpersonal relationships is associated with long-term employability.

The National Research Council (2011) proves that employers focus more on soft skills than on technical competencies. Hogan et al. (2013) later confirm the same. Van Laar et al. (2017) explore key competencies of the 21st century and conclude that they are precisely a combination of technical and social competencies.

2.2 Assessing social competencies

Some researchers equate social competency with social skills, leading to varied research methods and constructs for identical variables. Scholars exploring social competencies within social intelligence distinguish them from numerical intelligence (IQ), prompting the development of specialized measurement techniques (Matthews et al., 2004). IQ, a numerical indicator of cognitive capacity, assesses abilities like verbal comprehension and logical reasoning (Scoupe et al., 2023). However, IQ does not cover all aspects of intelligence such as creativity or emotional and practical skills. This gap has necessitated the introduction of tests tailored to assess social intelligence, complementing traditional IQ assessments (Scoupe et al., 2023).

The George Washington Social Intelligence Test, developed by Moss et al. (1955), is a prominent tool for assessing social competencies. It includes subtests like judgment in social scenarios, discerning mental states, and humor appreciation, showing a reliability coefficient of approximately 0.85, albeit with varying validity outcomes (Brdar & Smojver-Ažić, 1992).

The Six-Factor Social Intelligence Test, introduced by O'Sullivan and Guilford (1966), evaluates social intelligence through subtests such as drawing prediction and expression grouping, condensing over time to maintain reliability and validity (O'Sullivan et al., 1966).

Despite potential biases, assessments and self-assessments play crucial roles in appraising social behavior and emotional intelligence (Goldenthal, 1985).

Role-playing methods in social competency assessment, categorized into multiple-response and single-response scenarios, highlight challenges in replicating real-life interactions (Becker & Heimberg, 1988).

The Social Skills Test for Children (SST-C), with scenes categorized into assistance, praise, and as-

sertiveness, demonstrates satisfactory criterion validity (Williamson et al., 1983).

The Simulated Social Interaction Test, validated for reliability and validity, offers insights into social competencies across contexts (Campbell et al., 2010, as cited in Curran, 1982).

Overall, these methodologies have enriched the author's research by offering various tools to explore and measure social competencies across different contexts and populations, enhancing the depth and breadth of the author's investigations.

3. Materials and methods

3.1 Data

The research focuses on employers' perceptions in the Republic of Croatia regarding the importance of employees' socially competent behavior for career advancement and long-term employability. Specifically, the questions described empathetic behaviors of employees, behaviors indicating possession of communication skills, and behaviors demonstrating adept conflict management. Using

a Likert scale, respondents expressed the extent to which the aforementioned behaviors are important for advancement and long-term employability in their organization. Notably, the survey respondents consisted exclusively of human resource managers and top-level management personnel. The study involved 113 organizations from various economic sectors. The questionnaire was distributed via Google forms. The data collected from the study were processed using SPSS and Microsoft Excel software. In the analysis of primary data, statistical tests such as t-tests for dependent and independent samples, ANOVA and Cronbach's alpha were utilized. The survey is based on the self-assessment method.

3.2 Methodology

3.2.1 Assessing questionnaire reliability

The reliability of a questionnaire measuring desirable social behaviors of employees for the purpose of achieving career advancement and long-term employability for employers was further examined. The questionnaire consists of two different measurement scales.

Table 1 Assessment of the reliability of a questionnaire measuring desirable social behaviors of employees for the purpose of achieving career advancement and long-term employability

Categories	N	Cronbach's alpha
Behaviors demonstrating social competency_Career advancement	30	0.94
Behaviors demonstrating social competency_Long-term employability	30	0.95

Source: Research results

It is evident that the survey questionnaire assessing desirable employee social behaviors according to employers, for the purpose of career advancement and long-term employability, demonstrates a very high degree of reliability, thus confirming its validity as a measurement tool for social competencies. As a measure of reliability, Cronbach's alpha coefficient stands at 0.94 and 0.95. The closer Cronbach's alpha is to 1, the higher the reliability.

3.3 Hypothesis of the study

H1: Employers posit that behaviors reflecting empathy hold lesser significance for career progression and sustained employability compared to employee

behaviors demonstrating adept conflict resolution and effective communication.

The terms "career progression" and "sustained employability" were chosen as variables because social competencies significantly influence their achievement, in contrast to logical intelligence, which is associated with the concept of employment itself (Hogan et al., 2013).

This proposition stems from the emphasis placed by employers on social competencies, notably underscoring the value of communication proficiency and conflict resolution skills. Consequently, it is hypothesized that empathy, as a social competency within the business domain, carries diminished importance (Ivaldi et al., 2022).

Table 2 Intercomparison of the perceived importance of socially competent employee behaviors (empathy, communication skills, conflict management) among employers

Comparison	Social behaviors of employees	N	Mean	Standard deviation	t-value	df	p - value
1	Empathy_ career advancement	113	3.72	0.71	-1.86	112	0.06
	Communication skills _ career advancement	113	3.81	0.63			
2	Empathy_ career advancement	113	3.72	0.71	-0.33	112	0.73
	Conflict management _career advancement	113	3.74	0.70			
3	Communication skills _ career advancement	113	3.81	0.63	1.53	112	0.12
	Conflict management _career advancement	113	3.74	0.70			
4	Empathy_ long-term employability	113	3.55	0.76	-2.11	112	0.03
	Communication skills _long-term_ employability	113	3.66	0.70			
5	Empathy_ long-term employability	113	3.55	0.76	-0.97	112	0.33
	Conflict management _long term_ employability	113	3.61	0.70			
6	Communication skills _long-term_ employability	113	3.66	0.70	1.09	112	0.27
	Conflict management _long term_ employability	113	3.61	0.70			

Source: Author's own elaboration

Table 2 indicates that in the first comparison, there is no statistically significant difference in responses (p > 0.05) regarding the assessment of the importance of employee behaviors indicating empathy and those indicating possession of communication skills, for the purpose of career advancement according to employers' perceptions.

In the second comparison, there is no statistically significant difference in responses (p > 0.05) regarding the assessment of the importance of employee behaviors indicating empathy and those indicating effective conflict management, for the purpose of career advancement according to employers' perceptions.

In the third comparison, there is no statistically significant difference in responses (p > 0.05) regarding the assessment of employee behaviors indicating effective conflict management and those indicating possession of communication skills, for the pur-

pose of career advancement according to employers' perceptions.

In the fourth comparison, there is a statistically significant difference (p < 0.05) in responses regarding the assessment of the importance of employee behaviors indicating empathy and those indicating possession of communication skills, for the purpose of long-term employability according to employers' perceptions.

In the fifth comparison, there is no statistically significant difference in responses (p > 0.05) regarding the assessment of the importance of employee behaviors indicating empathy and those indicating effective conflict management, for the purpose of long-term employability according to employers' perceptions.

In the sixth comparison, there is no statistically significant difference in responses (p > 0.05) regarding

the assessment of employee behaviors indicating effective conflict management and those indicating possession of communication skills, for the purpose of long-term employability according to employers' perceptions.

H2: Behaviors denoting an employee's empathy, skilled communication, and conflict management proficiency hold equal significance across all economic sectors for both career progression and long-term employability.

The presumption is grounded in the unanimous consensus among employers regarding the paramount importance of social competencies across all sectors of industry, spanning from service-oriented to manufacturing enterprises. It is believed that organizations can enhance operational efficiency through the contributions of employees skilled in effective communication, conflict resolution, and empathy.

Table 3 Display of the significance of socially competent employee behaviors (empathy, communication skills, conflict management) across various economic sectors

Social behaviors of employees	Economic sector	N	Mean	Standard deviation	*F*-value	p-value
Empathy_ career advancement	primary	3	3.40	1.11	2.68	0.05
	secondary	27	3.75	0.65		
	tertiary	60	3.86	0.61		
	quaternary	23	3.39	0.91		
Empathy_ long-term employability	primary	3	3.23	1.27	1.55	0.20
	secondary	27	3.48	0.78		
	tertiary	60	3.69	0.64		
	quaternary	23	3.34	0.95		
Communication skills _ career advancement	primary	3	3.23	0.75	1.80	0.15
	secondary	27	3.77	0.65		
	tertiary	60	3.92	0.51		
	quaternary	23	3.68	0.84		
Communication skills _long-term_employability	primary	3	3.30	0.70	1.34	0.26
	secondary	27	3.58	0.77		
	tertiary	60	3.78	0.61		
	quaternary	23	3.51	0.83		
Conflict management _career advancement	primary	3	3.07	0.42	1.38	0.25
	secondary	27	3.87	0.61		
	tertiary	60	3.76	0.66		
	quaternary	23	3.65	0.89		
Conflict management _long term_ employability	primary	3	2.93	0.64	1.06	0.36
	secondary	27	3.63	0.71		
	tertiary	60	3.66	0.63		
	quaternary	23	3.57	0.88		

Source: Author's own elaboration

Table 3 provides a comparative analysis of self-assessed social competencies required in various economic sectors, delineated into empathy, communication skills, and conflict management. Upon scrutiny of each competency individually, it becomes apparent that there are no statistically significant disparities among responses (p > 0.05). This suggests the absence of a correlation between economic sectors and sought-after employee behaviors (their social competencies) conducive to career progression and sustained employability.

H3: Human Resources managers maintain that behaviors reflecting high empathy among employees,

effective communication, and conflict management are more significant for career progression and long-term employability than perceived by top management.

This proposition derives from the fact that Human Resources managers primarily interface with employees within the organization, evaluating their competencies, steering them through career development, facilitating additional training, mentor-

ship, etc. As a result, there is an evident disparity in the recognition of social competencies between Human Resources managers and top management within organizations. Human Resources managers demonstrate a nuanced appreciation of the role of these competencies in fostering a positive work environment and nurturing organizational culture. In contrast, top management tends to prioritize technical expertise and final outcomes to a greater extent.

Table 4 Assessment of the significance of socially competent employee behaviors (empathy, communication skills, conflict management) from the perspectives of human resources managers and top management

Social behaviors of employees	Respondents	N	Mean	Standard deviation	*t*-value	p-value
Empathy_ career advancement	Head of HR	56	3.78	0.69	0.75	0.45
	Top management	57	3.68	0.74		
Empathy_ long-term employability	Head of HR	56	3.56	0.71	0.07	0.95
	Top management	57	3.55	0.82		
Communication skills _ career advancement	Head of HR	56	3.93	0.55	1.80	0.08
	Top management	57	3.71	0.69		
Communication skills _ long-term_employability	Head of HR	56	3.65		-0.29	0.78
	Top management	57	3.68	0.75		
Conflict management _career advancement	Head of HR	56	3.86	0.69	1.65	0.10
	Top management	57	3.64	0.71		
Conflict management _long term_ employability	Head of HR	56	3.63	0.68	0.28	0.78
	Top management	57	3.60	0.74		

Source: Author's own elaboration

Table 4 shows that based on self-assessment, there are no statistically significant differences in responses between Human Resources managers and top management. Both groups believe that employee behaviors indicating empathy, high communication skills, and effective conflict management are equally important for career advancement and long-term employability ($p > 0.05$).

3.4 Results

The first hypothesis, *H1: Employers posit that behaviors reflecting empathy hold lesser significance for career progression and sustained employability compared to employee behaviors demonstrating adept conflict resolution and effective communica-*

tion, is rejected. Namely, based on the presented results and intercomparisons of competencies, it is evident that there is a statistically significant difference only in the comparison of empathy and communication skills for long-term employability. The comparison of empathy and communication skills for career advancement did not show a statistically significant difference in responses. Additionally, the comparison of empathy with conflict management competency did not show a significant difference in responses in the context of either long-term employability or career progression.

The second hypothesis, *H2: Behaviors denoting an employee's empathy, skilled communication, and conflict management proficiency hold equal*

significance across all economic sectors for both career progression and long-term employability, is accepted. Based on the presented results, there are no statistically significant differences in responses between employers from different economic sectors; hence it is concluded that the mentioned social competencies are equally important in all economic sectors.

The third hypothesis, *H3: Human Resources managers maintain that behaviors reflecting high empathy among employees, effective communication, and conflict management are more pivotal for career progression and long-term employability than perceived by top management,* is rejected. Based on the presented results, there are no statistically significant differences in responses between Human Resources managers and top management, thus it is concluded that Human Resources managers and top management have similar views regarding the importance of behaviors indicating social competency for career advancement and long-term employability.

3.5 Managerial implications and limitations of the study

Employers acknowledge the significance of empathetic conduct in advancing careers and ensuring sustained employability. They view empathetic behavior as equally important as behaviors that demonstrate communication skills and conflict management. This underscores the need for continuous promotion and cultivation of empathy among employees to bolster their prospects for long-term employability. It is paramount to underscore that empathy, communication, and conflict management hold equal importance across diverse economic sectors for career progression and sustained employability. This underscores the imperative for employers to invest in fostering these social competencies within their organizations, irrespective of their industry. While it was hypothesized that Human Resources managers attribute greater weight to empathy, communication, and conflict management compared to top management, research findings reveal no notable disparity in their perspectives. It is inferred that top management, despite their primary focus on non-human resources management functions, recognizes that social competencies are as indispensable as technical skills for long-term employability. Consequently, it is advisable for organizations to ensure the promotion of these competencies across all management echelons to opti-

mize employee career advancement and long-term employability outcomes.

The study may have been limited by the size and diversity of the sample (especially in the context of economic sectors). Future research could benefit from a larger and more diverse sample to ensure greater generalizability of the findings across different industries and organizational contexts. Research design and methodology used in this study may have certain limitations. Future studies could employ longitudinal designs or experimental approaches to establish causal relationships between variables of interest. As a limitation of the study, it is also possible to mention the use of parametric tests instead of non-parametric tests. The study relies on perceptual measures of the importance of certain behaviors for career advancement and long-term employability. Future research could incorporate objective measures or performance evaluations to provide a more comprehensive understanding of the impact of these behaviors on career outcomes. The study does not fully explore the influence of contextual factors such as organizational culture, industry norms, or leadership styles on the perceived importance of social competencies. Future research could investigate how these contextual factors may moderate or mediate the relationship between these behaviors and career outcomes.

4. Conclusion

Throughout one's lifetime, individuals are interconnected with others, both during working hours and leisure time. Social skills stand out among the key competencies currently in demand in the job market, characterized by swift job turnover and the concurrent necessity for adaptability. This paper underscores the exceptional significance of social competencies in today's business world and beyond. Crucial to the success of both individuals and organizations, social competencies encompass communication, empathy, conflict resolution, and team leadership. Employers increasingly appreciate employees with well-developed social skills, recognizing their pivotal role in creating a productive work environment, cultivating quality client relationships, and attaining business objectives. Additionally, social competencies play a vital role in the long-term employability of individuals, contributing to their adaptability, offering opportunities for career development, and fostering the

establishment of professional networks. Various methods for measuring social competencies enable the assessment and comprehension of these essential skills, thereby contributing to improved human resource development and management within organizations. Ultimately, understanding and applying social competencies are pivotal for successful business operations and the realization of personal and professional success in the dynamic contemporary business environment.

Previous research tends to focus more on the perception of the importance of social competencies alone and does not necessarily relate them to achieving career advancement and long-term employability. The focus of this research is on the importance of socially competent behaviors for career advancement and long-term employability. Employees are given the opportunity to identify behaviors preferred by employers to ensure long-term employability, thereby enhancing their competencies.

The conclusion of this research highlights several noteworthy findings. Primarily, it illustrates that employers value empathetic behavior as a crucial social competency for career progression and long-term employability of employees, ranking it alongside communication skills and conflict management skills. This underscores the imperative to actively promote and cultivate the development of these competencies within organizations to enhance the long-term prospects of employees. Furthermore, the research suggests that empathy, communication, and conflict management are equally vital across all economic sectors for career progression and long-term employability, prompting employers to invest in the development of these social competencies irrespective of their industry. Despite the assumption that Human Resources managers assign greater importance to empathy, communication, and conflict management compared to top management, the research indicates no significant difference in their perspectives. For future research endeavors, it is recommended to employ larger and more diverse samples, introduce objective measures or success assessments, and explore the influence of contextual factors on the perception of the importance of social competencies.

Acknowledgements

References

1. Arghode, V., Lakshmanan, G. & Nafukho, F. M. (2023). Emotional intelligence, intercultural competence and online instruction: Review and reflection. *European Journal of Training and Development, 47*(5/6), 437-455. https://doi.org/10.1108/EJTD-05-2021-0064

2. Bargsted, M., Yeves, J., Merino, C. & Venegas-Muggli, J. I. (2021). Career success is not always an outcome: its mediating role between competence employability model and perceived employability. *Career Development International, 26*(2), 119-139. https://doi.org/10.1108/CDI-06-2020-0141

3. *Becker, R. E.* & *Heimberg, R. G. (1988). Assessment of social skills. In Bellack, A. S. & Hersen, M. (Eds.), Handbook of Behavioral Assessment* (pp. 365-395). Pergamon Press.

4. Brdar, I. & Smojver-Ažić, S. (1992). Social competence and social status. *Annual of the Institute of Psychology*, *1*, 31-35.

5. Campbell, C., Hansen, D. & Nangle, D. W. (2010) Social Skills and Psychological Adjustment. In Nangle, D. W. et al. (Eds.), *Practitioner's Guide to Empirically Based Measures of Social Skills* (pp. 51-67). Springer. https://doi.org/10.1007/978-1-4419-0609-0_4

6. Collie, R. J. (2021). Motivation Theory and Its Yields for Promoting Students' Social and Emotional Competence. In Yoder, N. & Skoog-Hoffman, A. (Eds.), *Motivating the SEL Field Forward Through Equity* (Advances in Motivation and Achievement, Vol. 21) (pp. 43-59). Emerald Publishing Limited. https://doi.org/10.1108/S0749-742320210000021004

7. Fadhila, N. & Sulistiyani, E. (2022). The influence of motivation, working environment and career development toward employees' loyalty. *AFEBI Management and Business Review, 6*(2), 140-148. https://doi.org/10.47312/ambr.v6i2.503

8. Figueiredo, P. C. N., Sousa, M. J. & Tomé, E. (2023). Integrative model of the leader competences. *European Journal of Training and Development, 47*(5/6), 533-564. https://doi.org/10.1108/EJTD-08-2021-0121

9. Galleli, B. & Hourneaux Junior, F. (2021). Human competences for sustainable strategic management: evidence from Brazil. *Benchmarking: An International Journal, 28*(9), 2835-2864. https://doi.org/10.1108/BIJ-07-2017-0209

10. Gargie, O., Saunders, C. & Dickson, D. (1981). *Social Skills in Interpersonal Communication*. Routledge.

11. Goldenthal, P. (1985). Posing and judging facial expressions of emotion: The effects of social skills. *Journal of Social and Clinical Psychology, 3*(3), 325-338. https://doi.org/10.1521/jscp.1985.3.3.325

12. Goleman, D. (1998). *Working with Emotional Intelligence*. Bantam Books.

13. Herrera Granda, A., Yepes, S. M., Montes Granada, W. F. & Alvarez Salazar, J. (2023). Students' self-perception of social, emotional, and intercultural competences in a public higher education institution in Colombia. *Journal for Multicultural Education, 17*(1), 70-82. https://doi.org/10.1108/JME-02-2022-0032

14. Hirn, S. L., Thomas, J. & Zoelch, C. (2018). The role of empathy in the development of social competence: a study of German school leavers. *International Journal of Adolescence and Youth*, *24*(49), 395-407. https://doi.org/10.1080/02673843.2018.1548361

15. Hogan, R., Chamorro-Premuzic, T. & Kaiser, R. B. (2013). Employability and career success: Bridging the gap between theory and reality. *Industrial and Organizational Psychology, 6*(1), 3-16. https://doi.org/10.1111/iops.12001

16. Ivaldi, S., Scaratti, G. & Fregnan, E. (2022). Dwelling within the fourth industrial revolution: organizational learning for new competences, processes and work cultures. *Journal of Workplace Learning, 34*(1), 1-26. https://doi.org/10.1108/JWL-07-2020-0127

17. Konrad, T., Wiek, A. & Barth, M. (2021). Learning processes for interpersonal competence development in project-based sustainability courses – insights from a comparative international study. *International Journal of Sustainability in Higher Education, 22*(3), 535-560. https://doi.org/10.1108/IJSHE-07-2020-0231

18. Maras, A. (2023). *The significance of social competency and the quality of friendships as predictors of the quality of adolescent school life* [Doctoral dissertation, University of Zagreb]. University of Zagreb.

19. Matthews, G., Zeidner, M. & Roberts, R. D. (2004). *Emotional Intelligence: Science and Myth*. MIT Press.

20. McGee, A. D. (2014). How the Perception of Control Influences Unemployed Job Search. *ILR Review, 68*(1), 184–211. https://doi.org/10.1177/0019793914556245

21. Mezhoudi, N., Alghamdi, R., Aljunaid, R., Krichna, G. & Düştegör D. (2023). Employability prediction: a survey of current approaches, research challenges and applications. *Journal of Ambient Intelligence and Humanized Computing, 14*, 1489-1505. https://doi.org/10.1007/s12652-021-03276-9

22. Morehouse, M. M. (2007). An exploration of emotional intelligence across career arenas. *Leadership & Organization Development Journal, 28*(4), 296-307. https://doi.org/10.1108/01437730710752184

23. Moss, F. A., Hunt, T., Omwake, K. T. & Woodward, L. G. (1955). *Manual for the George Washington University Series Social Intelligence Test.* The Center for Psychological Service.

24. National Research Council. (2011). *Assessing 21st Century Skills: Summary of a Workshop.* The National Academies Press. https://doi.org/10.17226/13215

25. Nserko (2020). Comprehensive social competence and social entrepreneurial action: the mediating role of entrepreneurial tenacity. *World Journal of Entrepreneurship, Management and Sustainable Development, 17*(1), 16-29. https://doi.org/10.1108/WJEMSD-04-2020-0038

26. O'Sullivan, M. & Guilford, J. P. (1966). *Six Factor Tests of Social Intelligence: Manual of Instructions and Interpretations.* Sheridan Psychological Services.

27. Potts, D. (2022). Employability development and career outcomes from short-term learning abroad programmes. *Higher Education Research & Development, 41*(4), 1215-1230. https://doi.org/10.1080/07294360.2021.1901665

28. Samat, M. F. et al. (2020). The Relationship between Career Development, Compensation, Job Security, Work Environment and Employee Loyalty. *e-Academia Journal, 9*(1), 145-161. https://doi.org/10.24191/e-aj.v9i1.9525

29. Scoupe, R., Römgens, I. & Beausaert, S. (2023). The development and validation of the student's employability competences questionnaire (SECQ). *Education + Training, 65*(1), 88-105. https://doi.org/10.1108/ET-12-2020-0379

30. Selimović, Z., Selimović, H. & Opić, S. (2018). Development of social skills among elementary school children. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 6*(1), 17-30. https://doi.org/10.5937/ijcrsee1801017S

31. Van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M. & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior, 72*, 577-588. https://doi.org/10.1016/j.chb.2017.03.010

32. Wang, Y., Sommier, M. & Vasques, A. (2022). Sustainability education at higher education institutions: pedagogies and students' competences. *International Journal of Sustainability in Higher Education, 23*(8), 174-193. https://doi.org/10.1108/IJSHE-11-2021-0465

33. Williamson, D. A., Moody, S. C., Granberry, S. W., Lethermon, V. R. & Blouin, D. C. (1983). Criterion-related validity of a role-play social skills test for children. *Behavior Therapy, 14*(4), 466-481. https://doi.org/10.1016/S0005-7894(83)80071-9

Preliminary communications

Kenan Mahmutović
University of Bihać
Faculty of Economics
77000 Bihać, Bosnia
and Herzegovina
kenan.mahmutovic@efbi.unbi.ba

THE STATE OF SEO OF E-COMMERCE WEBSITES THAT SELL ELECTRONICS IN BOSNIA AND HERZEGOVINA AND THE IMPACT OF SEO FACTORS ON TRAFFIC AND REVENUE

Abstract

Purpose: The research aims to assess the SEO status of electronics e-commerce websites in Bosnia and Herzegovina and analyze the impact of certain SEO factors on website traffic and company revenue.

Methodology: Empirical research was conducted on 47 out of 50 identified e-commerce websites that sell consumer electronics in Bosnia and Herzegovina. Descriptive statistical analysis was used to determine the share and structure of organic search visits, the current state of SEO factors, and their impact on user experience and visitor behavior. Path analysis was used to identify and quantify the relationship between SEO factors and website traffic as well as between website traffic and company revenue.

Results: Organic search visits represent 23.6% of traffic, with branded visits comprising 48.9%. Mobile optimization needs attention, with 6.4% lacking mobile-friendly design and 75.4% exhibiting poor mobile performance. Desktop performance scores are generally low, with only 12.8% rated as good. An average number of pages per visit is 2.67, and the average bounce rate is 61.4%. The authority score, which measures the quantity and quality of backlinks, has a strong positive impact on the number of keywords. In turn, the number of keywords has a strong positive impact on the number of visits. Finally, visits have a strong positive impact on the revenue of e-commerce websites.

Conclusion: Path analysis determined the link between SEO factors, traffic, and revenue. Recommendations of the most critical tactics for attracting web visitors and improving user experience and retention through SEO optimization are given.

Keywords: SEO, search engine optimization, e-commerce, website traffic

1. Introduction

Business success of an online retailer largely depends on the number of visitors to the e-commerce website and the conversion rate. There are several possible sources of website visits: direct visits achieved by entering a known web address in a web browser, visits from online advertisements, visits through referral links (links on other web-

sites, social networks, or e-mail messages), and visits from search engines known as organic traffic (Mahmutović, 2021).

Organic traffic denotes the flow of visitors who use search engines to locate desired information at no cost (Roumeliotis et al., 2022). Following a keyword search, search engine results display a list of websites relevant to the searched terms (Luh et al., 2016), ranking them based on various factors, such as content relevance, the number and rank of incoming links and the link structure (Tomasi & Li, 2015). Out of 100 searches conducted for a specific keyword, 88 users clicked on one of the first ten results (Sgueo, 2023). A website will receive more visitors from a search engine (organic search) when websites rank higher on the search engine results page (SERP), and these visitors can then be converted into customers (Ortiz-Cordova & Jansen, 2012). Achieving a high organic search ranking is typically tricky since it cannot be bought (Matošević et al., 2021). Websites aiming for top positions must meet specific criteria, termed ranking factors, outlined in guidelines periodically issued by search engine companies (Google, 2023e). While these factors are publicly announced, the exact impact on search rankings remains undisclosed as search engines keep their algorithms and ranking criteria confidential (Iqbal et al., 2022).

According to BrightEdge's 2019 report, organic search is the channel that brings the most traffic to websites, on average, organic search accounts for 53.3% of traffic across all industries and 41% in retail and e-commerce (BrightEdge, 2019). The same report states that, in terms of revenue, organic search is the largest channel, accounting for 44.6% across all industries and 36.4% in retail and e-commerce. However, according to SemRush's (2023a) report, the last three years have seen a decrease in organic traffic to e-commerce sites worldwide. For example, Amazon lost 37.5% of organic visits in 2021 compared to 2020. The same report states that in 2022, organic search visits were the second most significant source of visits to e-commerce websites, with a share of 27.6% of visits in Europe and 12.7% in the U.S, while direct visits make up more than 85% of all U.S. e-commerce traffic, compared to 55.9% in Europe. Considering the primary sources of visits, building brand awareness (for direct visits), and optimizing websites for better positions on search engines (Search Engine Optimization - SEO) have become the most critical marketing tools for at-tracting visitors to online stores. This paper aims to investigate the SEO factors that affect higher rankings of e-commerce websites in search engines and how they impact the number of visits and revenue for online retailers.

2. Literature review

2.1 What is Search Engine Optimization (SEO) and why is it important?

While strengthening brand awareness through long-term investments in promotion, especially advertising and public relations, can increase the number of direct visits to the website, search engine optimization (SEO) can achieve more visits from search engines.

SEO optimizes web pages and edits website content and code to improve its visibility and ranking in the organic (non-paid) search results within one or more search engines (Li et al., 2014; Killoran, 2013).

Successful SEO should push a web page to the top of search results for the requested keyword or phrase. Being at the top of search results is essential because visitors mostly visit only the first links on search results pages. Research conducted by Sistrix, which analyzed over 80 million keywords and billions of search results, showed that out of 100 searches conducted, 88 users click on one of the first 10 results, 70 of them click on one of the first five results, and 55 users click on one of the first three results (Sgueo, 2023).

SEO includes on-page and off-page optimization (Chotikitpat et al., 2015). On-page optimization focuses on content optimization according to keywords, content and URL structure optimization, image optimization, internal and outgoing links (outbound links) optimization and technical website optimization to ensure the validity of the code, high loading speeds and adaptability for different screen size devices (Chotikitpat et al., 2015; Hui et al., 2012; Viney, 2008; Fuxue et al., 2011; Iqbal et al., 2022). This process should help search engines understand and present content (Google, 2023b) to users. Google claims that its algorithms look at many factors and signals, including the words of query, relevance, and usability of pages, the expertise of sources, and user location and settings, and when all things are relatively equal, page experience (mobile friendliness and page loading speed) will be the deciding factor for ranking (Google, 2023d). Mobile friendliness is ensured with responsive

design. Responsive design ensures dynamic adaptation of the contents to the display size, thus ensuring a good viewing experience regardless of the web access device (Almeida & Monteiro, 2017). The loading speed of websites reflects website performance and significantly influences user experience and satisfaction (Bartuskova & Krejcar, 2015). Outbound links refer to links originating from websites and directing users to external sources. Related outgoing links provide helpful information to the visitors, and a quantity of unique outbound links improves a website's ranking (Nath & Ahuja, 2014; Iqbal et al., 2022).

Off-site optimization relates to the practices through which the website and its contents are propagated over the Internet to increase its traffic, which enhances a website's ranking (Khan & Mahmood, 2018). It includes link building and website submission to search engines and link directories (Mahmutović, 2021). Link building aims to ensure that as many other sites as possible link to the e-commerce website, thus showing the web search engines that the website is authoritative and relevant for the targeted keywords. If other prominent websites link or refer to web page content, the search engine will prioritize that page, rating it positively for expertise, authoritativeness and trustworthiness (Google, 2023d)**.** Essential factors for link building are backlink quality (the total number of backlinks, the number of referring domains, and the quality of the web page where the link is placed), anchor text used and social media activity (Alfiana et al., 2023). Utilization of on-site optimization is known as the content SEO strategy, while off-site optimization is known as link building and social sharing strategies (Zhang & Cabbage, 2017).

SEO requires significantly lower investment than advertising, so it is no surprise that in recent years, it has become one of the favorite techniques of digital marketers to secure free traffic. The estimated value of the SEO services market increased from USD 62.75 billion in 2022 to USD 74.76 billion, with an expected growth of USD 146 billion by 2027 at a CAGR of 18.4% (The Business Research Company, 2023).

For a company to have insight into search engine optimization of its website and the amount of traffic it achieves through organic search, it is necessary to use digital analytics. Digital analytics includes collecting, measuring, analyzing, visualizing, and interpreting digital data that illustrate user behavior on websites and mobile applications. It includes web, social media, SEO, and user experience analytics (Mahmutović, 2021). While web analytics is focused on collecting and processing data related to website performance, SEO analytics is focused on tracking the signals that dictate a website's position on search engine results pages. These signals include the number of incoming links and their quality (backlinks), keywords, website indexing errors, broken links, and user experience measured through website performance and loading speed (Mahmutović, 2021).

According to StatCounter (2023), Google dominates the search engine market today, with a 91.85% market share. It is followed by Bing (3.02%), Yandex (1.49%), Yahoo (1.17%), Baidu (1.06%), and DuckDuckGo (0.54%). As the largest and most important player on the market, Google uses automated ranking systems that look at many factors and signals about hundreds of billions of web pages and other content in its search index to present the most relevant and valuable results. Over time, the company has developed several AI systems that help Google rank results more effectively in its index to:

- understand representations of concepts in queries and pages and match them to one another (Neural Matching);
- show original content prominently in search results, including original reporting, ahead of sources that merely cite it (Original Content Systems);
- show only the most relevant results to avoid unhelpful duplication (Deduplication Systems);
- understand how combinations of words express different meanings and intent (Bidirectional Encoder Representations from Transformers);
- show fresher content for queries where it is expected (Freshness System);
- ensure that users see original, helpful content written by people for people in search results, rather than content made primarily to gain search engine traffic (Helpful Content System);
- understand how pages link to each other as a way to determine what pages are about and which might be most helpful in response to a query (Link Analysis System and PageRank);

- understand how words are related to concepts (RankBrain);
- better reward high quality reviews, i.e., content that provides insightful analysis and original research (Reviews System);
- deal with content and behaviors that violate Google spam policies (Spam Detection systems) (Google, 2023c).

Although the importance and impact of specific signals on the ranking of websites in the Google index is a business secret, through its documentation, Google (2023b) indicates significant SEO factors and provides recommendations for content optimization and technical optimization. Google especially emphasizes the importance of creating compelling and valuable content, the use of unique and informative meta tags and structured data markup that describe the content of a website to the search engine, optimization of the website structure and navigation elements, image optimization, and content optimization for mobile devices.

2.2 The influence of SEO factors on SERP positions and traffic

Most researchers and practitioners have concluded that the vital SEO techniques for improving SERP positions are improving keyword prominence, i.e., by increasing keyword visibility on websites (Mwosa Kivuti, 2018; Krrabaj et al., 2017), and improving the page rank through the quality and quantity of backlinks (Su et al., 2014; O'Neill & Curran, 2011; Giomelakis & Veglis, 2015; Tomasi & Li, 2015; Seyfabad & Fard, 2019; Krrabaj et al., 2017). O'Neill and Curran (2011) stated that SEO has many factors that affect website ranking performance within search engines, such as meta tags, content, link popularity, click popularity, and longevity, which determine a site's ranking. The authors especially emphasized the importance of keywords, relationships, and communication with other websites (backlinks and outbound links). Similarly, Su et al. (2014) revealed that the quality and quantity of backlinks (page rank[1]) are the dominant SEO factor. Ziakis et al. (2019), Zilincan (2015), and Tsuei et al. (2020) determined the positive impact of the appearance of keywords and key phrases in the

hostname, main page headings, image descriptions, the path segment of the URL, main content and page meta tags, on the SEO results. By testing the existence and strength of the correlation between individual SEO factors and the website's rank, Ziakis et al. (2019) found that an additional factor for SEO success is domain age. Domain age can significantly impact its ranking by enhancing trust and credibility among website users, whereby domains registered for two years or more are perceived as more reliable than newer domains (Shenoy & Prabhu, 2016). The opposite conclusion was reached by Chakrabortty & Jose (2018), analyzing the relationship between domain age and website traffic, where they determined that the correlation between domain age and monthly page views is insignificant, i.e., that domain age has no significant role in generating page views. Krrabaj et al. (2017) confirmed the importance of link building and keyword optimization and emphasized the importance of content currency.

The influence of page experience on SEO results through mobile-friendly responsive design and page loading speed has been confirmed by Gao et al. (2017), Bartuskova and Krejcar (2015), Mustafa et al. (2015) and Egri and Bayrak (2014). Contrary to those studies, Ziakis et al. (2019), Marszałkowski et al. (2014), and Roumeliotis et al. (2022) found out that the website loading time does not affect the ranking of a website. However, Marszałkowski et al. (2014) found that crawl time measured by Google's indexing robot plays a role in the Google algorithm, although its impact is not big.

Some studies claim that outbound links can significantly influence website authority. Specifically, linking to relevant and authoritative websites positively impacts ranking in search engine results pages (Iqbal et al., 2022; Nath & Ahuja, 2014). While SEO optimization literature suggests that offering valuable information via outbound links enhances trust and credibility on a website (Shenoy & Prabhu, 2016), empirical research validating the extent of this association is still ongoing and remains challenging to determine.

Previous research has confirmed a positive impact of SEO on website traffic, visit duration and user

---

[1]   PageRank works by analyzing the links between web pages. It assigns a numerical weight to each page, representing its importance on the web. The underlying idea is that a page is considered more valuable if it is linked by other pages, especially if those pages are important. It was developed by Larry Page and Sergey Brin, the founders of Google, as an algorithm used by Google Search to rank web pages in their search engine results.

engagement (Giomelakis & Veglis, 2015; Baye et al., 2016; Tomasi & Li, 2015; Roumeliotis et al., 2022).

As visit duration and user engagement increase the probability of making a sale (Bhatnagar & Ghose, 2004.; Lee et al., 2011), through this influence, SEO also indirectly affects conversions and sales revenue (Roumeliotis et al., 2022; Sinaga et al., 2024; Tomasi & Li, 2015). A positive impact of SEO practice is not only limited to increasing traffic and sales revenue but also on other marketing variables like market share, brand equity, product awareness, purchase persuasion, consumer insights and business growth (Bhandari & Bansal, 2018; Sinaga et al., 2024).

Giomelakis and Veglis (2015) investigated the extent to which SEO factors influence website traffic. They researched 30 media websites in Greece, and the results showed that the websites with the most visits used SEO. It was established that there is a significant correlation between website visits and various SEO factors, such as backlinks, a bounce rate, and link exchange practice. Baye et al. (2016) found that the quality and brand awareness of a site increase organic clicks through the better position of higher-quality e-commerce websites on search results pages. The authors realized that a firm not appearing on the first five search results pages receives 90% fewer clicks for a given search term. Tomasi and Li (2015) investigated the impact of SEO initiatives on the performance of three SMEs in Maryland. They found that SEO leads to an increase

in the number of visitors to the site, an increase in the average time duration of users visiting the site, and increased user engagement. Sinaga et al. (2024) came to similar findings in the e-commerce industry in Indonesia, establishing the positive impact of SEO on business growth. Roumeliotis et al. (2022) analyzed the applied SEO techniques (optimization of meta tags, titles, use of SEO-friendly URL, use of structured data, responsiveness) and SEO metrics (domain authority, organic keywords, website speed, and backlinks) on 243 websites of airline carriers and their impact on website traffic. They concluded that SEO techniques and metrics correlate significantly with website traffic, increasing airline conversions and bookings. A particularly significant and strong relationship was established for backlinks and keywords. Seyfabad and Fard (2019) investigated how SEO metrics can affect the ranking of university websites. The results showed that metrics like backlinks, PageRank, link quality, the number of indexed pages, the number of referring domains and IPs, and the authority of page and domain significantly affect the website's visibility. Therefore, the authors conclude that universities can improve their ranking on the Webometrics list with on-page optimization and off-page optimization.

Figure 1 shows the connection between SEO factors, SEO performance indicators, and sales revenue based on the literature review.

Figure 1 Connection between SEO factors, SEO performance indicators and sales revenue
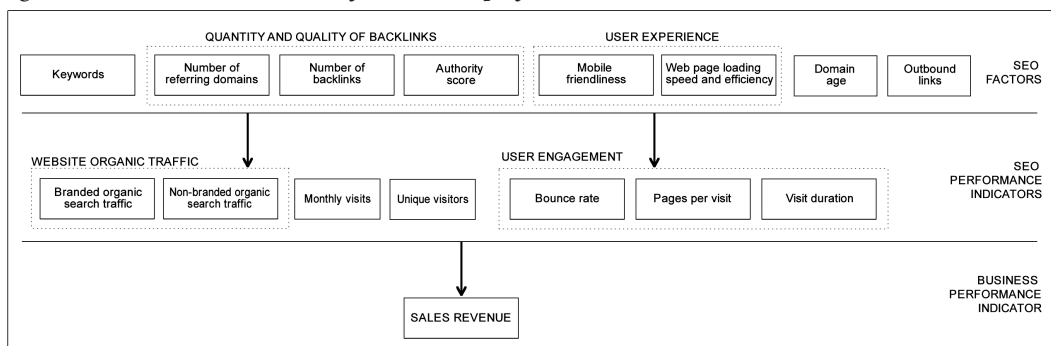
Source: Author

3. Method

Empirical research for this paper was conducted on the population of e-commerce websites in Bos-

nia and Herzegovina that sell consumer electronics (computers and computer equipment, smartphones, and TV sets). A Google search using the keywords in

the Bosnian language, "computer sale", "smartphone sale", and "TV sets sale", identified a total of 50 online shops in Bosnia and Herzegovina that sell the aforementioned products. Three online shops, which were found to belong to large retail chains in Bosnia and Herzegovina, were excluded from the sample. The remaining 47 e-commerce websites were analyzed with several online tools to collect data on individual SEO factors and SEO performance indicators.

Reliable quantitative data are imperative for effectively investigating and analyzing specific variables. That is why a significant number of researchers opt to employ external tools to obtain essential information, such as metrics like unique visitor count, bounce rate, link volume, and domain authority, among others—metrics that Google itself does not furnish directly (Reyes-Lillo et al., 2023).

The list of analyzed SEO factors, SEO performance indicators and the names of the services used to collect data are listed in Table 1 and Table 2.

One of the analyzed SEO factors is the domain authority score. Domain authority is a prominent metric frequently utilized within professional contexts that has garnered considerable attention within academic scholarship, as evidenced by studies conducted by Vyas (2019), Urosa-Barreto (2020), Nagpal and Petersen (2021), and Ganguly (2022).

Moz, Semrush, and Ahrefs are prominent SEO analytics companies providing valuable insights into domain authority scores. The domain authority values supplied by Moz (Domain Authority), SemRush (Authority Score), and Ahrefs (Domain Rating) can be considered reliable (Reyes-Lillo et al., 2023). We decided to use the SemRush indicator (Authority Score), which is based on multiple factors of trustworthiness and authority, including the quality and quantity of site's backlinks, organic search traffic, and indicators of manipulation or spam in the link profile (Reyes-Lillo et al., 2023; SemRush, 2023b). Varangouli (2020) explains that the authority score is assessed using a logarithmic scale ranging from 0 to 100, where higher scores are associated with increased traffic and improved rankings.

Data on the number of employees and income in 2022 for every company were collected through the web portal of the Financial Information Agency of the Federation of Bosnia and Herzegovina[2]. Statistical data processing, descriptive statistical analysis, and path analysis were performed using R statistical software.

Table 1 SEO factors and data sources

SEO factor	Description	Source
Mobile and Desk performance score	Performance score measures the speed and efficiency of loading a website. Google prioritizes websites with high speed because it improves user experience. The performance score is a weighted average of metric scores, and the weight of each metric is a representation of the user's perception of performance. Metrics included in the calculation are First Contentful Paint (FCP), Speed Index (SI), Largest Contentful Paint (LCP), Time to Interactive (TTI), Total Blocking Time (TBT), and Cumulative Layout Shift (CLS).	Google PageSpeed Insights
Largest Contentful Paint (LCP)	Measures loading performance. To provide a good user experience, LCP should occur within 2.5 seconds after the page first starts loading (Walton & Pollard, 2019).	Google PageSpeed Insights
First Input Delay (FID)	Measures *interactivity*. To provide a good user experience, pages should have an FID of **100 milliseconds** or less (Walton, 2019).	Google PageSpeed Insights
Cumulative Layout Shift (CLS)	The user-centric metric for measuring visual stability because it helps quantify how often users experience unexpected layout shifts. To provide a good user experience, pages should maintain a CLS of 0.1. or less (Mihajlija & Walton, 2019).	Google PageSpeed Insights

SEO factor	Description	Source
Authority Score	This is a metric for the overall quality of a website or webpage. The higher the score, the more weight domain or webpage outbound links to another site could have.	SemRush
Backlinks	The number of links from other websites to a specific website.	SemRush
Keywords	The number of keywords bringing users to an analyzed domain via Google's top 100 organic search results.	SemRush
Referring domains	The number of referring domains with at least one link points to an analyzed domain/URL.	SemRush
Outbound domains	The total number of domains the analyzed domain or URL points to.	SemRush
Domain age	Calculated domain age in days as of 25 September 2023.	https://nic.ba https://who.is/
Mobile-friendliness score	A categorical variable indicating whether the website is optimized for mobile device use (0=No, 1=Yes).	https://search.google.com/test/mobile-friendly

Source: Author

Table 2 SEO performance indicators and data sources

SEO performance indicator	Description	Source
Bounce Rate	Bounce rate is the percentage of visitors that leave a website after viewing only one page.	SemRush
Pages per visit	An average number of pages visited during one visit.	SemRush
Visit Duration	Average duration of the visit in seconds.	SemRush
Monthly visits	The number of unique visits to a root domain for the last month.	SemRush
Organic search traffic	It estimates organic traffic driven to a domain from Google search results.	SemRush
Branded organic search traffic	The amount of traffic that comes to an analyzed domain from keywords that include the domain's brand name.	SemRush
Non-branded organic search traffic	The amount of traffic that comes to an analyzed domain from non-branded keywords.	SemRush

Source: Author

The research questions in this paper are:

1. What is the share of organic search visits to e-commerce websites that sell consumer electronics in Bosnia and Herzegovina, and what is the ratio of branded vs. non-branded organic visits?

2. What is the state of SEO of e-commerce websites that sell consumer electronics in Bosnia and Herzegovina?

3. Which SEO factors affect e-commerce website traffic and sales revenues in Bosnia and Herzegovina, and to what extent?

4. Which SEO areas should e-commerce websites that sell consumer electronics in Bosnia and Herzegovina improve to achieve better business results?

4. Results and discussion

4.1 Sample profile

The sample consists of 47 e-commerce websites in Bosnia and Herzegovina that sell computers and computer equipment, smartphones, and TV sets. The following tables show the sample profile according to the number of employees, domain age and annual revenue.

Table 3 Profile of the sample according to the number of employees

Company type	Number of employees	N	%
Microenterprise	1-9	20	42.6
Small business	10-49	16	34.0
Medium enterprise	50-249	10	21.3
Large enterprise	> 250	1	2.1

Source: Author

Table 4 Profile of the sample according to the age of the web domain

Domain age	N	%
Up to 1 year	1	2.1
1 – 5 years	7	14.9
5 – 10 years	11	23.4
More than ten years	28	59.6

Source: Author

Table 5 Profile of the sample according to revenue per employee in 2022

N	47
Mean	387,317 BAM
Median	338,426 BAM
Std. Deviation	259,974 BAM
Minimum	60,703 BAM
Maximum	1,166,157 BAM
Percentile - 25	195,489 BAM
Percentile - 50	338,426 BAM
Percentile - 75	504,053 BAM

Source: Author

4.2 Share and structure of organic search visits

Organic search visits to the analyzed e-commerce websites in Bosnia account for an average of 23.6% of the total number of visits, compared to the European average of 27.6% for retail and e-commerce websites, as reported by SemRush (2023a). Out of 47 analyzed websites, 50% have up to 17.1% share of organic search visits, 25% of websites achieve between 17% and 27.5% share of organic search visits, while 25% achieve over 27.5%.

The amount of organic search traffic from keywords that include the domain's brand name (branded organic search traffic) is 45.8%, while the share of non-branded organic search traffic is 54.2% (n=46). Out of 46 websites with complete data on the number of branded organic search visits, 25% of websites achieve up to 22.45% of branded organic search visits, and 50% of analyzed websites have more than 48.6% of branded organic search visits in the total number of organic visits.

The results indicate that strengthening brand awareness is a crucial marketing tactic for e-commerce websites in Bosnia and Herzegovina to generate traffic. On average, a quarter of visits to e-commerce websites come from search engines, while the remaining 75% are direct visits or vis-

its from referral links. In addition, 46% of organic visits are branded organic visits, which indicates that through promotional activities strengthening brand awareness accounts for almost half of the organic visits to e-commerce websites.

4.3 Analysis of user experience through SEO factors

Mobile friendliness analysis shows that 93.6% of websites are usable on mobile devices, while 6.4% are not mobile-friendly.

A performance score was measured for each website regarding the speed and efficiency of loading websites as essential elements of the user experience. A score of 90 or above is considered good, 50 to 89 is a score that needs improvement, and below 50 it is considered poor (Google, 2023a). Descriptive analysis shows that the average performance score of the analyzed websites for mobile devices is 38, and for desktop devices it is 67. Additional frequency analysis indicates that 74.5% of the analyzed websites have a poor mobile performance score, and 25.5% have a mobile performance score that needs improvement. Regarding desktop devices, 17% of websites have a poor performance score, 70.2% need improvement, and only 12.8% have a good performance score.

The results tell us that e-commerce websites need to work on improving loading speed and improving user experience. Particular focus must be directed to the factors included in the calculation of the Google metric called Core Web Vitals, which focuses on three aspects of user experience: loading (LCP), interactivity (FID), and visual stability (CLS). Web Vitals is an initiative by Google to provide unified guidance for quality signals essential to delivering a great user experience on the web. Core Web Vitals are the subset of Web Vitals that apply to all web pages, should be measured by all site owners, and will be surfaced across all Google tools (Walton, 2020).

Descriptive statistics for Core Web Vitals metrics are shown in Table 6.

Table 6 Core Web Vitals assessment

Metric	Thresholds	Measured average value
Largest Contentful Paint (LCP) – Mobile	Good (≤ 2.5 s) Needs Improvement (2.5s - 4s) Poor (> 4 s)	3.48s
First Input Delay (FID) – Mobile	Good (≤ 100 ms) Needs improvement (100 ms - 300 ms) Poor (> 300 ms)	38.11 ms
Cumulative Layout Shift (CLS) – Mobile	Good (≤ 0.1) Needs Improvement (0.1 – 0.25) Poor (> 0.25)	0.1362
Largest Contentful Paint (LCP) – Desktop	Good (≤ 2.5 s) Needs Improvement (2.5s - 4s) Poor (> 4 s)	2.77s
First Input Delay (FID) – Desktop	Good (≤ 100 ms) Needs improvement (100 ms - 300 ms) Poor (> 300 ms)	3.81ms
Cumulative Layout Shift (CLS) – Desktop	Good (≤ 0.1) Needs Improvement (0.1 – 0.25) Poor (> 0.25)	0.1340

Source: Author

Based on actual user data for the last 28 days, the Core Web Vitals assessment shows that e-commerce websites should improve user experience regarding loading performance (LCP) and visual stability (CLS) for mobile and desktop devices. In general, optimization of LCP includes optimizing images, removing unnecessary characters from HTML/CSS code, using a faster hosting provider, using a CDN (content delivery network), and techniques for delaying the loading of images un-

til they are needed (Walton & Pollard, 2020). The most common reasons for high CLS are undefined image dimensions, dynamically injected content that causes content to shift down, embeddable widgets such as videos from YouTube, maps from

Google Maps, social media posts, and ads with dynamic ad sizing (Osmani & Pollard, 2020). Those are primarily technical issues that e-commerce marketers must know when they plan and develop their websites.

Table 7 Metrics of user behavior on a website

		Bounce rate (%)	Pages per visit	Visit duration in seconds
N	Valid	39	44	42
	Missing	8	3	5
Mean		**61.4**	**2.67**	**508.1**
Median		63.4	2.20	250.50
Std. Deviation		31.05	1.62	649.62
Range		97.61	5.86	2,489
Minimum		2.39	1.00	10
Maximum		100.00	6.86	2,499
Percentiles	25	37.17	1.17	97.00
	50	63.38	2.20	250.50
	75	92.23	3.69	617.75

Source: Author

The analysis shows that the average percentage of visitors who leave an e-commerce website after viewing only one page is 61.4%. This bounce rate is significantly higher than the benchmark bounce rate average for e-commerce and retail websites, which is 20%-45% (CustomMediaLabs, 2023). A high bounce rate is a sign that website content is irrelevant to the searched keyword, that users need help finding the content they are looking for or need help navigating, or that the site offers a poor user experience. Through digital analytics, marketers should gain insight into the bounce rate by different sources of visits (direct visits, organic search visits, referral links) to understand which keywords and phrases the content on the website should be optimized for. With this approach, it is also possible to gain insight into whether a high bounce rate for visits from referral links indicates backlinks of poor quality.

The average number of pages visitors view on a site within a single session is 2.67, which is 45% less than the average for mobile users in the retail industry (4.87) and 51% less than the average for desktop users (5.5) (Contentsquare, 2023). The average visit duration is 8 minutes and 28 seconds, while on 50% of the analyzed websites, the duration is half as long (4 minutes and 21 seconds). Although the average pages per visit can vary depending on

several factors, such as the type of e-commerce website, a target audience, and website design and content, marketers should analyze the reasons for lower retention, with particular emphasis on simplicity of navigation, easier finding of related products, personalized recommendations and use of social proof, such as customer reviews and ratings.

4.4 Backlinks, keywords, referring domains, and outbound links

Descriptive statistics for backlinks, keywords, referring domains, outbound links, and authority scores are shown in Table 8. The average number of keywords bringing users to an analyzed e-commerce website via Google's top 100 organic search results is 848. Given several outliers with significantly more backlinks than other websites, the median better represents the average value of backlinks and it is 12,400 links. The average number of referring domains with at least one link pointing to an e-commerce website is 1,841, while the average authority score of the analyzed websites is 18. Outbound links from e-commerce sites signal trustworthiness and relevance to search engines. The average number of outbound domains on the analyzed e-commerce websites is 28.7. Google evaluates the quality of a web page through experience,

expertise, authoritativeness, and trustworthiness (Varangouli, 2020). Linking to other sources can help e-commerce websites to show their expertise. External links can help retailers validate the quality and reliability of their products by referring to

product reviews, ratings, testimonials, manufacturer specifications, user manuals, or external resources that explain payment security, privacy policies, and other essential trust factors.

Table 8 Backlinks, keywords, referring domains and outbound domains statistics

		Authority score	Keywords	Backlinks	Referring domains	Outbound domains
N	Valid	47	47	47	47	45
	Missing	0	0	0	0	2
Mean		18.11	848	54678	1841	28.7
Median		17.00	411	12400	1600	15.00
Std. Deviation		8.45	1140	113038	1441	26.6
Range		32	4761	502954	6174	93
Minimum		5	39	46	26	0
Maximum		37	4800	503000	6200	93
Percentiles	25	10.00	185	2800	676	8.00
	50	17.00	411	12400	1600	15.00
	75	25.00	858	46400	2600	51.50

Source: Author

4.5 Technical on-page SEO practice

How e-commerce websites follow best practices for on-page SEO was measured by the PageSpeed Insights SEO score. On a scale of 1-100, this score measures how much a website follows essential search engine optimization advice. A score of 90 or above is considered good. A score of 50 to 89 needs improvement and a score below 50 is considered poor (Google, 2023a).

Table 9 The SEO score of e-commerce websites for mobile and desktop devices

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Mobile SEO	47	31	69	100	88.70	7.99
Desktop SEO	47	30	70	100	88.34	8.41

Source: Author

An average SEO score of 88 indicates that some improvements are necessary for technical on-page SEO optimization (Google, 2023a). Each web page must have an adequate title and meta description, links should be crawlable and have descriptive text, image elements should have descriptive "alt" attributes, the HTTPS protocol is used, the code should be valid, and pages should be optimized for mobile devices.

4.6 Impact of SEO factors on traffic and sales revenue

We have used path analysis to identify the causal relationships between SEO factors and website traffic and between website traffic and company revenue and to quantify the strength of those relationships.

First, correlation analysis was conducted between every SEO factor identified through a literature review and the SEO performance indicator presented in Figure 1 to identify the most critical variables and potential causal relationships.

Knowing that the SemRush Authority Score is calculated based on three main facets: Link Power (the quality and quantity of backlinks – referring domains), Organic Traffic (an estimated monthly average of organic traffic) and Spam Factors (indicators of manipulation or spam in the link profile) (SemRush, 2023b), correlation analysis was conducted between backlinks and the Authority Score, referring domains and the Authority Score, and organic search traffic and the Authority Score.

Table 10 Correlation matrix – SEO factors/performance indicators and the Authority Score

		Authority Score
Referring Domains	Pearson Correlation	**.630****
	Sig. (2-tailed)	0
	N	47
Backlinks	Pearson Correlation	**.332***
	Sig. (2-tailed)	0.023
	N	47
Organic Search Traffic	Pearson Correlation	**.714****
	Sig. (2-tailed)	0
	N	47

Source: Author

Table 11 Correlation matrix – SEO factors and SEO performance indicators

SEO factors	SEO performance indicators >	Organic traffic	Branded organic traffic	Non-branded organic traffic	Monthly visits	Unique visitors	Pages per visit	Avg. visit duration	Bounce rate
Keywords	Pearson Correlation	**.850****	**.904****	0.012	**.919****	**.833****	**.372****	**.345***	**-.267***
	Sig. (2-tailed)	0	0	0.469	0	0	0.006	0.013	0.05
	N	47	46	46	43	43	44	42	39
Referring domains	Pearson Correlation	**.558****	**.589****	0.103	**.381***	**.455****	0.249	0.124	**-.346***
	Sig. (2-tailed)	0	0	0.247	0.006	0.001	0.051	0.216	0.016
	N	47	46	46	43	43	44	42	39
Backlinks	Pearson Correlation	**.309***	**.427****	-0.028	0.153	0.202	0.028	-0.049	-0.132
	Sig. (2-tailed)	0.017	0.002	0.427	0.163	0.097	0.43	0.38	0.211
	N	47	46	46	43	43	44	42	39
Authority score	Pearson Correlation	**.714****	**.743****	0.075	**.690****	**.725****	**.403****	**.351***	**-.397****
	Sig. (2-tailed)	0.000	0.000	0.309	0.000	0.000	0.003	0.011	0.006
	N	47	46	46	43	43	44	42	39
Mobile friendliness	Pearson Correlation	0.066	0.01	0.048	0.03	0.031	-0.198	0.047	**.401****
	Sig. (2-tailed)	0.33	0.474	0.376	0.424	0.422	0.099	0.385	0.006
	N	47	46	46	43	43	44	42	39
Mobile Performance Score	Pearson Correlation	-0.045	-0.044	0.017	0.002	0.015	-0.093	0.079	0.187
	Sig. (2-tailed)	0.383	0.385	0.454	0.496	0.463	0.274	0.309	0.127
	N	47	46	46	43	43	44	42	39
Desktop Performance Score	Pearson Correlation	-0.129	-0.07	-0.025	-0.07	-0.139	-0.2	-0.044	0.208
	Sig. (2-tailed)	0.194	0.323	0.434	0.328	0.187	0.096	0.39	0.102
	N	47	46	46	43	43	44	42	39

SEO factors	SEO performance indicators >	Organic traffic	Branded organic traffic	Non-branded organic traffic	Monthly visits	Unique visitors	Pages per visit	Avg. visit duration	Bounce rate
Domain age	Pearson Correlation	0.062	0.149	0.195	0.056	0.093	0.201	0.154	-0.142
	Sig. (2-tailed)	0.339	0.161	0.097	0.362	0.278	0.096	0.164	0.195
	N	47	46	46	43	43	44	42	39
Out-bound domains	Pearson Correlation	0.226	**.293***	0.218	**.289***	**.282***	0.033	0.223	-0.02
	Sig. (2-tailed)	0.068	0.027	0.078	0.032	0.035	0.416	0.08	0.453
	N	47	46	46	43	43	44	42	39

Source: Author

In Tables 10 and 11, bold text and one star indicate significant correlation coefficients at the 0.05 level, and two stars indicate significant correlation coefficients at the 0.01 level.

For the interpretation of the correlation coefficients shown in Tables 10 and 11, we have used a conventional approach to evaluating weak, moderate, strong, and very strong correlations, as shown in Table 12.

Table 12 Example of a conventional approach to interpreting a correlation coefficient

Absolute Magnitude of the Observed Correlation Coefficient	Interpretation
0.00–0.10	Negligible correlation
0.10–0.39	Weak correlation
0.40–0.69	Moderate correlation
0.70–0.89	Strong correlation
0.90–1.00	Very strong correlation

Source: Schober et al. (2018)

Given that the SemRush Authority Score is a score calculated through multiple factors of trustworthiness and authority, including the number of backlinks/referring domains and their quality (SemRush, 2023b), we decided to use this factor as a mediator, through which the mentioned variables indirectly affect SEO performance. Table 10 shows a significantly strong correlation between organic search traffic and the Authority Score, a moderate correlation between referring domains and the Authority Score, and a weak correlation between backlinks and the Authority Score. As SemRush also takes organic search traffic into Authority Score calculation, we want to highlight that the literature review identified the variables backlinks and referring domains as SEO factors. In contrast, organic search traffic is the SEO performance indicator that measures website visits due to the influence of SEO factors: SEO optimization. In order to satisfy the assumptions for conducting path analysis (Land, 1969), especially the exogeneity principle, organic search traffic, a measure of website visits (a SEO performance indicator similar to monthly visits), was omitted as a predictor variable that affects the authority score. Furthermore, given that the list of SEO performance indicators presented in Table 11 contains several indicators for measuring website traffic (organic search traffic, unique visitors and monthly visits) and that the research was conducted on a relatively small sample of e-commerce websites in Bosnia and Herzegovina, in order to reduce the number of variables in the path model, we decided to use only one website traffic indicator: the number of monthly visits.

Correlation analysis did not confirm the existence of any significant correlation between backlinks and SEO performance indicators: monthly visits, pages per visit, visit duration and a bounce rate. The same was determined for the SEO factors of the mobile performance score, the desktop performance score, and domain age, and those factors were excluded from further path analysis.

We started path analysis with an initial model that included five SEO factors (keywords, referring do-

mains, the Authority Score, mobile friendliness and outbound domains) presented in Table 11. Every selected SEO factor has a significant positive correlation with at least one of the SEO performance indicators (monthly visits, pages per visit, visit duration and bounce rate). The initial path model did not have a good model fit, so we improved the model fit based on modification indices by correlating error terms e3-e4 and e6-e7.

Figure 2 Path model
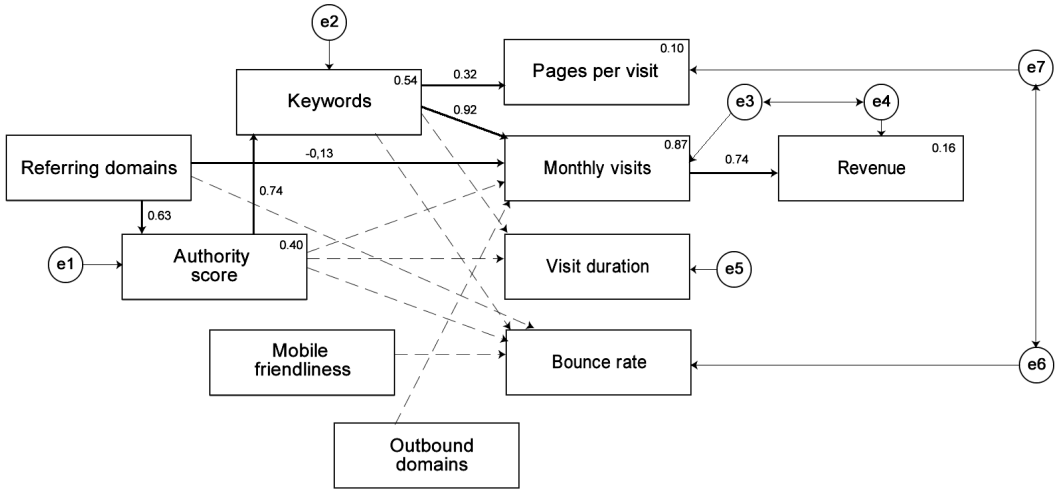
Source: Author

Figure 2 presents a path diagram of the final causal model we developed. Independent and dependent variables are shown in rectangles, and error terms are shown in circles. Above each straight solid line with an arrow there is a standardized regression parameter for that path, and below the upper right corner of each dependent variable there is a squared multiple correlation (the coefficient of determination) for each variable. Solid lines present paths with significant standardized regression parameters, and dashed lines present paths with non-significant standardized regression parameters at $p < 0.05$.

Table 13 shows the expected model fit indicators according to Hu and Bentler (1995) and Joreskog and Sorbom (1982) and the measured indicators of the proposed initial and final models after correlating error terms e3-e4 and e6-e7. The final model shows a good fit for all parameters.

Table 13 Model fit indices

Indices	Expected value	Initial model value	Final model value
X2	> 0.05	0.000	0.566
X2/df	< 3	3.347	0.930
CFI	> 0.90 (0.95 good fit)	0.727	1.000
NFI	> 0.90	0.672	0.915
RMSEA	< 0.10 (0.06 good fit)	0.226	0.000
PCLOSE	> 0.05	0.000	0.707

Source: Author, based on Hu and Bentler (1995) and Joreskog and Sorbom (1982)

The standardized regression coefficient of 0.63 between the number of referring domains and the authority score indicates a moderately strong positive relationship between these two variables. For every

standard deviation increase in the number of referring domains, the authority score tends to increase by 0.63 standard deviations. The coefficient of determination (0.40) suggests that 40% of the variability in the authority score of e-commerce websites can be attributed to the number of domains linking to the website. Correlation analysis presented in Table 10 shows a very weak correlation between backlinks and organic search traffic and the absence of a significant correlation between backlinks and monthly visits. A website with backlinks from a wide range of reputable referring domains will likely be seen as more trustworthy and authoritative by search engines.

Our model indicates that keywords mediate the influence of the authority score on monthly visits. A strong positive relationship between the authority score and keywords, with a standardized coefficient of 0.74, suggests that websites with higher authority scores tend to rank for more keywords. The coefficient of determination of 0.54 means that the authority score explains 54% of the variance in the number of keywords bringing users to the website via Google's top 100 organic search results. E-commerce websites with higher authority scores may rank for more competitive keywords, leading to increased visibility in SERPs and, consequently, more monthly visits. This finding underscores the role of the "authority score" in influencing the domain's organic search visibility. Building and maintaining authority in SEO efforts with a well-optimized set of keywords can substantially impact the number of monthly visits a website receives.

We have confirmed a very strong positive relationship between keywords and monthly visits. The standardized coefficient of 0.92 indicates that websites ranking for a larger number of keywords tend to attract more monthly visits. The coefficient of determination suggests that 87% of the variance in monthly visits is explained by the number of keywords bringing users to the website via Google's top 100 organic search results.

Finally, our model shows a moderately strong positive relationship between e-commerce website monthly visits and company revenue. For every standard deviation increase in monthly visits, there tends to be a 0.74 standard deviation increase in revenue. Furthermore, 16% of the variance in company revenue is explained by e-commerce website monthly visits. Although we expect a higher impact of e-commerce store visits on revenue, there

is a logical explanation for these results. Most of the analyzed e-commerce websites in Bosnia and Herzegovina are just one of the sales channels these companies use to sell products and generate income. Through interviews with managers of several companies included in our analysis, we have found that significant sales are realized through traditional stores and specialized online classified websites such as www.olx.ba. The analysis used data on the companies' total revenue, rather than just the sales revenue generated exclusively through the website. Although e-commerce websites can indirectly influence sales in brick-and-mortar stores or through other online channels by providing information to consumers, we see this as one of the limitations of the research that could be avoided in future research through an additional survey of the companies to determine the share of income generated through their e-commerce website.

5. Conclusion

To the best of the author's knowledge, this empirical research represents the first research of this kind on the sample of e-commerce websites in Bosnia and Herzegovina, as well as in the countries of the wider region. We have successfully provided answers to all research questions. The state of SEO optimization of e-commerce websites that sell consumer electronics in Bosnia and Herzegovina was determined. SEO factors with a significant impact on SEO performance indicators and indirectly on the sales revenue were determined. At the same time, clear recommendations were given to managers for the improvement of certain SEO areas.

The average share of organic search visits to analyzed e-commerce websites in Bosnia and Herzegovina is 23.6%, i.e., 4% less than the European average for retail and e-commerce websites (SemRush, 2023a). Branded organic search visits account for, on average, 45.8% of total organic search visits, with 50% of the analyzed websites achieving more than 48.6% of organic search visits thanks to branded organic searches. These data tell us that the most significant number of visits to the analyzed e-commerce websites is due to brand awareness, which affects direct visits and branded organic search visits. At the same time, it is a message to all e-commerce companies that one of the key tactics for increasing the number of website visitors should be investing in promotion to build and strengthen brand recognition.

The state of optimization of e-commerce websites for mobile devices requires significant improvements since 6.4% of websites need to be mobile-friendly optimized, 75.4% have a poor mobile performance score, and the remaining 25.5% have a mobile performance score that needs improvement. In order to positively influence the user experience, all analyzed companies must pay considerable attention to improving the speed and efficiency of loading websites. The results are similar for desktop devices, too, since 70.2% of websites have a performance score that needs improvement, and 17% have a poor performance score. In improving user experience, particular focus must be on technical optimization of web pages to improve the LCP (Largest Contentful Paint) and CLS (Cumulative Layout Shift) metrics.

The analysis of visitor behavior showed that they spend an average of 508 seconds on e-commerce websites, viewing an average of 2.67 pages. The average number of pages viewed per visit, 48% lower than the retail industry average (Contentsquare, 2023), is partly due to a high bounce rate of 61.4%. Companies should analyze the bounce rate according to different sources to determine which content needs to be optimized according to search keywords and which referrals provide low-quality backlinks. User retention should be increased through improved navigation, simplified search methods, offering personalized recommendations, and using social proof such as customer reviews and ratings.

The average number of outbound links on the analyzed websites is 28.7. Having in mind that many practitioners agree that 3-5 outbound links per 1,000 words of content are a good rule of thumb, it is clear that e-commerce websites should provide more links to useful websites, such as manufacturer websites, product information pages, and product reviews or buyer guides. In this way, they can contribute to building expertise, authoritativeness, and trustworthiness - essential features that determine SEO success.

Path analysis determined that the number of referring domains and their quality strongly affect the website authority score. A significant influence of the number of backlinks on SEO performance indicators (traffic and user engagement) has not been proven. This finding implies a managerial recommendation regarding the link-building off-site search engine optimization strategy. From the SEO point of view, it is much more important to have one link from ten different domains than ten links or more from the same domain, assuming that the websites have the same backlink quality. When considering potential link exchange partners, plac-

ing the link on only one partner web page, such as the home page, is sufficient.

Furthermore, it was determined that the keywords mediate the authority score impact on the number of monthly visits. The authority score has a significant strong and positive influence (0.74) on the number of keywords bringing users to a website via Google's top 100 organic search results; that is, it positively affects the website's organic search visibility. This finding confirms the claims about the reliability of the SemRush Authority Score for assessing the competitiveness and visibility of a website in search engine results, which should lead to a higher number of monthly visits, which is in line with the claims of Reyes-Lillo et al. (2023).

Finally, visits have a strong positive significant impact (0.74) on the revenue of e-commerce websites, and 16% of the variance in company revenue is explained by e-commerce website monthly visits.

Path analysis did not show a significant influence of website mobile friendliness or page speed on any SEO performance indicator, contrary to the recommendations of Google and the findings of earlier research (Gao et al., 2017; Bartuskova & Krejcar, 2015; Mustafa et al., 2015; Egri & Bayrak, 2014). One possible explanation is that the websites analyzed have similar mobile friendliness scores and speed performance indicators, and there may not be enough variability in the data to detect a significant influence. However, a cross-sectional study on a larger sample should verify this.

We did not find a direct significant impact of the number of outbound domains or domain age on website traffic, contrary to the findings of some earlier research (Iqbal et al., 2022; Nath & Ahuja, 2014). Companies like SemRush are known not to provide enough information about their algorithm and the exact way the authority score is calculated to minimize the possibility of manipulation. Therefore, in future research, it would be desirable to check whether there is a positive significant influence of outbound links and domains, domain age, and mobile and desktop performance scores on the authority score.

Some of the research limitations can also provide ideas for future research. Some recommendations are to use a larger sample, samples of e-commerce companies from other niches, and other tools for collecting analytical data. The model could be improved further by collecting data on sales revenue generated exclusively from the website rather than the total revenue used in this study. Furthermore, the model could be extended by analyzing the impact of social media signals on SEO results.

REFERENCES

1. Alfiana, F., Khofifah, N., Ramadhan, T., Septiani, N., Wahyuningsih, Azizah, N. N. & Ramadhona, N. (2023). Apply the Search Engine Optimization (SEO) Method to Determine Website Ranking on Search Engines. *International Journal of Cyber and IT Service Management*, *3*(1), 65-73. https://doi.org/10.34306/ijcitsm.v3i1.126

2. Almeida, F. & Monteiro, J. (2017). The Role of Responsive Design in Web Development. *Webology*, *14*(2), 48-56.

3. Bartuskova, A. & Krejcar, O. (2015). Loading Speed of Modern Websites and Reliability of Online Speed Test Services. In Núñez, M. et al. (Eds.). *Proceedings of Computational Collective Intelligence: 7th International Conference* (pp. 65-74). Cham: Springer. https://doi.org/10.1007/978-3-319-24306-1_7

4. Baye, M. R., De los Santos, B. & Wildenbeest, M. R. (2014). *Search Engine Optimization: What Drives Organic Traffic to Retail Sites?* Available at SSRN: https://ssrn.com/abstract=2277077 or http://dx.doi.org/10.2139/ssrn.2277077

5. Bhandari, R. S. & Bansal, A. (2018). Impact of Search Engine Optimization as a Marketing Tool. *Jindal Journal of Business Research*, *7*(1), 23-36. https://doi.org/10.1177/2278682117754016

6. Bhatnagar, A. & Ghose, S. (2004). Segmenting consumers based on the benefits and risks of Internet shopping. *Journal of Business Research*, *57*(12). 1352-1360. https://doi.org/10.1016/S0148-2963(03)00067-5

7. BrightEdge (2019). *BrightEdge Channel Report for 2019.* https://www.brightedge.com/resources/re-search-reports/channel_share

8. Chakrabortty, K. & Jose, E. (2018). Relationship Analysis between Website Traffic, Domain Age and Google Indexed Pages of E-commerce Websites. *IIM Kozhikode Society & Management Review*, *7*(2), 171-177. https://doi.org/10.1177/2277975218770028

9. Chotikitpat, K., Nilsook, P. & Sodsee, S. (2015). Techniques for improving website rankings with search engine optimization (SEO). *Advanced Science Letters*, *21*(10), 3219-3224. https://doi.org/10.1166/asl.2015.6503

10. Contentsquare (2023). *The 2023 Digital Experience Benchmark Explorer.* https://explore.con-tentsquare.com/2023-benchmark-en/2023-benchmark-explorer

11. CustomMediaLabs (2023). *Bounce Rates: What's Good, What's Bad, and Why You Should Give a Damn.* https://www.customedialabs.com/blog/bounce-rates/

12. Egri, G. & Bayrak, C. (2014). The role of search engine optimization on keeping the user on the site. *Procedia Computer Science*, *36*, 335-342. https://doi.org/10.1016/j.procs.2014.09.102

13. Fuxue, W., Li, Y. & Yiwen, Z. (2011). An Empirical Study on the Search Engine Optimization Tech-nique and Its Outcomes. In *Proceedings of the 2011 2nd International Conference on Artificial Intel-ligence, Management Science and Electronic Commerce (AIMSEC)* (pp. 2767-2770). IEEE. https://doi.org/10.1109/AIMSEC.2011.6011361

14. Ganguly, S. (2022). Correlation of SEO score and domain authority with NIRF ranking for top engi-neering institutes in India. *MetaArXiv preprints.* https://doi.org/10.31222/osf.io/crhxu

15. Gao, Q., Dey, P. & Ahammad, P. (2017). Perceived performance of top retail webpages in the wild: Insights from large-scale crowdsourcing of above-the-fold qoe. In *Internet QoE '17: Proceedings of the Workshop on QoE-based Analysis and Management of Data Communication Networks* (pp. 13-18). https://doi.org/10.1145/3098603.3098606

16. Giomelakis, D. & Veglis, A. (2015). Investigating Search Engine Optimization Factors in Media Web-sites. *Digital Journalism.* *4*(3), 379-400. https://doi.org/10.1080/21670811.2015.1046992

17. Google (2023a, May 10). *About PageSpeed Insights.* https://developers.google.com/speed/docs/insights/v5/about

18. Google (2023b, May 23). *Search Engine Optimization (SEO) Starter Guide.* https://developers.google.com/search/docs/fundamentals/seo-starter-guide

19. Google (2023c, September 27). *A guide to Google Search ranking systems.* https://developers.google.com/search/docs/appearance/ranking-systems-guide

20. Google (2023d, October). *How results are automatically generated.* https://www.google.com/search/howsearchworks/how-search-works/ranking-results

21. Google (2023e, October). *Webmaster Guidelines.* https://developers.google.com/search/docs/advanced/guidelines/webmaster-guidelines

22. Hu, L. T. & Bentler, P. M. (1995). Evaluating model fit. In Hoyle, R. H (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications* (pp. 76-99). Sage.

23. Hui, Z., Shigang, Q., Jinhua, L. & Jiahli, C. (2012). Study on website search engine optimization. In *Proceedings of the 2012 International Conference on Computer Science and Service System* (pp. 930-933). IEEE Computer Society. https://doi.org/10.1109/CSSS.2012.236

24. Iqbal, M., Khalid, M. N, Manzoor, A., Abid, M. M. & Shaikh, N. A. (2022). Search Engine Optimization (SEO): A Study of important key factors in achieving a better Search Engine Result Page (SERP) Position. *Sukkur IBA Journal of Computing and Mathematical Sciences*, *6*(1), 1-15. https://doi.org/10.30537/sjcms.v6i1.924

25. Joreskog, K. & Sorbom, D. (1982). Recent developments in structural equation modeling. *Journal of Marketing Research*, *19*(4), 404-416. https://doi.org/10.1177/002224378201900402

26. Khan, M. N. A. & Mahmood, A. (2018). A distinctive approach to obtain higher page rank through search engine optimization. *Sādhanā*, *43*(3), 43. https://doi.org/10.1007/s12046-018-0812-3

27. Killoran, J. B. (2013). How to Use Search Engine Optimization Techniques to Increase Website Visibility. *IEEE Transactions on Professional Communication*, *56*(1), 5066. https://doi.org/10.1109/TPC.2012.2237255

28. Krrabaj, S., Baxhaku, F. & Sadrijaj, D. (2017). Investigating search engine optimization techniques for effective ranking: A case study of an educational site. In *Proceedings of the 2017 6th Mediterranean Conference on Embedded Computing (MECO)*. IEEE. https://doi.org/10.1109/MECO.2017.7977137

29. Land, K. C. (1969). Principles of Path Analysis. *Sociological Methodology*, *1*, 3-37. https://doi.org/10.2307/270879

30. Lee, S., Ungson, G. R. & Russo, M. V. (2011). What determines an engaging website?: An empirical study of website characteristics and operational performance. *The Journal of High Technology Management Research*, *22*(1), 67-79. https://doi.org/10.1016/j.hitech.2011.04.002

31. Li, K., Lin, M., Lin, Z. & Xing, B. (2014). Running and chasing – The competition between paid search marketing and search engine optimization. In *Proceedings of the 2014 47th Hawaii International Conference on System Sciences* (pp. 3110-3119). IEEE. https://doi.org/10.1109/HICSS.2014.640

32. Luh, C.-J., Yang, S.-A. & Huang, T.-L. D. (2016). Estimating Google's search engine ranking function from a search engine optimization perspective. *Online Information Review*, *40*(2), 239-255. https://doi.org/10.1108/OIR-04-2015-0112

33. Mahmutović, K. (2021). Digitalni marketing: strategije, alati i taktike. University of Bihać, Faculty of Economics.

34. Marszałkowski, J., Marszałkowski, J. M. & Drozdowski, M. (2014). Empirical study of load time factor in search engine ranking. *Journal of Web Engineering*, *13*(1-2), 114-128.

35. Matošević, G., Dobša, J. & Mladenic, D. (2021). Using Machine Learning for Web Page Classification in Search Engine Optimization. *Future Internet*, *13*(1), 9.

36. Mihajlija, M. & Walton, P. (2019, June, 11). *Cumulative Layout Shift (CLS).* https://web.dev/articles/cls

37. Mustafa, R. U., Nawaz, M. S. & Lali, M. I. (2015). Search engine optimization techniques to get high score in SERP's using recommended guidelines. *Science International*, *27*(6), 5079-5086.

38. Mwosa Kivuti, E. (2018). An Valuation of the Impact of Keyword Frequency on Keyword Prominence. *International Journal of Online Marketing*, *8*(3), 52-70. https://doi.org/10.4018/IJOM.2018070104

39. Nagpal, M., Petersen, J. A. (2021). Keyword Selection Strategies in Search Engine Optimization: How Relevant is Relevance?. *Journal of Retailing*, *97*(4), 746-763. https://doi.org/10.1016/j.jretai.2020.12.002

40. Nath, C. & Ahuja, L. (2014). Search Engine Optimization (SEO): Improving Website Ranking. *International Journal of Engineering Research & Technology*, *3*(4), 2652-2656.

41. O'Neill, S. & Curran, K. (2011). The Core Aspects of Search Engine Optimisation Necessary to Move up the Ranking. *International Journal of Ambient Computing and Intelligence*, *3*(4), 62-70. https://doi.org/10.4018/jaci.2011100105

42. Ortiz-Cordova, A. & Jansen, B. J. (2012). Classifying Web Search Queries in Order to Identify High Revenue Generating Customers. *Journal of the American Society for Information Sciences and Technology*, *63*(7), 1426-1441. https://doi.org/10.1002/asi.22640

43. Osmani, A. & Pollard, B. (2020, May 5). *Optimize Cumulative Layout Shift*. https://web.dev/articles/optimize-cls

44. Reyes-Lillo, D., Morales-Vargas, A. & Rovira, C. (2023). Reliability of domain authority scores calculated by Moz, Semrush, and Ahrefs. *Profesional de la información*, *32*(4), e320403. https://doi.org/10.3145/epi.2023.jul.03

45. Roumeliotis, K. I., Tselikas, N. D. & Nasiopoulos, D. K. (2022). Airlines' Sustainability Study Based on Search Engine Optimization Techniques and Technologies. *Sustainability*, *14*(18). https://doi.org/10.3390/su141811225

46. SemRush (2023a). *Futureproof Your E-commerce Strategy: Exploring Key 2022 E-commerce Trends*. https://static.semrush.com/blog/uploads/files/b6/4e/b64ecf23529acf20d940c3e826c35bd8/futureproof_your_ecommerce_strategy.pdf

47. SemRush (2023b). *What is the Authority Score?.* https://www.semrush.com/kb/747-authority-score-backlink-scores

48. Seyfabad, M. K. & Fard, M. J. S. (2019). An Analysis of the Impact of SEO on University Website Ranking. *Iranian Journal of Information Processing and Management*, *34*(4), 1787-1810. https://doi.org/10.48550/arXiv.2009.12417

49. Schober, P., Boer, C. & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763-1768. https://doi.org/10.1213/ANE.0000000000002864

50. Sgueo, G. (2023). The Aesthetics of Consumer Tech. In *The Design of Digital Democracy* (pp. 13-47). Springer. https://doi.org/10.1007/978-3-031-36946-9_2

51. Shenoy, A. & Prabhu, A. (2016). Ranking in SEO. In *Introducing SEO.* Apress. https://doi.org/10.1007/978-1-4842-1854-9_3

52. Sinaga, H. D. E., Syamsulbahri, S., Muhajir, A. & Darmayanti, N. (2024). Analysis of the Effect of Social Media Usage, Search Engine Optimization, and Content Strategy on Business Growth in the E-Commerce Industry in Indonesia. *West Science Journal Economic and Entrepreneurship*, *2*(01), 11-18. https://doi.org/10.58812/wsjee.v2i01.647

53. StatCounter (2023, September). *Search Engine Market Share Worldwide.* https://gs.statcounter.com/search-engine-market-share

54. Su, A.-J., Hu, Y. C., Kuzmanovic, A. & Koh, C.-K. (2014). How to Improve Your Search Engine Ranking: Myths and Reality. *ACM Transactions on the Web*, *8*(2), 8. http://dx.doi.org/10.1145/2579990

55. The Business Research Company (2023, January). *Search Engine Optimization Services Global Market Report 2023.* https://www.thebusinessresearchcompany.com/report/search-engine-optimization-services-global-market-report

56. Tomasi, S. & Li, X. (2015). Influences of Search Engine Optimization on Performance of SMEs: A Qualitative Perceptive. *Journal of Electronic Commerce in Organizations*, *13*(1), 27-49. https://doi.org/10.4018/jeco.2015010103

57. Tsuei, H., Tsai, W., Pan, F. & Tzeng, G. (2020). Improving search engine optimization (SEO) by using hybrid-modified MCDM models. *Artificial Intelligence Review, 53, 1-16.* https://doi.org/10.1007/s10462-018-9644-0

58. Urosa-Barreto, F. (2020). Posicionamiento en Google: análisis SEO de los sitios web de agencias de comunicación en España. *aDResearch ESIC International Journal of Communication Research, 23*(23). https://doi.org/10.7263/adresic-023-02

59. Varangouli, E. (2020, November, 25). *Outbound Links: What Are They & Do They Affect SEO?.* https://www.semrush.com/blog/outbound-links

60. Viney, D. (2008). *Get to the Top on Google*. Nicholas Brealey.

61. Vyas, C. (2019). Evaluating state tourism websites using Search Engine Optimization tools. *Tourism Management, 73*, 64-70. https://doi.org/10.1016/j.tourman.2019.01.019

62. Walton, P. (2019, November 7). *First Input Delay (FID).* https://web.dev/articles/fid

63. Walton, P. (2020, April 30). *Web Vitals.* https://web.dev/articles/vitals#overview

64. Walton, P. & Pollard, B. (2019, August 8). *Largest Contentful Paint (LCP).* https://web.dev/articles/lcp

65. Walton, P. & Pollard, B. (2020, May 5). *Optimize the Largest Contentful Paint.* https://web.dev/articles/optimize-lcp

66. Zhang, S. & Cabage, N. (2017). Search Engine Optimization: Comparison of Link Building and Social Sharing. *Journal of Computer Information Systems, 57*(2), 148-159. https://doi.org/10.1080/08874417.2016.1183447

67. Ziakis, C., Vlachopoulou, M., Kyrkoudis, T. & Karagkiozidou, M. (2019). Important Factors for Improving Google Search Rank. *Future Internet, 11*(2), 32. https://doi.org/10.3390/fi11020032

68. Zilincan, J. (2015). Search engine optimization. In *CBU International Conference Proceedings.* Vol. 3 (pp. 506-510). https://doi.org/10.12955/cbup.v3.645

Dejan Mirčetić
The Institute for Artificial Intelligence of Serbia
University of Novi Sad
Faculty of Technical Sciences
21000 Novi Sad, Serbia
dejanmircetic@gmail.com

Svetlana Nikoličić
University of Novi Sad
Faculty of Technical Sciences
21000 Novi Sad, Serbia
cecan@uns.ac.rs

Marinko Maslarić
University of Novi Sad
Faculty of Technical Sciences
21000 Novi Sad, Serbia
marinko@uns.ac.rs

Sanja Bojić
University of Novi Sad
Faculty of Technical Sciences
21000 Novi Sad, Serbia
s_bojic@uns.ac.rs

ASSESSMENT OF TECHNOLOGICAL AND ORGANISATIONAL IMPROVEMENTS IN THE LOGISTIC DISTRIBUTION HUB USING DISCRETE EVENT SIMULATION

Abstract

Purpose: The viability of companies in today's turbulent and dynamic market depends on their ability to change and develop together with the market. Primarily under the influence of technological innovations and new customer demands, the logistics sector is under great pressure of change. The focus of this paper is on the functioning of the distribution hub as an important node in the distribution network and the operational processes it implements. The aim of the paper is to investigate different variants of the development of logistics processes in the distribution hub and thus help decision-makers in deciding on the level and performance of potential technological improvement.

Methodology: The research was conducted in a real distribution hub using a simulation approach. In particular, discrete event simulation (DES) was used, which allows supporting the dynamics of logistics processes, examining the impact of random variables on processes and determining process performance in different scenarios. Two models were formed in the paper: the As-Is model that imitates the existing processes in the distribution hub, and the To-Be model that assumes a higher technological level of the process.

Results: Findings from this paper have two practical implementations. The first one is to use As-Is models as an insight into the current situation in the distribution hub. The second implementation is to use a To-Be model as a decision support system for evaluating different solutions in the execution of logistics processes.

Conclusion: The obtained research results provide an opportunity to examine different ways of the functioning of processes in the distribution hub before making changes in the real system, which is especially important for managers of such systems.

Keywords: Distribution hub, discrete event simulations, logistics processes

1. Introduction

Logistics management activities typically include inbound and outbound transportation management, fleet management, warehousing, materials handling, order fulfillment, logistics network design, inventory management, supply/demand planning, and management of third-party logistics services providers (CSCMP, 2013). By managing the flow of goods, logistics supports and connects the functioning of various economic sectors and is one of the key factors in the development of both individual companies and national economies. At the core of these different material flows (raw materials, products, shipments, waste, etc.) are the material handling processes (loading/unloading, transhipment, transfer, packaging, labeling, formation/dismantling of cargo units, sorting, commissioning, etc.), storage and transportation, as well as that efficient logistics is of fundamental importance for the functioning of various economic sectors. In the 21st century, primarily under the influence of technological development and changed customer requirements, there is no sector in any industry that has remained untouched by major changes and challenges in business. The ambient turbulence has created a need for dynamic business processes and companies are looking for models that can evolve and adapt efficiently business processes to the changing conditions and the changing business strategies (Clempner, 2014). Different company goals, which reflect different market conditions (economic, political, social) and business priorities (capital return, customer satisfaction, employee satisfaction, company growth, cost reduction, etc.) are reflected in the structure of logistics systems and processes. In such conditions, logistics is under constant pressure of changes and faces many dilemmas aimed at improving the operational efficiency and sustainable business of companies.

The focus of this paper is on one segment of logistics, i.e. on the functioning of the distribution hub as an important node in the distribution network. Hubs are special facilities that serve as switching, transhipment and sorting points in many-to-many distribution systems (Alumur & Kara, 2008) so that intensive operational processes take place in them, which are an integral part of the functioning of the distribution system. In this paper, through a case study in the postal center as a distribution hub that exists within the postal logistics network of the Republic of Serbia, different variants of the unfolding of logistics processes are investigated in order to increase their productivity and overcome operational difficulties. Discrete event simulation (DES) was used in the research, which makes it possible to support the dynamics of logistics processes, examine the impact of randomly changing variables on processes and determine process performance in different scenarios. Two models were formed in the paper: the As-Is model that imitates the existing processes in the distribution hub, and the To-Be model that assumes a higher technological level of the process.

The remainder of this paper is organized as follows: Section 2 describes the research background and research approach. Section 3 describes the case study and the structure of the developed simulation models, and discusses the obtained results. Finally, concluding remarks are given in Section 4.

2. Background

2.1 Logistics processes in postal logistics center

The key characteristics of postal services are reflected in a massive user demand for the transfer of postal items, which is realized through postal processes that generate physical moving of mails to the recipient. From an organizational point of view, for the implementation of postal services that demand the market (national operators are conditioned to provide a universal service), it is necessary to establish a uniform postal network (on the national and international level), and use unique technologies and standardization of equipment. The structure of the postal network and its equipment is conditioned by the economic capabilities of operators, traffic volume, and in the case of a national operator, government investment policies. In recent decades, postal operators are faced with the challenges of rapid technological development, market liberalization, segmentation and increasing competition. In such conditions, the industry has evolved to include traditional post (like package and mail delivery), courier services, freight services and e-services (Chan et al., 2006) in logistics fright

flows. Accordingly, leading European postal operators (such as Deutsche Post DHL, La Poste and Royal Mail) have expanded their operations in the logistics sector. In support of these transformations is the fact that domestic letter traffic in Europe (28 countries) between 2013 and 2021 fell by about 38 percent and in the same period domestic parcel volumes increased by 122 percent (23 countries) (Publications Office of the European Union, 2022). The majority of mail in the universal postal services are letters with standard dimensions (about 70%). The decline of these services will be greater in the years to come, due to the development of electronic services. Today, electronic services are already being used that allow citizens to pay all the bills and taxes via their mobile phones, and the recent surge in e-commerce and world trade has led the parcel delivery industry to be one of the fastest growing industries (Kulkarni et al., 2021). These trends can help the Post of Serbia public enterprise plan and innovate its business. Today, competition leads companies to improve their functionality or durability of innovation and survival (Liu et al., 2020).

Postal centers represent an important link in the postal network and realize one stage in the delivery of items between the sender and the recipient. In line with the market-oriented and customer-driven development of postal operators, many traditional postal centers have expanded their range of services and grown into postal logistics centers (PLC). PLCs are located in the traffic hubs to achieve the concentration and diffusion of shipments to the geographical area that they cover, and from the logistics point of view, a PLC performs the functions of intralogistics (Lisec & Rihter, 2007). The incoming flows of shipments as independent quantities are the result of various external factors, while the outgoing flows can be seen as dependent quantities that reflect the functioning of the PLC. The technology and organization of the process in a PLC depends on the type and category of postal items and the processing differs accordingly: ordinary LC items of standard dimensions, AO items and non-standard LC items, registered and valuable items, packages, express items, etc. In general, the basic processes implemented in shipments in a PLC are: taking

postal conclusions, classification of conclusions (dismantling) according to the type of shipment, transfer to a sorting center, sorting, formation of conclusions (consolidation of the load unit), waiting for shipment, transfer to the place of shipment, loading and shipping (issuance of formed conclusions). So, the functioning of a PLC is based on fundamental logistic processes with postal items. The processes of consignments arriving in the PLC and the dispatch processes of processed consignments occur according to a defined time schedule, and the processing processes engage the resources of the PLC (space, equipment, employees) and must be fulfilled by the defined deadlines. The number and structure of postal items that are processed daily in the PLC are variables that affect the duration of the mail processing process and, due to the defined processing deadlines, the required resources as well.

2.2 Simulation modeling of logistics processes

Modeling is a problem-solving method, in which the system under study is replaced by a simple object that describes the real system or its behavior. It is based on abstraction, simplification, quantification, and analysis. According to Sterman (2000), modeling is a part of the learning process and it is iterative, a continual process of formulating hypotheses, testing, and revision, of both formal and mental models. As an object of modeling, the system is defined as an aggregation or assemblage of objects joined in some regular interaction or interdependence toward the accomplishment of some purpose (Banks et al., 2004). The model of a system is the replica of the system, physical or mathematical, which has all the properties and functions of the system (Singh, 2009).

The logistics system is a complex socio-technical system that supports the rational flow of various material goods through the implementation of logistics processes. Two approaches can generally be used to model logistics processes: analytical and simulation (Hung et al., 2006). While analytical models provide accurate and often optimal solutions, which generally imply significant simplifications (compared to the real system) and mathematical formulations complicated to solve, simulation models can support the real

characteristics of logistics processes and enable objective evaluation of performance in different functioning scenarios, which is especially important when planning new and redesigning existing processes. Simulation of the effects of redesigned processes before their implementation creates the opportunities for decision-making in relation to the planned changes. For the research in this paper, simulation models were used that respect the time dimension of the process and allow the modeling of complex processes in a logistics system such as a PLC. More precisely, from the existing approaches to simulation modeling (such as discrete-event simulation, dynamic simulations, agent-based simulations, and hybrid simulations), discrete-event simulation (DES) was used in this research for simulating operational processes in the PLC.

DES is "modeling of systems in which the state variable changes only at a discrete set of points in time" (Banks et al., 2004). The basic idea of DES is to abstract from the continuous nature of the system and consider only some "important moments", "events" in the system lifetime. DES gives an insight into the business processes of a particular system by describing the system through process maps. Process maps provide a critical assessment of what really happens inside a given company (process identification or As-Is process mapping) and they also serve for designing a new solution (To-Be process modeling). Process modeling and the evaluation of different alternative scenarios (a To-Be model) for improvement by simulation are usually the driving factors of the successful reengineering process (Trkman et al. 2007). DES is an established tool for the design and management of large-scale mail sorting and distribution systems (White et al., 2001) that allows a better understanding of the dynamic behavior of a postal system, as pieces of mail flow through from receipt to dispatch. This approach incorporates the variability and interdependence of the system and allows the analysis and understanding of both the current behavior of the system and the future as a result of numerous changes in the system (Greasley, 2003).

3. PLC case study

The case study was carried out on the example of the PLC Novi Sad (Republic of Serbia). The PLC can be considered as a post hub with the basic function of receiving, sorting, delivery and transhipment of incoming mail. In the observed PLC, different mail classes are processed (priority mail express[1], first-class mail[2] and standard mail[3] and international mail). Accordingly, during processing in the PLC, different shipments (i.e. mail classes) have a different sequence of activities through the PLC, and an average of 7,000,000 pieces of mail are processed in the PLC in one month. The focus of the research is on the quantification of the temporal characteristics of operational processes in order to analyze the engagement of employees and find opportunities to increase their productivity in order to eliminate difficulties in the PLC functioning.

3.1 Data collection

Based on the process approach and a hierarchical structure of the process, all activities and resources (employees and equipment) engaged in mail processing processes in the PLC were identified. For the observed problem, the number of employees and their time commitment to the identified activities and processes are crucial, which, in accordance with the postal operator's business policy, must be

---

1 *Priority mail express* represents service which allows fastest transfer of all kind of mail (letters, parcels, advertisements, bills, etc.). This service also provides tracking of shipments during transfer and it is available only in domestic traffic (i.e. within the boundaries of the Republic of Serbia).

2 *First-class mail* includes valuable parcels and valuable letters. By the business practice of the Post of Serbia, medium and large packages are commonly considered as valuable parcels. Valuable letters are special kinds of postal shipments that are processed in separate PLC sections, since these are confidential contracts between companies, international mail, government shipments, etc.

3 The Post of Serbia considers advertisements, circulars, newsletters, small parcels, merchandise, bills, etc. as *standard mail*. Standard postal items are classified on the basis of their dimensions, and in the practice of the Post of Serbia, there is a classification into ordinary and nonordinary mail. Ordinary mail includes letters with standard dimensions that consist of recommended letters and priority mail. Recommended letters are similar shipments as post mail express. They also allow tracking of letters and but are slower than post mail express shipments. Priority mail is mail that has priority in terms of delivery to the final consumer.

completed within precisely specified deadlines. The number and structure of shipments, organization and work technology, and time frames for the realization of processing directly affect the number of employees. The number and structure of postal items that are processed daily is a randomly variable quantity (it depends on a number of factors and primarily on the month of the year and the week of the month), whose daily volume ranges from about 15,000 to over 800,000 items.

In order to identify and quantify the time potential of PLC employees, simulation models of existing processes (As-Is) and projected processes (To-Be) have been created. Data were used to form these models from the document "Postal Statistics" (Djakovački, 2006) on the number and structure of postal items (for the selected month) downloaded from the PLC and determined by direct monitoring of the process in the PLC and the technical parameters of the sorting machine (Toshiba, 2023).

3.2 Structure of the model

In order to understand the existing processes, there is a common practice to create a process map model of existing processes in the organization. Process maps are a graphical way to describe the process, which helps to structure the information collected during the case analysis or process improvement project (Cachon & Terwiesch, 2009). For business process modeling, an academic version of iGrafx, a software package for business process mapping and simulating, was used, which is often used by researchers who have dealt with similar problems of process modeling (Trkman et al., 2007; Maslaric et al., 2012; Mircetic et al., 2013).

The DES model consists of elements connected by links that describe the flow of processes with shipments in the PLC. The basic structural elements of the diagram are the activities that take place and enable the process to take place. Each activity is defined through the type and number of resources (technical means and personnel) and the duration of the activity (constant or stochastic). In the To-Be model, activities are either modified from the aspect of time and resources needed for their realization, or they are transformed into new activities according to the requirements of the sorting machine, or they are eliminated as redundant. The paper presents the results of two models, namely:

1. As-Is model - imitates existing processing processes with postal items;

2. To-Be model - assumes a higher technological level of the process, i.e. introducing a sorting machine, which requires significant investment in equipment and employee training.

The simulation covered the period of March 2023. The relevant output from the model is the duration of the processes (active and total).

3.3 Model description and discussion of results

3.3.1 As-Is model

The process map for the As-Is model (Figure 1) presents four sectors in which incoming mail is processed according to mail type (i.e. mail class). Each mail class has different specificities and requirements for processing. Green color objects represent the generators of mail classes. The daily incoming mail distribution is highly stochastic, where the biggest picks occur at the beginning of the month.

After incoming, the mail is then directed to particular sectors according to their requirements for further processing. Grey color objects represent the activities where the duration of the activity is determined by measurement at the PLC during vocational training carried out in the PLC Novi Sad. Blue color objects represent the activities related to manual sorting of incoming mail and are based on the standard workload of the postal operator (1,000 pieces of mail per hour). Brown color objects represent the activities for which the execution time was taken from the "Postal Statistics" document (Djakovački, 2006).

Figure 1 As-Is model with manual mail sorting and processing
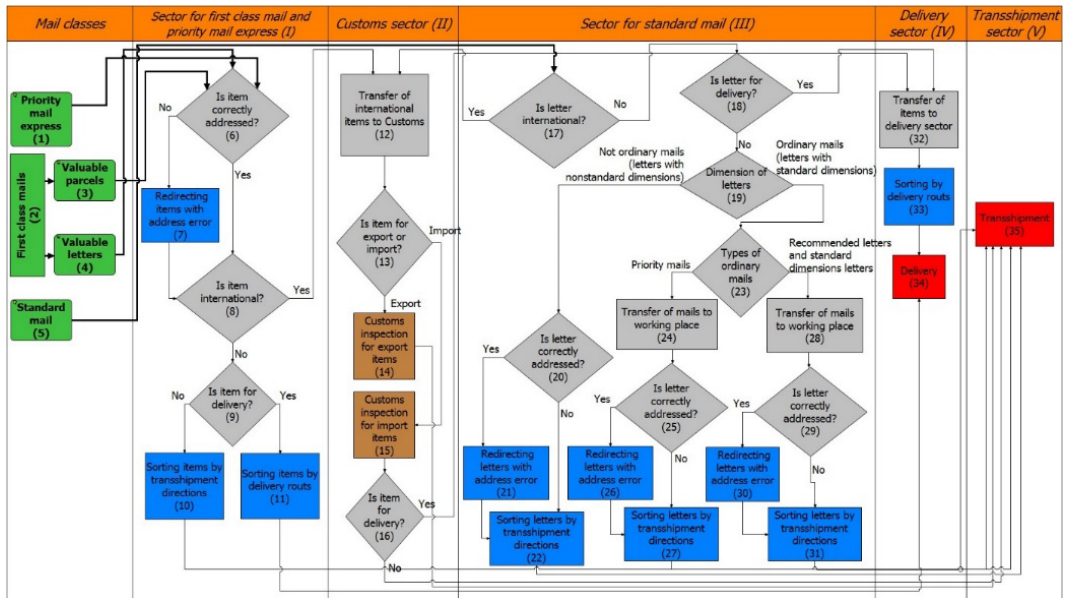
Simulating the existing activities by means of the As-Is model, mail processing times for each PLC sector are determined and presented in Table 1. It is obvious that mail spends the most time in the standard mail sector (≈54 minutes), whose task is to sort incoming mail by dividing it into mail for delivery (to local end-users) and mail for transhipment to the other PLC. The simulation also shows that in this sector mail spent considerable time waiting for resources to be served. The reason is a lack of manual labor in this sector, as well as small oper- ating capacity, which consequently causes mail to queue. As a result, mail spends time waiting for its turn to be processed.

Processing lead time is the time that elapses from the moment mail enters the PLC to the time when it is ready for delivery or transhipment (i.e. exit the PLC). The goal is to keep processing lead time as low as possible, so that the center could process all incoming mail without long delays[4]. In As-Is, the average overall processing time per one mail is ≈57 minutes (Table 1).

Table 1 Average processing time in minutes per postal mail by the As-Is model sectors

Sector	Number of pieces of mail	Average processing time	Average working time	Average time spent waiting for processing
I	146,714	1.06	0.03	1.03
II	10,471	33.27	8.01	25.26
III	5,242,636	53.83	0.09	53.75
IV	2,276,721	15.34	0.3	15.04
Total number of pieces of mail in March 2023	7,519,357	Total number of pieces of mail for transhipment		5,242,636
Processing lead time				57.49 minutes

4   Level of service defined by the Post of Serbia is that received mail should be processed within 24 hours from the moment of receipt.

Processing time through sectors is a good indicator of bottlenecks in the terminal, and the implied utilization was used to identify bottlenecks, Table 2 (column 8).

It can be seen in Table 2 that the standard mail sector has the highest implied utilization, and that other sectors have relatively small utilizations and could handle additional mail traffic. Consequently, it could be concluded that the standard mail sector represents the bottleneck of the terminal. Therefore, the maximum throughput of the entire PLC is by causality directly determined by the capacity of its weakest link, in this case, by the capacity of the standard mail sector. However, mail does not arrive evenly throughout the month. Every month there are peak load days in which the load is three times higher than the average (usually the second week of the month), when large companies deliver bills to the population (e.g. bills for electricity, cable TV, utilities, etc.), which additionally burdens PLC processes.

Table 3 shows the results of the simulation of the PLC operation on peak load days. Again, the standard mail sector has the highest processing time (it takes an average of 1,792 minutes to finish mail sorting), so the total time is also significantly higher and amounts to 1,778 minutes.

Table 2 Bottleneck identification in the As-Is model during mail processing[5]

Sector	Activities in the As-Is model	Number of workers per shift	Resource capacity[6] (pieces of mail/hour)	Resource capacity per day[7]	The average number of pieces of mail per day	Total number of pieces of mail (month)	Implied utilization[8]
I	Sorting parcels for transhipment and delivery (activities 10 and 11)	1	1,000	21,500	261.6	8,111	>1%
	Sorting letters for transhipment and delivery (activities 10 and 11)	3	3,000	64,500	4,470.1	138,574	7%
II	Sorting parcels for transhipment and delivery-import (activities 14 and 15)	1	4.6	98.8	6.2	195	6%
	Sorting parcels-export (activities 14 and 15)		9.2	197.2	4.2	131	2%
	Sorting letters for transhipment and delivery-import (activities 14 and 15)	3	20	430	196.3	6,087	46%
	Sorting letters-export (activities 14 and 15)		36	774	130.9	4,058	17%

---

5   Table 2 consists only of the activities that have the longest processing time in the observed sectors.

6   Capacity for sectors I, III, IV are taken from the literature, which recommends that one worker can process 1,000 pieces of mail/hour. Sector II is more specific than other sectors since there the goal is not to process mail as quickly as possible, but to control mail, following the government customs regulations. Accordingly, for sector II, capacity is determined by measuring the average time needed for mail inspection, and for activities (sorting parcels for transhipment and delivery-import), (sorting parcels-export), (sorting letters for transhipment and delivery-import), and (sorting letters-export), it takes 13, 6.5, 9, and 5 minutes, respectively. Therefore, resource capacity is determined by dividing one hour by the average time needed for the execution of activities. For example, 13 minutes are needed (activities 14 and 15) for the parcel sorting task, and capacity is determined as 60/13=4.6 parcels per hour.

7   Resource capacity for one day is obtained by multiplying the worker's capacity and the number of engaged workers (column 5) by 21.5 hours, which represents the total working time of employees with breaks, i.e. 3 shifts with half-hour breaks, and one hour lost in the preparation for work.

8   In logistics, the resource implied utilization is calculated by the following formula (Cachon & Terwiesch, 2009): implied utilization=current throughput/capacity*100%, or adjusted to our case study, implied utilization=average number of shipments per day/resource capacity per one day*100%.

Sector	Activities in the As-Is model	Number of workers per shift	Resource capacity[6] (pieces of mail/hour)	Resource capacity per day[7]	The average number of pieces of mail per day	Total number of pieces of mail (month)	Implied utilization[8]
III	Sorting nonordinary mail for transhipment and delivery (activities 22, 27 and 31)	2	2,000	43,000	24,946.9	773,354	58%
	Sorting priority mail for transhipment and delivery (activities 22, 27 and 31)	5	5,000	107,500	98,955.9	3,067,635	92%
	Sorting recommended letters and standard dimension letters for dispatch and delivery (activities 22, 27 and 31)	3	3,000	64,500	42,409.7	1,314,701	66%
IV	Sorting postal items for delivery (activity 33)	5	5,000	107,500	71,357.9	2,212,095	66%
Total number of workers per shift		23					
Total number of workers per day (three shifts)		69					

Source: Authors

Table 3 Average mail processing time in minutes through the As-Is model for peak load days

Sector	Number of pieces of mail	Average processing time	Average working time	Average time spent waiting for processing
I	11,416	1.09	0.03	1.06
II	393	33.95	7.96	25.99
III	1,259,967	1,791.92	0.1	1,791.81
IV	194,118	17.03	0.29	16.73
Total number of pieces of mail on peak load days	1,271,383	Total number of pieces of mail for transhipment		1,077,265
Processing lead time				1,778.45 minutes
Total time required to complete the process				5 days 20 hours and 52 minutes

Source: Authors

During the peak load days (Table 4), in the standard mail sector, the implied utilization has drastically increased, and overreached capacity in the standard mail sector, which intuitively explains a long waiting time for resources in that sector (Table 4, column 8).

Table 4 Bottleneck identification of the As-Is model during the peak load days

Sector	Activities in the As-Is model	Number of workers per shift	Resource capacity (pieces of mail/hour)	Resource capacity per day	The average number of pieces of mail per day	Total number of pieces of mail (on peak days)	Implied utilization
I	Sorting parcels for transhipment and delivery (activities 10 and 11)	1	1,000	21,500	325.5	651	2%
	Sorting letters for transhipment and delivery (activities 10 and 11)	3	3,000	64,500	5,382.5	10,765	8%
II	Sorting parcels for transhipment and delivery-import (activities 14 and 15)	1	4.6	98.8	9	18	9%
	Sorting parcels-export (activities 14 and 15)		9.2	197.2	7.5	15	4%
	Sorting letters for transhipment and delivery-import (activities 14 and 15)	3	20	430	99	198	23%
	Sorting letters-export (activities 14 and 15)		36	774	81	162	10%
III	Sorting nonordinary mail for transhipment and delivery (activities 22, 27 and 31)	2	2,000	43,000	64,245.5	128,491	149%
	Sorting priority mail for transhipment and delivery (activities 22, 27 and 31)	5	5,000	107,500	424,019.5	848,039	394%
	Sorting recommended letters and standard dimension letters for dispatch and delivery (activities 22, 27 and 31)	3	3,000	64,500	47,113.5	94,227	73%
IV	Sorting postal items for delivery (activity 33)	5	5,000	107,500	94,522	189,044	88%
Total number of workers per shift		23					
Total number of workers per day (three shifts)		69					

Source: Authors

As-Is model analysis showed that the PLC has difficulties in mail processing, especially on peak load days, due to a large number of mail and small operational capacity. The analysis also showed that the PLC is a transit center due to a large number of transhipped mail. The main indicator of the operation of each transit center is the duration of the overall processing lead time, which should be as short as possible to fulfil a specified level of service.

3.3.2 To-Be model

The To-Be model is a combination of manual and automated mail processing in the PLC. Automation involves the use of computers to control machines and processes to replace human labor. Automated processing only includes mail of standard sizes (dimensions ranging between 90x140 mm and 120x235 mm). This model was developed because the Post of Serbia public enterprise planned to purchase a machine for sorting mail of standard sizes (despite the price of this machine and a decline in the volume of mail of standard sizes).

The To-Be model is structurally equivalent to the As-Is model. The differences are in a reduced number of manual workers (from 73 to 54), and in a sorting machine allocated to the standard mail sector, which sorts only mail with standardized dimensions. Figure 2 presents the process map for the To-Be model.

Figure 2 To-Be model (automated model)
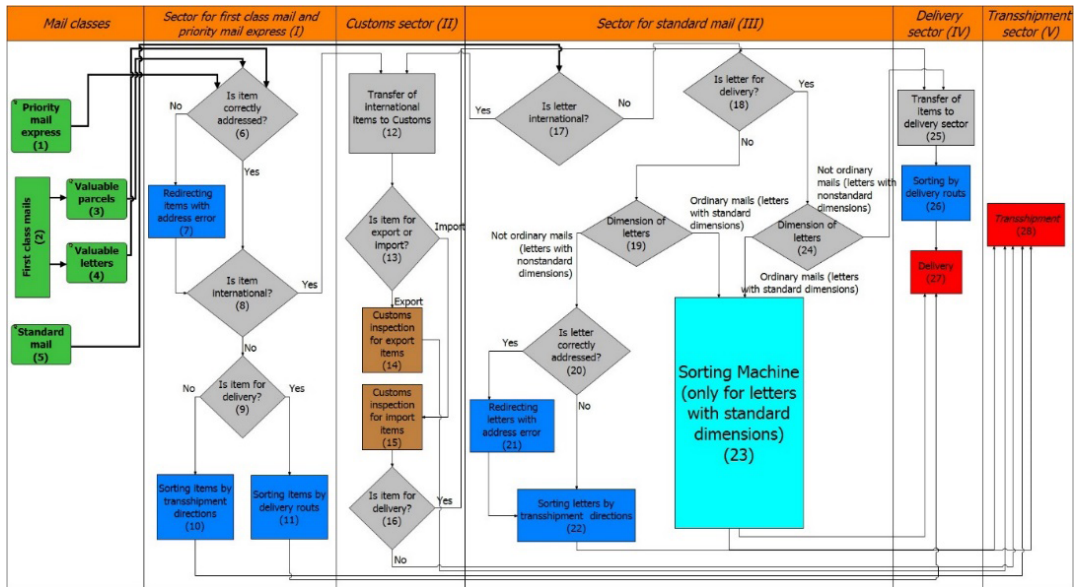
Source: Authors

The To-Be model results are summarized in tables 5 and 6. In To-Be, the longest processing time is in the customs sector, which is related to the government policies and it is not possible to make changes to government policies to optimize the PLC mail processing. The second and a much more interesting change is a big reduction in processing time in the standard mail sector caused by inserting a sorting machine.

Table 5 Average processing time per mail by To-Be model sectors

Sector	Number of pieces of mail	Average processing time	Average working time	Average time spent waiting for processing
I	146,714	1.04	0.02	1.01
II	10,471	33.27	8.01	25.26
III	7,372,646	4.45	0.03	4.41
IV	2,350,375	4.83	0.11	4.71
Total number of pieces of mail in March 2023	7,519,357	Total number of pieces of mail for transhipment		5,242,636
Processing lead time				5.93 minutes

Source: Authors

Table 6 Bottleneck identification in the To-Be model during mail processing

Sectors	Activities in the To-Be model	Number of workers per shift	Resource capacity (mails/ hour)	Resource capacity per one day	The average number of pieces of mail per day	Total number of pieces of mail	Implied utilization
I	Sorting parcels for transhipment and delivery (activities 10 and 11)	1	1,000	21,500	262	8,111	1%
	Sorting letters for transhipment and delivery (activities 10 and 11)	3	3,000	64,500	4,470	138,576	7%
II	Sorting parcels for transhipment and delivery-import (activities 14 and 15)	1	4.6	99	6	195	6%
	Sorting parcels-export (activities 14 and 15)		9.2	197	4	131	2%
	Sorting letters for transhipment and delivery-import (activities 14 and 15)	4	26.7	573	196	6,087	34%
	Sorting letters-export (activities 14 and 15)		48	1,032	131	4,058	13%
III	Sorting nonordinary mail for transhipment and delivery (activities 22 and 23)	3	3,000	64,500	21,312	660,665	33%
	Sorting machine activities (activities 22 and 23)	2	40,000	860,000	190,499	5,905,476	22%
IV	Sorting postal mail for delivery (activities 26)	5	5,000	107,500	27,944	866,272	26%
Total number of workers per shift		19					
Total number of workers per day (three shifts)		57					

Source: Authors

The use of machine dramatically reduces processing time in the standard mail sector and makes work easier. However, it can be concluded from Table 6 that monthly machine utilization is 22%. For this scenario, the simulation model results for peak load days are listed in tables 7 and 8.

Table 7 Average processing time by the To-Be model sectors for peak load days

Sectors	Number of pieces of mail	Average processing time	Average working time	Average time spent waiting for processing
I	11,416	1.1	0.03	1.07
II	393	33.77	7.95	25.82
III	1,259,967	5.7	0.03	5.67
IV	194,118	4.6	0.1	4.5
Total number of pieces of mail on peak load days	1,271,383	Total number of pieces of mail for transhipment		1,077,265
Processing lead time			6.38 minutes	
Total time required to complete the process			Within a period (2 days)	

Source: Authors

The results show that the To-Be model successfully sorts mail on peak load days and that machine capacity utilization on these days is 62%. Based on the analysis, it is determined that the To-Be model has almost five times the mail processing capacity of the actual monthly mail load.

Table 8 Bottleneck identification of the To-Be 1 model during the peak load days

Sectors	Activities in the To-Be model	Number of workers per shift	Resource capacity (mails/hour)	Resource capacity per one day	The average number of pieces of mail per day	Total number of pieces of mail	Implied utilization
I	Sorting parcels for tran-shipment and delivery (activities 10 and 11)	1	1,000	21,500	326	651	2%
	Sorting letters for tran-shipment and delivery (activities 10 and 11)	3	3,000	64,500	5,383	10,765	8%
II	Sorting parcels for tran-shipment and delivery-import (activities 14 and 15)	1	4.6	99	9	18	9%
	Sorting parcels-export (activities 14 and 15)		9.2	197	8	15	4%
	Sorting letters for tran-shipment and delivery-import (activities 14 and 15)	4	26.7	573	99	198	17%
	Sorting letters-export (activities 14 and 15)		48	1,032	81	162	8%
III	Sorting nonordinary mail for transhipment and delivery (activities 22 and 23)	3	3,000	64,500	64,246	128,491	100%
	Sorting machine activities (activities 22 and 23)	2	40,000	860,000	534,434	1,068,868	62%
IV	Sorting postal mail for delivery (activities 26)	5	5,000	107,500	31,221	62,442	29%
Total number of workers per shift		19					
Total number of workers per day (three shifts)		57					

Source: Authors

3.4 Comparative analysis of the As-Is and To-Be models

Table 9 shows a comparative analysis that helps determine which model is the best for application in the observed PLC. It can be seen in the table that the To-Be model is better, but that comes with the price since the To-Be model is significantly more expensive than As-Is, as it generates costs related to the investment in the sorting machine. These data call into question the need to automate the postal system of Serbia.

Table 9 Comparative analysis of the models

Models	Processing lead time (month) [minutes]	Processing lead time (peak load days) [minutes]	Total time required to complete the process (peak load days)	Engaged resources
As-Is	57.49	1,778.45	5 days 20 hours and 52 minutes	69 skilled workers
To-Be	5.93	6.38	On time (within 2 days)	57 skilled workers + sorting machine

Source: Authors

4. Conclusion

In business logistics micro- and meta-systems, processes can be represented as a series of related logistics activities executed using the working system resources to create value for the customers. The efficiency of the executed processes directly affects the overall efficiency of the entire logistics subject. This paper deals with this kind of problem.

In this paper, through a case study in the postal center as a distribution hub that exists within the postal logistics network of the Republic of Serbia, different variants of the unfolding of logistics processes were investigated. In order to raise productivity and overcome operational difficulties, two process models were developed, and the results showed that through organizational and structural changes of the existing processes via the To-Be model, operational problems can be eliminated sufficiently well. Besides, research results show that the use of DES in the PLC is fully justified and recommended. Research limitations are related to the implementation of reengineering changes in the To-Be model, which was beyond the scope of this paper.

Findings from this paper have two practical implementations. The first implementation is to use As-Is models as an insight into the current situation in the PLC. The second implementation is to use a To-Be model as a decision support system for evaluating different solutions in the execution of logistics processes and also for argumentation in the planning of investments in machinery and equipment.

Acknowledgment

References

1. Alumur, S. & Kara, B. Y. (2008). Network hub location problems: The state of the art. *European Journal of Operational Research*, *190*(1), 1-21. https://doi.org/10.1016/j.ejor.2007.06.008

2. Banks, J., Carson, J. S., Nelson, B. L. & Nicol, D. M. (2004). *Discrete Event System Simulation* (4th ed). Prentice Hall.

3. Cachon, G. & Terwiesch, C. (2009). *Matching Supply with Demand: An Introduction to Operations Management*. McGraw-Hill.

4. Chan, F. T., Chan, H., Lau, H. C. & Ip, R. W. (2006). An AHP approach in benchmarking logistics performance of the postal industry. *Benchmarking: An International Journal*, *13*(6), 636-661. https://doi.org/10.1108/14635770610709031

5. Clempner, J. (2014). Verifying soundness of business processes: A decision process Petri nets approach. *Expert Systems with Applications*, *41*(11), 5030-5040. https://doi.org/10.1016/j.eswa.2014.03.005

6. CSCMP (2013). *Supply Chain Management Definitions and Glossary*. https://cscmp.org/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms/CSCMP/ Educate/SCMDefinitions_and_Glossary_of_Terms.aspx?hkey=60879588-f65f-4ab5-8c4b-6878815ef921

7. Djakovački, V. (2006). *Pravilnik o statistici i normama, Javno preduzeće PTT-saobraćaja Srbija*. Službeni PTT glasnik.

8. Greasley, A. (2003). Using business-process simulation within a business-process reengineering approach. *Business Process Management Journal*, *9*(4), 408-420. https://doi.org/10.1108/14637150310484481

9. Hung, W. Y., Samsatli, N. J. & Shah, N. (2006). Object-oriented dynamic supply-chain modelling incorporated with production scheduling. *European Journal of Operational Research, 169*(3), 1064-1076. https://doi.org/10.1016/j.ejor.2005.02.009

10. Kulkarni, O., Cohen, Y. M., Dahan, M. & Montreuil, B. (2021). Resilient Hyperconnected Logistics Hub Network Design. In *Proceedings of the 8th International Physical Internet Conference*. Atlanta, GA: Georgia Institute of Technology.

11. Lisec, A. & Rihter, A. (2007). Logistical operations in postal logistics centres. *Logistics & Sustainable Transport*, *1*(2), 1-12.

12. Liu, J., Chen, C., Liu, Z., Jermsittiparsert, K. & Ghadimi, N. (2020). An IGDT-based risk-involved optimal bidding strategy for hydrogen storage-based intelligent parking lot of electric vehicles. *Journal of Energy Storage*, *27*, 101057. https://doi.org/10.1016/j.est.2019.101057

13. Maslaric, M., Groznik, A. & Brnjac, N. (2012). Distribution Channel Reengineering: A Case Study. *Promet – Traffic&Transportation*, *24*(1)**,** 35-43. https://doi.org/10.7307/ptt.v24i1.266

14. Mircetic, D., Maslaric, M. & Nikolicic, S. (2013). Reengineering of terminal processes in inland container terminal. In Pawar, K. S. & Rogers, H. (Eds.). *Proceedings of the 18th International Symposium on Logistics* (pp. 447-457). Vienna: Nottingham University Business School.

15. Publications Office of the European Union (2022). *Main developments in the postal sector (2017-2021)*. https://op.europa.eu/en/publication-detail/-/publication/e8348d3f-79d0-11ed-9887-01aa75ed71a1/language-en/format-PDF

16. Singh, V. P. (2009). *System Modeling And Simulation*. New Age International Publishers.

17. Sterman, J. (2000). *Business Dynamics: System Thinking and Modeling for a Complex World*. Irwin-McGraw-Hill.

18. Toshiba (2023). *OCR Letter Sorting Machine (LSM)*. https://www.global.toshiba/ww/products-solutions/security-automation/postal-logistics/ocr-letter-sorting-machine.html

19. Trkman, P., Stemberger, M. I., Jaklic, J. & Groznik, A. (2007). Process approach to supply chain integration. *Supply Chain Management: An International Journal*, *12*(2), 116-128. https://doi.org/10.1108/13598540710737307

20. White, K. P., Barney, B., Keller, S., Schwieters, R., Villasenor, J., Terry, W. S., Fairbrother, R. G. & Saxton, R. D. (2001). An object-oriented paradigm for simulating postal distribution centers. In Peters, B. A. (Eds.). *Proceedings of the 2001 Winter Simulation Conference* (pp. 1007-1012). Arlington, VA: IEEE.

Kübra Saka Ilgin
Erzincan Binali Yıldırım University
Faculty of Economics and Admini-
strative Sciences
24050 Erzincan, Turkey
kubra.saka@erzincan.edu.tr

CAN THE MAJOR CRYPTOCURRENCIES BE USED AS A PORTFOLIO DIVERSIFIER?: ANALYSIS OF THE RELATIONSHIP BETWEEN BITCOIN, ETHEREUM, AND GLOBAL FINANCIAL ASSET CLASSES

Abstract

Purpose: It can be stated that in today's competitive conditions, where portfolio management is very important, it has become necessary to examine the relationship between global financial assets and major cryptocurrencies, such as Bitcoin and Ethereum. This paper aims to investigate the cointegration and causality relationships between Bitcoin, Ethereum, and global financial assets such as gold, oil, the S&P Global 100, the Dow Jones Commodity, and the US Dollar Indices, and to determine the diversification role of Bitcoin and Ethereum comparatively for the period between April 2016 and January 2024.

Methodology: The ADF Unit Root, Johansen Cointegration, Granger Causality, Rolling Window Causality tests, and Variance Decomposition Analysis methods were used in the analysis process.

Results: Based on the findings obtained from the paper, it was determined that Bitcoin and Ethereum have no cointegration with selected financial asset classes. Granger causality analysis results indicated that there were unidirectional causalities from Bitcoin and Ethereum prices to Dow Jones Commodity Index prices. In addition to the results of the Rolling Window causality tests, it was also determined that there are some causalities between Bitcoin, Ethereum, and other variables, especially after the 2021-2022 period.

Conclusion: It can be concluded that Bitcoin and Ethereum are effective portfolio diversifiers throughout the entire period; however, the diversification effects of Bitcoin and Ethereum weakened towards the end of the review period. Therefore, it can be said that Bitcoin and Ethereum act similarly in the global investment portfolio.

Keywords: Bitcoin, Ethereum, global financial asset classes, diversifier, causality

1. Introduction

The authority of central banks to control the amount of banknotes in the market can lead to negative economic conditions such as inflation. The Central Bank of the USA reduced the interest rates to stimulate the US economy in 2008, which unexpectedly led to significant losses. The effects of the 2008 global crisis that emerged in the same year spread all over the world; a large number of companies from various sectors were driven into bankruptcy, which indirectly accelerated the beginning of a new currency era (Bhuiyan et al., 2021, pp. 1-2). Cryptocurrencies that emerged with these developments differ from traditional currencies because they are not issued or controlled by any government.

Bitcoin has dominated the cryptocurrency market from its inception to the present; it is a form of virtual money with an intrinsic value of zero, issued through computer code in electronic wallets, cannot be converted into anything, and does not have support of any Central Bank or any government. Bitcoin's value cannot be evaluated as a convertible material asset like gold or a currency like the dollar. It is determined by the mutual interaction of demand and supply. Since its inception in 2008, Bitcoin has gained significant international recognition due to the potential of its underlying technology to develop applications beyond currency. A new currency called Bitcoin facilitates person-to-person business transactions worldwide without the need for any intermediaries, reducing trade barriers and increasing efficiency. However, Bitcoin has always been approached with concern due to many reasons such as its highly volatile structure, speculative behavior, coding with mathematical formulas, inelastic money supply, and lack of legal security (Bouoiyour & Selmi, 2015, p. 3). Cryptocurrencies, which have become widespread with the emergence of Bitcoin, are experiencing increasing competition day by day, with new types of assets entering the market after this year. Several cryptocurrencies have been developed since then, but Bitcoin has dominated the market at all times. For instance, by mid-September 2023, Bitcoin achieved a value of over $500 billion, more than half of the total market value of cryptocurrencies, while Ethereum ranked second with approximately $190 billion (CoinMarketCap, 2023).

Developments in the cryptocurrency market have led to an increase in academic studies aimed at understanding the financial structure of Bitcoin as a currency, investment instrument, and commodity. The biggest obstacle to Bitcoin's ability to serve as a unit of account, a medium of exchange, and a means of savings—key functions it must fulfill to be accepted as a currency—is the high volatility of its prices. However, it has not reached widespread use as a currency. If Bitcoin becomes more widely used as a currency, it will compete with other fiat currencies, affecting the value of the fiat currency and ultimately central bank monetary policies. On the other hand, it can be stated that if it is used as an investment, it can compete with many other assets such as government bonds, stocks and commodities (Baur et al., 2018a, pp. 187-188).

The evaluation of Bitcoin as a new investable asset class has led to the investigation of its relationship with other financial assets and other cryptocurrencies and its adequacy as a safe haven, hedging tool and diversification tool as a financial asset. Some features distinguish assets that offer hedging, safe haven, and diversification benefits. Hedging is the situation where an asset is, on average, unrelated or negatively related to another asset or portfolio. Such an asset cannot effectively mitigate losses during periods of market pressure or financial turmoil because it may be negatively correlated, on average, with a positive correlation during such periods and a negative correlation during normal times. A diversifier is a situation where an asset has an imperfect but, on average, positive correlation with another asset or portfolio. Like the hedging feature, the diversification feature does not specifically reduce losses in extremely adverse market conditions. However, a safe haven is a situation where an asset is negatively correlated or uncorrelated with another asset or portfolio during periods of market pressure and financial turmoil. An asset with such a feature establishes a non-positive relationship with the portfolio in extremely negative market conditions and creates a safe haven for investors (Baur & Lucey, 2010, p. 219). Volatility movements in the price of Bitcoin and Ethereum, appetite for profit as a result of price changes, and curiosity factors have made them financial assets that attract the attention of investors. As major cryptocurrencies like Bitcoin and Ethereum began to provide depth in the financial markets, it became necessary to reveal the relationship between major cryptocurrencies and

other investment instruments. Studies continue to discuss whether cryptocurrencies can be considered as an investment instrument, a diversification instrument in portfolios, a hedging instrument, and whether they are a currency.

It can be stated that it has become necessary for rational investors who prioritize the efficiency of their investments to examine the relationship between the major cryptocurrencies Bitcoin, Ethereum, and other global financial asset classes in today's competitive conditions, where portfolio management is very important. In this context, this paper aims to examine the cointegration and causality relationships between Bitcoin, Ethereum, and global financial asset classes such as gold, oil, the Dow Jones commodity, the S&P100 stock, and the US dollar indices. The research findings aimed to make recommendations to investors, financial advisors, policymakers, portfolio managers, and especially Bitcoin and Ethereum investors in determining investment horizons. The study seeks to answer the following questions: Are there cointegration and causality relationships between major cryptocurrencies—specifically Bitcoin and Ethereum—and global financial asset classes? Additionally, do major cryptocurrencies exhibit diversifying features relative to other assets, and can they serve as portfolio diversifiers throughout the selected period?

2. Literature review

According to Van Wijk (2013), the Dow Jones index makes contributes significantly to Bitcoin prices in both the short and long term, as well as the euro-dollar parity and oil prices in the long term. Briere et al. (2015) investigated the portfolio performances that included traditional assets and alternative investments, with or without Bitcoin. They applied correlation analysis, some portfolio performance measurement techniques, and spanning tests. They found that Bitcoin has a low correlation with other assets, offering diversification benefits to investors. As a result, it can be said that Bitcoin improves the risk-return balance of well-diversified portfolios. Georgoula et al. (2015) reported that time series analyses conducted to determine the relationship of Bitcoin prices with important variables, e.g. the S&P 500 index, revealed a negative relation and a good diversification alternative. Dyhrberg (2016) investigated the suitability of adding Bitcoin to a portfolio as a risk-hedging tool and argued that Bit-

coin has certain characteristics of gold in terms of its risk-hedging ability. She concluded that Bitcoin should be seen as a hedging tool in a portfolio that includes the dollar and stocks. As a result of their analysis of three commodity indices (S&P GSCI general commodity, energy commodity, and non-energy commodity indices) and Bitcoin, including Bitcoin's 2013 price collapse, Bouri, Jalkh, Molnár, and Roubaud (2017) found that before 2013 Bitcoin experienced a significant decline in relation to two indices. They stated that it had the feature of a risk-hedging tool and a safe haven, but after 2013, it only offered a diversification feature. However, they argued that it was only diversifying for the non-energy commodity index throughout the entire period. The study of Bouri et al. (2017) stated that Bitcoin can be used for diversification purposes for many asset classes consisting of stocks, bonds, currencies, and commodities, but it is a weak hedging tool. Bouoiyour and Selmi (2017) determined that Bitcoin and Ethereum have a negative relationship with oil, S&P 500, and US bonds, and that cryptocurrencies are good diversifiers. Baur et al. (2018b) found that there was no relationship between assets, including gold, paper banknotes, and commodities, and Bitcoin. Güleç et al. (2018) researched the relationship between Bitcoin and stock markets, interest rates, exchange rates, and commodity markets in Turkey. According to the analyses, a relationship was found between Bitcoin prices and interest rates, but no significant relationship was found with other variables. Henriques and Sadorsky (2018) examined the consequences of replacing gold in a portfolio with Bitcoin, using some of the GARCH models for their analysis. They concluded that the performance of portfolios that include Bitcoin is higher than the others, so Bitcoin is a good diversifier. It is essential to say that the weight of Bitcoin is considerably lower. Öztürk et al. (2018) researched the relationship between Bitcoin and some asset groups with cointegration analysis to determine whether Bitcoin may be used as a new hedging tool. As a result of the study, they found that Bitcoin moves only with gold prices and is independent of other assets. This finding indicates that Bitcoin can be a good portfolio diversification instrument. Baumöhl (2019) concluded that Bitcoin and other cryptocurrencies provide diversification benefits for Euro, Yuan, Swiss Franc, Yen and Canadian Dollar investors. Giudici and Abu-Hashish (2019) stated that the correlation between conventional assets and Bitcoin is low; Bitcoin can be used

for diversification purposes in portfolios created with gold, oil, S&P 500, Euro, and Yuan. Kajtazi and Moro (2019) investigated the effects of including Bitcoin in US, European, and Chinese asset portfolios. Analysis results showed that performance increased in portfolios where Bitcoin was included, and this was due to an increase in returns rather than a decrease in volatility. Although Bitcoin has speculative features, it has been stated that it may be a good portfolio diversifier. Kliber et al. (2019) investigated which Bitcoin has hedging, safe haven, and diversification features against the stock index of five countries. Their analysis concluded that Bitcoin is a safe haven in Venezuela, a diversifier in Japan and China, and a weak hedging tool in Sweden and Estonia. Akhtaruzzaman et al. (2020) examined the effect of Bitcoin's diversification on global industry portfolios and bond indexes. They applied the VARMA DCC-GARCH method. According to the results, lower correlations were found between the variables so it can be said that Bitcoin is a hedge instrument. Bouri et al. (2020) examined the safe haven and hedging properties of cryptocurrencies during a decline in ten stocks. They argued that cryptocurrencies are valuable digital assets, but there is significant heterogeneity among them. They stated that Bitcoin is a safe haven against all US stocks, while some cryptocurrencies can be used as hedging tools for several sectors. Charfeddine et al. (2020) examined the diversification and hedging properties of Bitcoin and Ethereum against S&P500, gold, and oil prices. They concluded that cryptocurrencies are suitable for financial diversification, risk protection features remain weak, and the relationship between cryptocurrencies and traditional assets is affected by external shocks. Das et al. (2020) compared the qualities of Bitcoin as a safe haven and a hedging tool with gold, commodities and the US dollar. As a result of their analysis, they stated that the hedging and safe haven features of each asset differ for different economic conditions and market situations; therefore, it is not possible to achieve both features under all conditions with a single asset. Dutta et al. (2020) found that both Bitcoin and gold can be safe havens for hedging or diversification purposes during oil price volatility. Kang et al. (2020) investigated the relationship between gold futures, the US dollar, US stocks (S&P 500), and treasury bills in a portfolio, and Bitcoin. Asymmetric Granger causality was found between Bitcoin and gold, indicating a relationship between

the two. The findings suggest that Bitcoin can be used as a safe haven by investors to reduce risk and provide diversification benefits in portfolio risk management. Bakry et al. (2021) examined the diversifier effect of Bitcoin with different portfolio choices. They applied the Sharpe ratio for portfolio optimization. The results of the paper showed that Bitcoin acts as a diversifier and hedge for risk-seeking investors, especially in relation to safe havens. Huang et al. (2021) analysed the diversification effect of Bitcoin in the periods before and after COVID-19. They determined that Bitcoin contributes to diversification benefits to traditional assets among different major economies but the pandemic has altered the diversification role of Bitcoin in the markets, with the exception of the United States of America. Qarni and Gulzar (2021) examined the diversification effect of Bitcoin against currency market portfolios. They applied the spillover index and frequency connectedness methods. According to the results, there was a low level of integration and asymmetric volatility spillover between Bitcoin and other currency pairs. Bitcoin is found to provide significant diversification benefits to other currency portfolios, especially euro portfolios. Maghyereh and Abdoh (2022) examined the volatility connectedness between Bitcoin and traditional financial assets during the COVID-19 period. The findings of the analysis indicated that the volatility dynamics were negative and weak before the pandemic and positive during the pandemic. The volatility connectedness of bitcoin-gold and bitcoin-foreign exchange pairs is significant in the short term, but bitcoin-oil and bitcoin-stock pairs are significant in the intermediate term. Bhuiyan et al. (2023) investigated the performance of Bitcoin with stock markets advanced economies and its diversification potential by wavelet analysis in the 2014-2022 period. The results indicated that Bitcoin is a diversifier against gold in all indices; Bitcoin showed a low increase with all indices, with the exception of gold, especially in the short term. Bouri et al. (2023) investigated the relationship between Bitcoin and US stock markets in the 2017-2021 period. The results showed that Bitcoin prices have significant predictability for US stock volatility. Hanif et al. (2023) investigated the connectedness between cryptocurrencies (Bitcoin and Ethereum), stock markets, and gold and oil prices by using time-frequency models in the 2020-2022 period. They concluded that cryptocurrencies, stock markets and

commodities are highly interconnected in terms of volatility. Özbek (2023) examined the relationship between Bitcoin and BIST100, and S&P500 indexes to assess Bitcoin's features as a diversification or hedging instrument from 2020 to 2022. The results indicated no relationship between Bitcoin and the BIST100 index over the whole period, while a relationship was observed between Bitcoin and the S&P500 index in three-fifths of the period. It can be said that Bitcoin is a good diversifier but it loses this feature towards the end of the period for S&P500 investors.

It can be seen in the literature that there are many studies examining the interaction between Bitcoin and other investment instruments and the diversifying role of Bitcoin. However, in these studies, the existence of the relationship between Bitcoin and investment instruments, the existence and direction of the causality relationship, and the difference in the study results regarding the diversifying role of Bitcoin were the motivations for research presented in this paper. Comparative analysis of the findings in terms of Bitcoin and Ethereum, as well as the use of two different methods in causality analysis, such as the Granger and Rolling Window Causality test, aimed at determining whether the causality relations between the variables and diversifier effects differ during the review period, distinguish the paper from similar ones and it can be stated that the paper will contribute to the literature in this direction.

3. Data

The relationship between global financial asset classes and Bitcoin and/or Ethereum was examined in the study. Gold prices, the US dollar index, WTI crude oil prices, Dow Jones Commodity, and S&P Global 100 were used as financial asset classes. The data set consists of daily data for the period from April 2016 to January 2024. Spot gold price (dollar/ounce) shows the gold price in dollars per ounce. Logarithmic transformation was applied to the data used in the study. The variables used in the study, their abbreviations, and the sources they were obtained from are given in Table 1.

Table 1 Variables used in the study

Variable	Abbreviation	Source
Bitcoin Prices	BTC	Coinmarketcap
Ethereum Prices	ETH	Coinmarketcap
Gold Prices	GLD	World Gold Council[1]
US Dollar Index	USDX	Investing[2]
Oil Prices	WTI	US Energy Information Administration[3]
Dow-Jones Commodity Index	DJCI	Investing
S&P Global100 Index	SP100	Investing

Source: Author

Time-dependent oscillation graphs of the variables used in the study are given in Figure 1.

---

1   World Gold Council, https://www.gold.org

2   Investing, https://tr.investing.com

3   US Energy Information Administration, https://www.eia.gov

Figure 1 Time-dependent oscillation graphs of variables

Source: Author's calculations

As can be seen in Figure 1, Bitcoin and Ethereum prices started to increase in 2017. Prices, which decreased towards the end of 2018, started to increase again after 2019 and peaked. The fact that Bitcoin and Ethereum were unusual investment tools when they were first introduced contributed to their initial lack of demand. However, sharp price increases that occurred later along with the increasing demand caused it to be seen as an important investment tool by investors. The developments in the course of other global financial asset classes discussed in the study as alternatives to Bitcoin and Ethereum are also shown in Figure 1. There are sharp decreases in stock indices, oil prices, and the dollar index. Sharp increases in gold, Bitcoin, and Ethereum prices are noteworthy, especially during the COVID-19 pandemic period.

4. Research method and findings

The presence of unit roots in econometric analysis shows that a time series is not stationary. Since non-stationary time series will cause spurious regression problems in the analyses, stationarity testing must be done and non-stationary series must be made stationary (Gujarati, 1999, p. 713). Unit root tests commonly used in stationarity testing of time series are Augmented Dickey and Fuller - ADF (1979), Phillips (1987) and Perron (1988) - PP, and Kwiatkowski, Phillips, Schmidt, and Shin (1992) - KPSS. In the study, the stationarity of the time series was tested using the ADF unit root test. Distribution theory supporting the ADF test assumes that the error terms are statistically independent and have a constant variance. The regression equation used for the ADF unit root test is equation (1) (Mushtaq, 2011, pp. 10-11). The fact that the ADF-t statistical values obtained as a result of the ADF unit root test are greater than the MacKinnon critical values in absolute terms indicates that the time series are stationary. Otherwise, the time series must be differentiated to ensure their stationarity:

$$\Delta Y_t = \beta_0 + \beta_1(t-T/2) + \beta_2 Y_{t-1} \sum_{i=1}^{m} \Delta Y_{t-i} + u_t \ . \ (1)$$

Determining whether the time series are stationary at the same level is a prerequisite for investigating cointegration, i.e. the long-term relationship, between these series. Long-term relationships between time series can be investigated with the cointegration test developed by Johansen and Juselius (1990). The Johansen cointegration test is a maximum likelihood approach applied to deter-

mine the presence of cointegration vectors in time series. Additionally, this method is based on linear vector autoregression (VAR) (Balke & Fomby, 1997, p. 636). The regression equations used for the Johansen cointegration test are equations (2) and (3):

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-n} + \prod \Delta X_{t-n} + \varepsilon_t \tag{2}$$

$$\Gamma_i = -1 + \prod_1 + \cdots + \prod_i \qquad i:1\ldots\ldots n \tag{3}$$

VAR (Vector Autoregression) models based on the appropriate lag length must first be established in order to apply the Johansen cointegration test (Johansen and Juselius, 1990, p. 170). Trace and maximum eigenvalue statistics obtained by the cointegration test show whether there is a cointegration between the variables. In case there is a cointegration vector that indicates the presence of a long-term relation between time series, a VECM (Vector Error Correction Model) model should be established and it should be examined whether there is a short-term relationship.

Statistically, causality refers to the ability to estimate future values of a time series variable based on the influence of its past values or those of another related time series variable (Işığıçok, 1994, p. 94). Causality in the sense of Granger means that if past values of variable X improve the prediction accuracy of variable Y, then X is said to Granger-cause Y. If a cointegration relationship exists between the variables, a VECM-based Granger causality analysis is applied; if there is no cointegration relationship, a VAR-based Granger causality analysis is used. Granger causality analysis expresses the direction of relationships between time series. The regression equations used for the Granger causality test are equations (4) and (5):

$$X_t = \sum_{i-1}^{n} \alpha_i Y_{t-i} + \sum_{i-1}^{n} \beta_i X_{t-i} + U_{1t} \tag{4}$$

$$Y_t = \sum_{i-1}^{n} \gamma_i X_{t-i} + \sum_{i-1}^{n} \theta_i X_{t-i} + U_{2t} \tag{5}$$

The Granger causality test is an analysis method in which one-way or two-way relationships between variables are investigated without distinguishing between dependent and independent variables (Tarı, 2015, p. 436).

The rolling window causality test developed by Balcilar et al. (2010) is important in terms of showing

the change in the causality relationship in different periods. The null hypothesis based on the causality relationship predicts that there is no Granger causality between the variables. According to the test results, the Bootstrap-p value must be below the 5% or 10% critical value for the null hypothesis to be rejected and the alternative hypothesis to be accepted (Balcilar et al., 2010, p. 1403). In this context, rolling window causality test results will be evaluated in terms of determining the mutual causality relationships between the dependent variables and each independent variable.

Variance decomposition analysis shows the severity of the error variance estimated for the mobility caused by each of the independent variables at different time horizons beyond the selected period (Ahad, 2017, p. 820). After determining the significant causality relationships between variables with the Granger causality test, variance decomposition analysis is applied to examine the effect level of these causalities.

In the analysis of this paper, the relationship and the causality between Bitcoin, Ethereum, and global financial asset classes was examined for the period between April 2016 and January 2024. The econometric models are as follows:

Model 1:

$$LBTC_t = \alpha_t + \beta_1 LDJC_t + \beta_2 LGLD_t + \beta_3 LSPG100_t + \beta_4 LUSDX_t + \beta_5 LWT_t + \varepsilon_t, \quad (6)$$

Model 2:

$$LETH_t = \alpha_t + \beta_1 LDJC_t + \beta_2 LGLD_t + \beta_3 LSPG100_t + \beta_4 LUSDX_t + \beta_5 LWT_t + \varepsilon_t, \quad (7)$$

where α_t is the constant coefficient, β_t is the slope coefficient, t are the periods, and ε_t is the error term in equations 6 and 7.

In time series analyses, firstly, the ADF unit root test was applied to ensure that there is no spurious regression between the variables, in other words, to check the stationarity of the series. The Johansen cointegration test was applied to the series that were stationary at their first differences. Then, the Granger causality test was applied to test the existence of causality between asset classes and Bitcoin, and Ethereum. It is aimed to strengthen the findings by reanalyzing the causal relationships between variables in terms of different periods with rolling window causality analysis. The variance decomposition test was applied to determine the extent of the effect of this causality on Bitcoin and Ethereum, if causality was present. ADF unit root test results are given in Table 2.

Table 2 ADF unit root test

Variable	ADF unit root test (trend & constant)			
	Level t$_{statistics}$	p	1.Difference t$_{statistics}$	p
LBTC	-1.985	0.601	-7.895	0.000*
LETH	-1.843	0.675	-7.986	0.000*
LDJCI	-1.902	0.645	-7.730	0.000*
LGLD	-2.622	0.271	-9.950	0.000*
LSPG100	-2.839	0.187	-10.27	0.000*
LUSDX	-2.012	0.586	-8.578	0.000*
LWTI	-2.942	0.154	-7.076	0.000*
Critical values	1% 5% 10%		-2.590 -1.944 -1.614	
* indicates a 1% statistical significance level.				

Source: Author's estimate

According to Table 2, when the first differences of all variables were taken at the 1% significance level, it was determined that t-statistics in absolute terms were greater than the MacKinnon critical values. This finding shows that all-time series are not stationary at I(0), but time series are stationary at I(1) at the 1% significance level. The VAR (Vector Autoregression) model was created by first determining the lag length to investigate the long-term relationship between time series determined to be stationary at the same level. Table 3 presents the lag length criteria and optimal lag lengths for the established models.

Table 3 Lag length criteria for Model 1 and Model 2

Lag	LogL	LR	FPE	AIC	SC	HQ
Model 1						
0	324.3309	NA	2.466e-11	-7.403045	-7.231811	-7.334131
1	851.4145	968.3628	2.70e-16*	-18.82359*	-17.62496*	-18.34120*
2	884.4094	56.01462*	2.93e-16	-18.75371	-16.52767	-17.85783
Model 2						
0	286.7775	NA	5.88e-11	-6.529709	-6.358476	-6.460795
1	817.0772	974.2716*	6.00e-16*	-18.02505*	-16.82642*	-17.54266*
2	844.4409	46.45458	7.42e-16	-17.82421	-15.59817	-16.92833

* indicates the optimal lag length.
Source: Author's estimate

It was found that various information criteria for the study's models reached their minimum values for lag 1, as shown in Table 3. To determine the optimum lag length of the models, hypothetical tests of the model were carried out based on the appropriate lag length. It was examined whether the inverse roots of the AR-characteristic polynomials were within the unit circle to determine whether the model established according to appropriate lag lengths was stable.

Figure 2 Inverse roots graph of AR-characteristic polynomials of the models

Model 1

Inverse Roots of AR Characteristic Polynomial

Model 2

Inverse Roots of AR Characteristic Polynomial
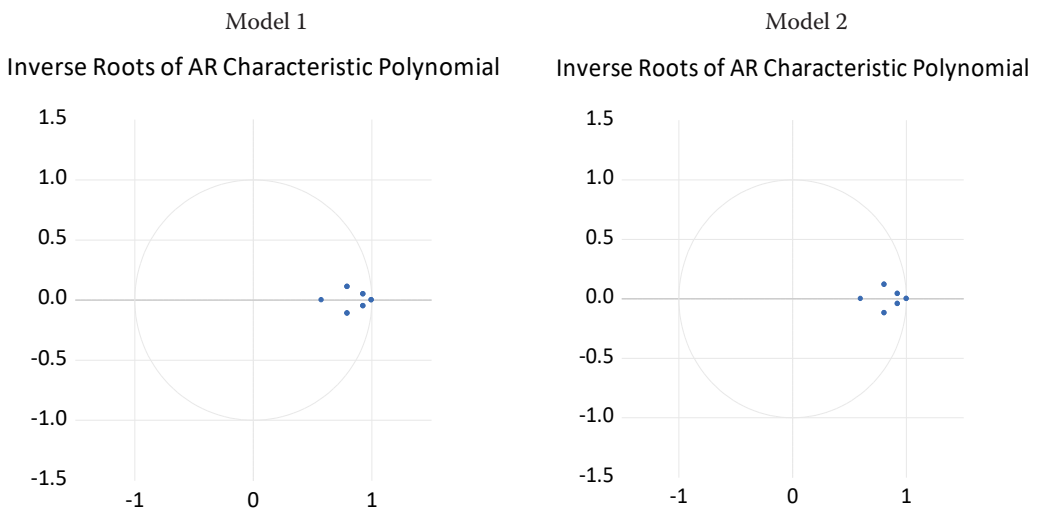
Source: Author's estimate

Figure 2 shows that the inverse roots of the AR-characteristic polynomials are located within the unit circle and the established models are stable.

After determining that the models were stable, LM and White tests were performed to test the auto-correlation and heteroscedasticity problems in the models presented in Table 4, which indicates that there were no autocorrelation and heteroscedasticity problems in the models since the p-probability values of the tests were greater than 0.05.

Table 4 Autocorrelation and heteroscedasticity test results

	Lag length	LM test p-probability	White test p-probability
Model 1	1	0.1966	0.2374
Model 2	1	0.1737	0.3518

Source: Author's estimate

VAR(1) models were established and the VAR model-based Johansen cointegration test was applied. Johansen cointegration test results are presented in Table 5.

Table 5 Johansen cointegration test

Model 1							
		Trace			Maximum eigenvalue		
Cointegration number	Eigenvalue	Trace statistic	0.05 Critical value	p	Max-eigen statistic	0.05 Kritik Değer	p
None	0.267559	82.06125	95.75366	0.2990	28.64622	40.07757	0.5161
At most 1	0.229132	53.41503	69.81889	0.4870	23.94190	33.87687	0.4598
At most 2	0.139737	29.47313	47.85613	0.7454	13.84760	27.58434	0.8330
At most 3	0.090908	15.62554	29.79707	0.7383	8.768437	21.13162	0.8507
At most 4	0.071801	6.857099	15.49471	0.5943	6.854807	14.26460	0.5065
At most 5	2.49E-05	0.002292	3.841465	0.9598	0.00229	3.841465	0.9598
Model 2							
		Trace			Maximum eigenvalue		
Cointegration number	Eigenvalue	Trace statistic	0.05 Critical value	p	Max-eigen statistic	0.05 Kritik Değer	p
None	0.273406	82.30789	95.75366	0.2917	29.38365	40.07757	0.4651
At most 1	0.232287	52.92424	69.81889	0.5085	24.31927	33.87687	0.4325
At most 2	0.153746	28.60497	47.85613	0.7868	15.35806	27.58434	0.7194
At most 3	0.086801	13.24690	29.79707	0.8800	8.353706	21.13162	0.8808
At most 4	0.051576	4.893196	15.49471	0.8201	4.871780	14.26460	0.7579
At most 5	0.000233	0.021417	3.841465	0.8836	0.021417	3.841465	0.8836

Source: Author's estimate

Based on the cointegration analysis results presented in Table 5, it was found that there were no cointegration equalities at a 5% significance level for the established models. Therefore, it can be said there are no long-term relationships between the variables used in the established models.

The VAR-based Granger causality test was applied to the model to examine the causality relationship and it was not found to have a cointegration relationship. Granger causality test results are presented in Table 6.

Table 6 Granger causality test results

H$_0$ hypothesis	F-statistics	p-probability
LDJCI →LBTC.	0.34035	0.5611
LBTC →LDJCI.	4.42326	0.0382**
LGLD →LBTC.	1.48726	0.2258
LBTC →LGLD.	1.59241	0.2102
LSPG100 →LBTC.	0.59903	0.4410
LBTC →LSPG100.	2.63308	0.1082
LUSDX →LBTC.	0.26403	0.6086
LBTC →LUSDX.	0.19364	0.6610
LWTI →LBTC.	1.53751	0.2182
LBTC →LWTI.	2.06117	0.1546
LDJCI →LETH.	0.01497	0.9029
LETH →LDJCI.	4.04539	0.0473**
LGLD →LETH.	2.05614	0.1551
LETH →LGLD.	0.18879	0.6650
LSPG100 →LETH.	1.45198	0.2314
LETH →LSPG100.	0.44902	0.5045
LUSDX →LETH.	0.01102	0.9166
LETH →LUSDX.	0.22864	0.6337
LWTI →LETH.	1.10472	0.2960
LETH →LWTI.	3.40931	0.0681

** indicates 5% statistical significance levels.
Source: Author's estimate

According to Table 6, unidirectional causality relationships have been identified from the LBTC and LETH variables to the LDJCI variable. The results obtained from the Granger causality test indicate that Bitcoin and Ethereum prices Granger-cause the Dow Jones Commodity Index. It can be stated that the change in the Bitcoin and Ethereum prices caused a change in Dow Jones Commodity Index prices. The results obtained from the Granger causality test indicate that there is no causality relationship between Bitcoin, Ethereum, and other variables.

The rolling window causality test (Balcilar et al., 2010) was applied to determine whether there is a causal relationship between the variables in the estimated models for different periods. The results were also compared with classical Granger causality analysis. Figures 3 and 4 show the rolling window causality test results for Model 1 and Model 2, respectively. Bootstrap p (probability)-values of LR statistics were calculated to test the null hypothesis that the dependent variable/independent variable does not Granger-cause the independent variable/dependent variable.

Figure 3 Rolling window causality results of Model 1

a) DJI-BTC

b) BTC-DJI

c) GLD-BTC d) BTC-GLD

e) SPG100-BTC

f) BTC-SPG100

g) USDX-BTC

h) BTC-USDX

i) WTI-BTC

j) BTC-WTI
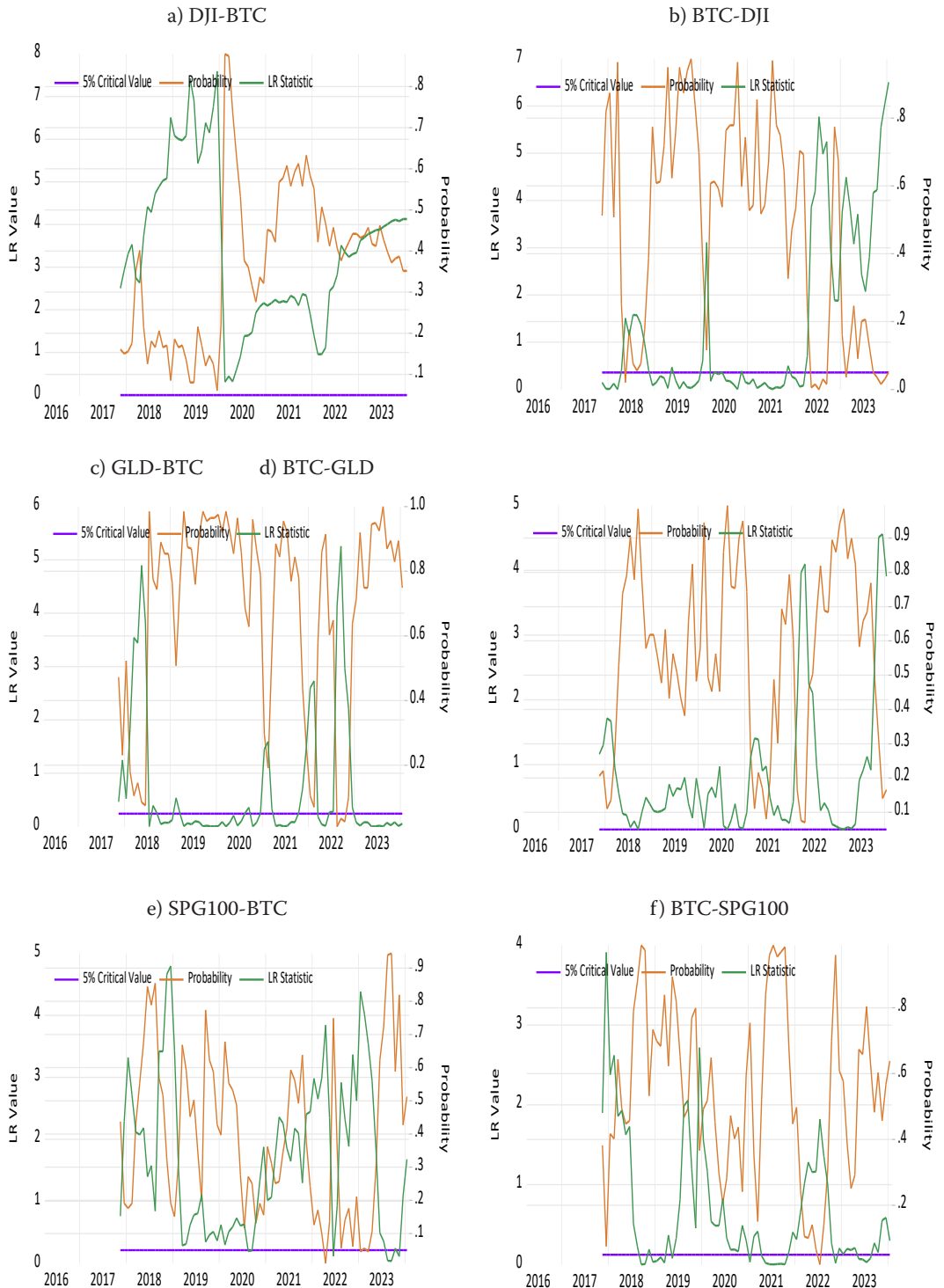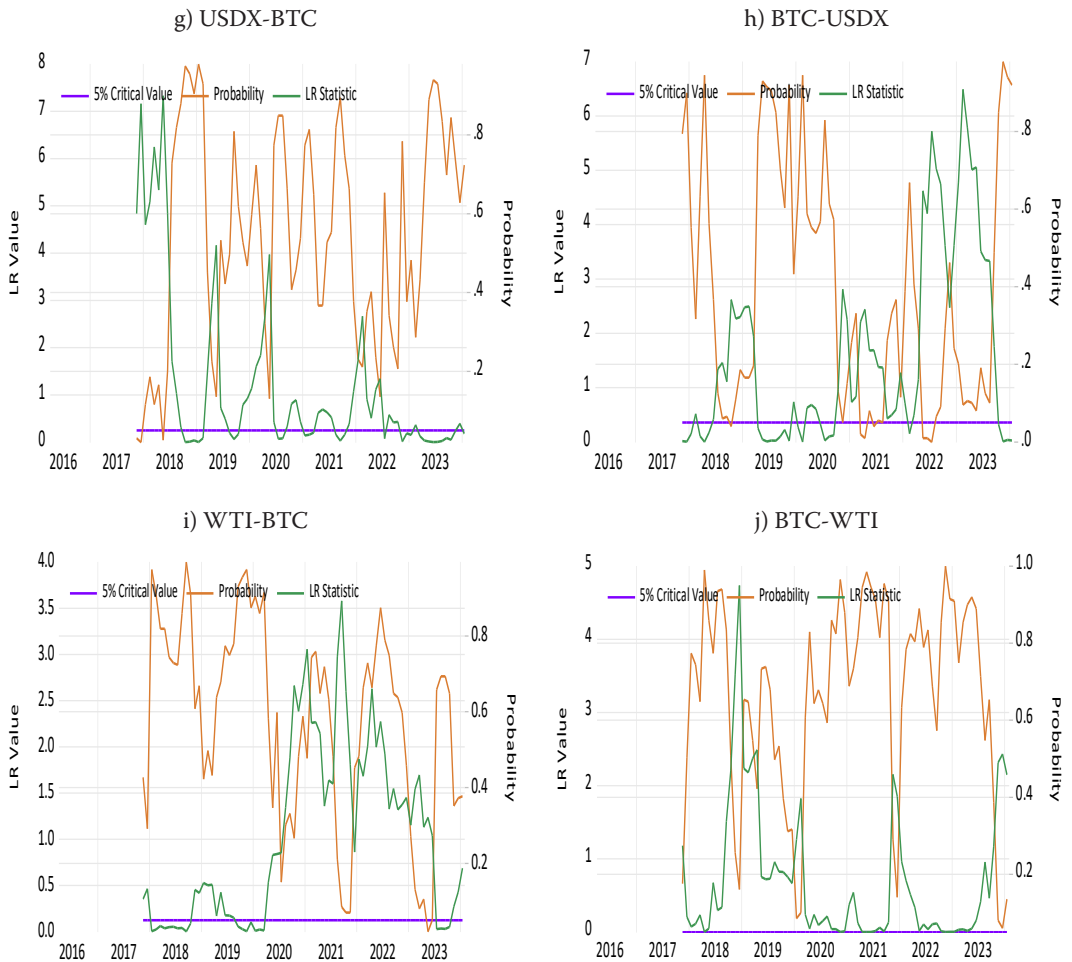
Source: Author's estimate

Figure 4 Rolling window causality results of Model 2

a)DJI-ETH

b) ETH-DJI
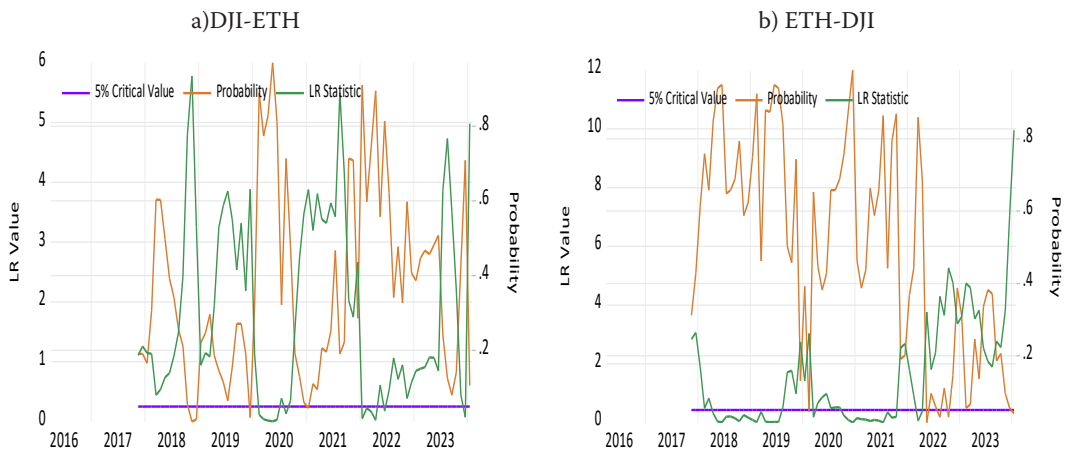
c) GLD-ETH

d) ETH-GLD

e) SPG100-ETH
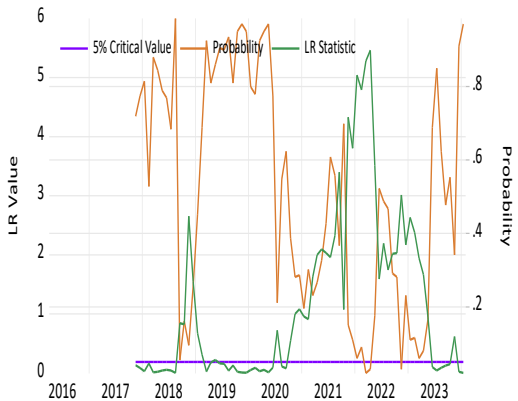
f) ETH-SPG100
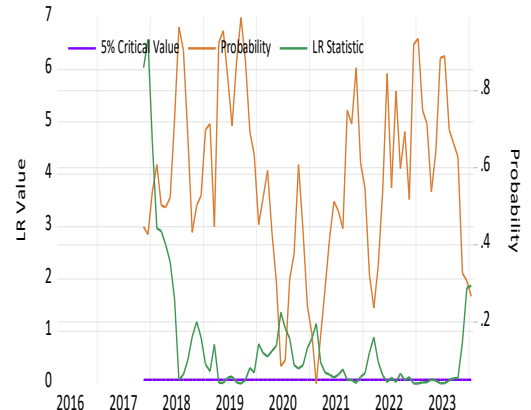
g) USDX-ETH

h) ETH-USDX

i) WTI-ETH

j) ETH-WTI
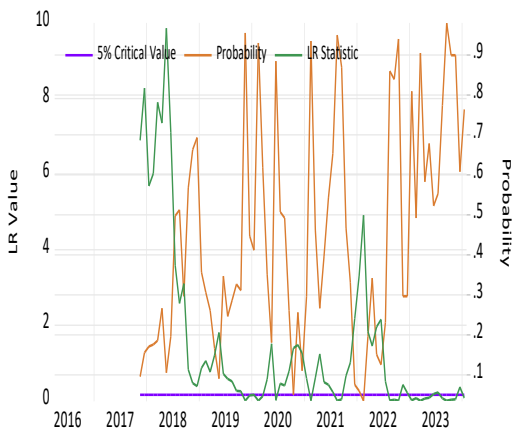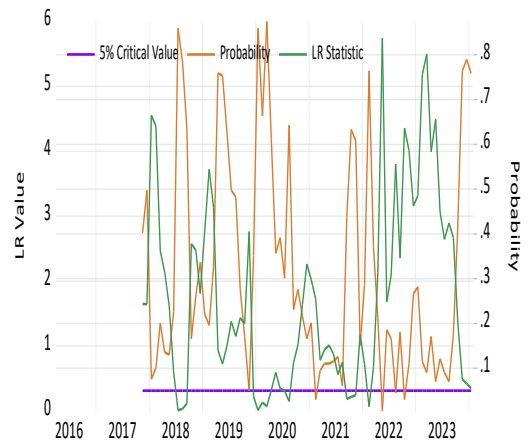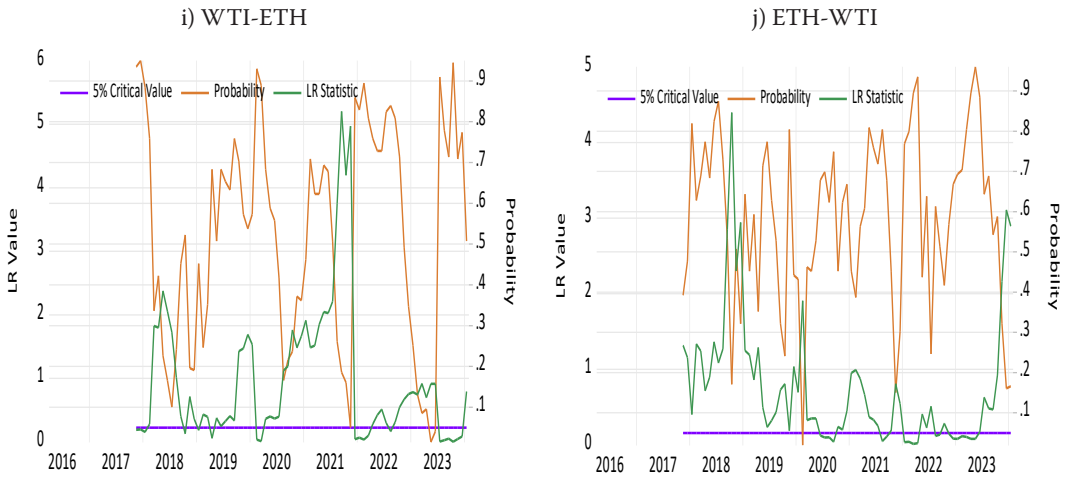
Source: Author's estimate

According to Figure 3, Rolling window causality test results of Model 1, there is a one-way causality between LBTC-LDJCI variables. It means that Bitcoin prices Granger-cause the Dow Jones Commodity Index prices. There are causalities in May 2018, May to September 2022, February and September to December 2023, and January 2024. This finding is in line with the classic Granger causality test. In contrast to the classic Granger causality test, there is a causality between LBTC-LUSDX in October 2018, March, April, June and August 2021, and May-July 2022. It means that Bitcoin prices Granger-cause the US Dolar Index prices. However, unlike the classic Granger causality test, it can be said that there are some weak causalities, too. There is a weak one-way causality between LGLD-LBTC in August-October 2022. There is a weak bidirectional causality between LSPG100-LBTC variables in April 2022, January and March 2023, and LBTC-LSPG100 in July 2022. There is a bidirectional causality between LBTC and LUSDX variables, and a weak causality between LUS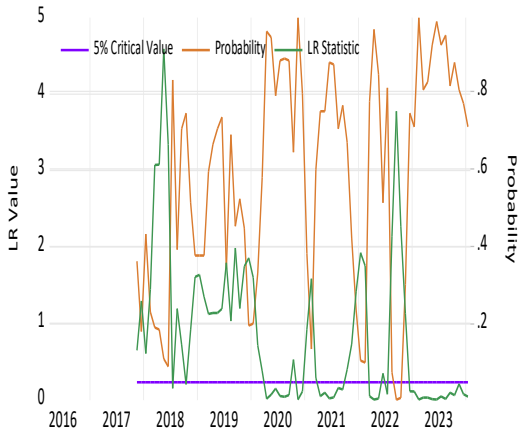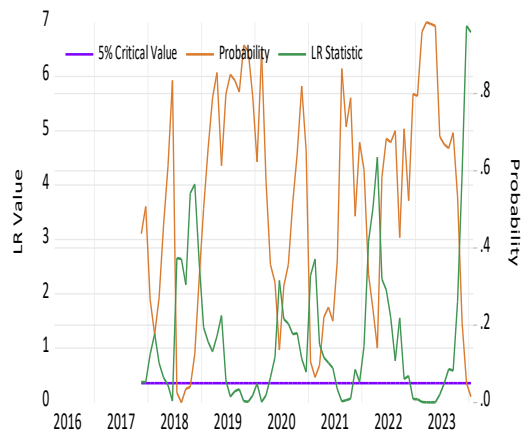DX-LBTC in November and December 2017, and May 2018. Furthermore, there is a weak causality between LWTI-LBTC only in May 2023.

According to Figure 4, Rolling window causality test results of Model 2, it can be stated that there are bidirectional causalities between all variables, although some are weak. There is a causality between LETH-LDJCI variables. It means that Ethereum

prices Granger-cause the Dow Jones Commodity Index prices. There are causalities in February 2020, May-October 2022, February, March, November and December 2023, and January 2024. This finding is parallel to the classic Granger causality test. However, unlike the Granger causality test, there is a weak causality between LDJCI-LETH variables at the end of 2018 and 2019. In contrast to the classic Granger causality test, there is a causality between LETH-LGLD in the July-October 2018 period, and in December 2023 and January 2024. It means that Ethereum prices Granger-caused the gold prices in these periods. Also, there is a causality between LETH-LUSDX in November 2019, February 2021, and May, August, and October 2022. Unlike the classic Granger causality test, the rolling window causality test suggests that there are some other weak causalities, too. For example, there is a weak LGLD-LETH causality in September and October 2022, a weak LSPG100-LETH causality in March, April, and November 2022, a weak LETH-LSPG100 causality in February 2021, a weak LWTI-LETH causality in May and June 2023, and a weak LETH-LWTI causality in February 2020.

The aim is to determine the extent to which Bitcoin and Ethereum prices are affected by their shocks and the shocks of the independent variables considered, with variance decomposition analysis. Figures 5 and 6 show the variance decomposition analysis results of Models 1 and 2, respectively.

Figure 5 Variance decomposition analysis results of Model 1

Period	S.E.	LBTC	LDJCI	LGLD	LSPG100	LUSDX	LWTI
1	0.206486	100.0000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.284327	99.42769	0.003313	0.072994	0.452112	0.011219	0.032676
3	0.340685	98.60005	0.004028	0.229966	1.076226	0.015626	0.074107
4	0.385914	97.71978	0.036723	0.449988	1.678233	0.012957	0.102318
5	0.424013	96.82484	0.139050	0.712505	2.200199	0.013047	0.110361
6	0.457061	95.89656	0.335423	1.000408	2.637590	0.027420	0.102594
7	0.486321	94.90566	0.634068	1.300614	3.003450	0.065588	0.090625
8	0.512635	93.82938	1.029537	1.603669	3.314158	0.133860	0.089400
9	0.536593	92.65640	1.506909	1.903069	3.584380	0.235369	0.113876
10	0.558620	91.38650	2.046001	2.194575	3.825744	0.370630	0.176550
Cholesky ordering:		LBTC	LDJCI	LGLD	LSPG100	LUSDX	LWTI

Source: Author's estimate

Figure 6 Variance decomposition analysis results of Model 2

Period	S.E.	LETH	LDJCI	LGLD	LSPG100	LUSDX	LWTI
1	0.315431	100.0000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.433453	99.61692	0.017539	0.051308	0.144783	0.001536	0.167914
3	0.517965	98.99865	0.021069	0.164671	0.368777	0.001119	0.445716
4	0.585009	98.29333	0.016517	0.328236	0.613065	0.003565	0.745288
5	0.640724	97.55949	0.025946	0.529114	0.855263	0.017188	1.012999
6	0.688219	96.81239	0.069802	0.755851	1.089985	0.050070	1.221901
7	0.729341	96.04911	0.160352	0.999267	1.319002	0.108416	1.363853
8	0.765311	95.26115	0.300908	1.252477	1.546573	0.196126	1.442764
9	0.797004	94.44030	0.487543	1.510568	1.777340	0.314975	1.469270
10	0.825082	93.58096	0.711548	1.770177	2.015452	0.465048	1.456813
Cholesky ordering:		LETH	LDJCI	LGLD	LSPG100	LUSDX	LWTI

Source: Author's estimate

Examination of Figure 5 reveals that Bitcoin prices are influenced by approximately 91% of changes in their lagged values, with 3.8% attributed to the SPG100 index, 2.2% to gold prices, 2.0% to the DJCI index, 0.4% to the USDX index, and 0.2% to oil prices. Initially, Bitcoin prices were entirely influenced by their shocks, but this rate decreased to approximately 91% with a 10-period delay. In contrast, Figure 6 shows that Ethereum prices are affected by about 94% of changes in their lagged values, with 2.0% attributed to the SPG100 index, 1.8%

to gold prices, 1.5% to oil prices, 0.7% to the DJCI index, and 0.5% to the USDX index. Like Bitcoin, Ethereum prices were initially fully affected by their shocks, but this rate decreased to around 89% after a 10-period delay.

5. Conclusion

This paper aims to reveal the cointegration and causality relationships between major cryptocurrencies, Bitcoin and Ethereum, and global finan-

cial asset classes, the diversifying feature of cryptocurrencies for other assets, as well as to make recommendations to investors, financial advisors, policymakers, and especially Bitcoin and Ethereum investors, in line with the findings. The findings relate to asset allocation, hedging, risk management, financial stability, investment decisions, and portfolio diversification. This paper aims to fill a gap in the literature by conducting a comparative analysis of the major cryptocurrencies Bitcoin and Ethereum, and examining whether the diversification effect remains consistent throughout the review period. Although cryptocurrencies are considered highly speculative, they have been included in the investment asset class, especially in recent years. In this way, the relationship between two important cryptocurrencies and global financial assets was comparatively analyzed, and depending on whether there was a relationship or not, the study aimed to determine whether they would provide diversification benefits in the portfolios created with these assets. In this paper, Johansen cointegration, Granger causality, rolling window causality, and variance decomposition analyses were applied to reveal the relationships between selected global financial asset classes and Bitcoin and Ethereum, as well as the diversification features of these cryptocurrencies.

The results of the research study indicate that there are no cointegration relationships between Bitcoin and Ethereum and selected global financial asset classes. One of the main reasons for the lack of relationship may be that Bitcoin and Ethereum exhibit high variance, while the other variables have low variance. Another reason may be because financial asset classes other than Bitcoin and Ethereum depend on a certain central authority and are traded under certain rules. Because of that, cyrptocurrencies represent virtual assets created without any authority, as stated. When we look at volatility movements in the markets in recent years, it appears that the low transaction volume of cryptocurrencies, compared to other traditional investment instruments, contributes to the lack of relationship between Bitcoin, Ethereum and these instruments. This finding is compatible with similar studies in the literature (Bouri et al. (2017), Baur et al. (2018b), Güleç et al. (2018), Giudici & Abu-Hashish (2019) and Özbek (2023)). However, this finding is opposite to a similar study in the literature by Van Wijk (2013). Granger causality analysis results indicated that there are no significant

causality relationships between the variables, with the exception of significant unidirectional causalities from Bitcoin and Ethereum prices to the Dow Jones Commodity Index. The causalities between asset classes, except Bitcoin, Ethereum, and DJCI. This finding aligns with the findings in the literature ((Bouri et al. (2017a), and Kang et al. (2020)). As of December 2017, cryptocurrencies, especially Bitcoin, started to be traded in the futures market of the US Chicago Commodity Exchange, revealing that cryptocurrencies can be considered as commodities. Because of that, there may be a causality relationship between Bitcoin, Ethereum, and the commodity index. Hence, it can be stated that Bitcoin and Ethereum are not good diversifiers for a portfolio containing the Dow Jones Commodity Index, but they are good diversifiers for a portfolio containing gold, oil, the S&P100 stock, and the US dollar investments. The unidirectional causality obtained from Bitcoin and Ethereum prices to Dow Jones Commodity Index prices shows that the major cryptocurrency prices have a significant impact on Dow Jones Commodity Index prices. That is why Bitcoin and Ethereum cannot be used as diversification instruments within a portfolio containing the Dow Jones Commodity Index. Therefore, because the Bitcoin and Ethereum markets and the global financial system have no long-term relationship and causality relationships, with the exception of the Dow Jones Commodity Index, it can be stated that these major cryptocurrencies act separately from selected global financial asset classes, other than the Dow Jones Commodity Index, throughout most of the review period. It includes a suggestion for investors to include Bitcoin or Ethereum in a portfolio created from selected global financial assets so that they can have a better diversified, and hence better risk-balanced, portfolio. These findings are parallel to Briere et al. (2015), Georgoula et al. (2015), Bouoijour and Selmi (2017), Baumöhl (2019), Kajtazi and Mono (2019), Kang et al. (2020), Charfeddine et al. (2020), Dutta et al. (2020), Qarni and Gulzar (2021), Bhuiyan et al. (2023), Bouri et al. (2023), and Hanif et al. (2023). However, a causality relationship found between most of the selected financial assets and Bitcoin and Ethereum, especially after 2021-2022, obtained as a result of rolling window causality analysis, suggests that there is a loss in the diversification feature of Bitcoin and Ethereum. Given that this period coincides with negative developments in the economy after the COVID-19 period, which is a global epidemic

disease, it can be stated that the negative atmosphere in the cryptocurrency markets has spread to other global investment instruments. However, the significant increases observed in the transaction volume of Bitcoin and Ethereum in recent years, similar to other assets, can be seen as an element contributing to the causality observed between Bitcoin, Ethereum, and selected global financial assets. Similarly, in their studies, Huang et al. (2021), Maghyereh and Abdoh (2022), Hanif et al. (2023), and Özbek (2023) concluded that there was a loss in Bitcoin's diversification ability after COVID-19. Variance decomposition analysis indicates that Bitcoin prices are affected by approximately 91% of changes in their lagged values, while 9% is attributed to the selected global financial asset prices. Ethereum prices are affected by approximately 94% of the changes in their lagged values, and 6% are attributed to the selected global financial asset prices. These findings indicate that compared to Ethereum, Bitcoin prices are more affected by these global financial asset classes; however, it can be stated that the level of influence is low for both Bitcoin and Ethereum. Bitcoin and Ethereum can provide diversification benefits to portfolios as a result of their inclusion in investment portfolios, with the finding that Bitcoin and Ethereum act separately from selected global financial assets. It can be said that Bitcoin and Ethereum act similarly in the global investment portfolio. However, it can be concluded that investors should be more careful when including these assets in their portfolios, especially during periods of sharp declines in Bitcoin and Ethereum markets, such as after crisis periods that can affect the whole world, and the effect may spread to other assets during these periods.

In addition, policymakers should take into account Bitcoin, Ethereum, which has millions of users, and the cryptocurrency market in this context, in terms of economic reforms and cryptocurrency regulations. Our findings are important to investors, financial advisors, Bitcoin and Ethereum investors, and policymakers for making decisions regarding asset allocation, hedging, risk management, financial stability, investment decisions, and portfolio diversification. The findings serve as a reminder to those who manage portfolios that they plan to include Bitcoin and Ethereum in their investment portfolios as a diversification instrument to keep abreast of global geopolitical conditions and other global economic events when trading investment instruments. This paper is serves as a valuable reference for investors and portfolio managers in determining investment horizons, managing risks, and aiming for profitable opportunities.

The interaction between the two major cryptocurrencies, Bitcoin and Ethereum, and global gold, oil prices, the US Dollar Index, the Dow Jones Commodity Index, and the S&P Global 100 Index was analyzed in this paper. The interaction between selected financial investment instruments and other cryptocurrencies can be examined in further research to determine whether the findings also apply to other cryptocurrencies. Within the scope of the study, it was concluded that Bitcoin and Ethereum are generally good diversifiers for selected financial investment instruments. It is recommended to examine the interaction between Bitcoin, Ethereum, and other cryptocurrencies and different investment instruments in further research to determine whether Bitcoin and Ethereum are good diversifiers for other investment instruments within the scope of the paper.

References

1. Ahad, M. (2017). Financial development and money demand function: Cointegration, causality and variance decomposition analysis for Pakistan. *Global Business Review*, *18*(4), 811-824. https://doi.org/10.1177/0972150917692209

2. Akhtaruzzaman, M., Sensoy, A. & Corbet, S. (2020). The influence of Bitcoin on portfolio diversification and design. *Finance Research Letters*, *37*, 101344. https://doi.org/10.1016/j.frl.2019.101344

3. Bakry, W., Rashid, A., Al-Mohamad, S. & El-Kanj, N. (2021). Bitcoin and portfolio diversification: A portfolio optimization approach. *Journal of Risk and Financial Management*, *14*(7), 282. https://doi.org/10.3390/jrfm14070282

4. Balcilar, M., Ozdemir, Z. A. & Arslanturk, Y. (2010). Economic growth and energy consumption causal nexus viewed through a bootstrap rolling window. *Energy Economics*, *32*(6), 1398-1410. https://doi.org/10.1016/j.eneco.2010.05.015

5. Balke, N. S. & Fomby, T. B. (1997). Threshold Cointegration. *International Economic Review*, *38*(3), 627-645. https://doi.org/10.2307/2527284

6. Baumöhl, E. (2019). Are cryptocurrencies connected to forex? A quantile cross-spectral approach. *Finance Research Letters*, *29*, 363-372. https://doi.org/10.1016/j.frl.2018.09.002

7. Baur, D. G. & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review*, *45*(2), 217-229. https://doi.org/10.1111/j.1540-6288.2010.00244.x

8. Baur, D. G., Hong, K. & Lee, A. D. (2018a). Bitcoin: Medium of exchange or speculative assets?. *Journal of International Financial Markets, Institutions and Money*, *54*, 177-189. https://doi.org/10.1016/j.intfin.2017.12.004

9. Baur, D. G., Dimpfl, T. & Kuck, K. (2018b). Bitcoin, gold and the US dollar – A replication and extension. *Finance Research Letters*, *25*, 103-110. https://doi.org/10.1016/j.frl.2017.10.012

10. Bhuiyan, R. A., Husain, A. & Zhang, C. (2023). Diversification evidence of bitcoin and gold from wavelet analysis. *Financial Innovation*, *9*, 100. https://doi.org/10.1186/s40854-023-00495-1

11. Bouri, E., Jalkh, N., Molnár, P. & Roubaud, D. (2017a). Bitcoin for energy commodities before and after the December 2013 crash: diversifier, hedge or safe haven?. *Applied Economics*, *49*(50), 5063-5073. https://doi.org/10.1080/00036846.2017.1299102

12. Bouri, E., Gupta, R., Tiwari, A. K. & Roubaud, D. (2017b). Does Bitcoin hedge global uncertainty? Evidence from wavelet-based quantile-in-quantile regressions. *Finance Research Letters*, *23*, 87-95. https://doi.org/10.1016/j.frl.2017.02.009

13. Bouri, E., Shahzad, S. J. H. & Roubaud, D. (2020). Cryptocurrencies as hedges and safe-havens for US equity sectors. *The Quarterly Review of Economics and Finance*, *75*, 294-307. https://doi.org/10.1016/j.qref.2019.05.001

14. Bouri, E., Salisu, A. A. & Gupta, R. (2023). The predictive power of Bitcoin prices for the realized volatility of US stock sector returns. *Financial Innovation*, *9*, 62. https://doi.org/10.1186/s40854-023-00464-8

15. Bouoiyour, J. & Selmi, R. (2017). Ether: Bitcoin's competitor or ally?. hal-01567277f. https://doi.org/10.48550/arXiv.1707.07977

16. Briere, M., Oosterlinck, K. & Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with bitcoin. *Journal of Asset Management*, *16*, 365-373. http://dx.doi.org/10.1057/jam.2015.5

17. Charfeddine, L., Benlagha, N. & Maouchi, Y. (2020). Investigating the dynamic relationship between cryptocurrencies and conventional assets: Implications for financial investors. *Economic Modelling*, *85*, 198-217. https://doi.org/10.1016/j.econmod.2019.05.016

18. CoinMarketCap (2023). *Top 100 Cryptocurrencies by Market Capitalization*. https://coinmarketcap.com

19. Das, D., Le Roux, C. L., Jana, R. K. & Dutta, A. (2020). Does Bitcoin hedge crude oil implied volatility and structural shocks? A comparison with gold, commodity and the US Dollar. *Finance Research Letters*, *36*, 101335. https://doi.org/10.1016/j.frl.2019.101335

20. Dutta, A., Das, D., Jana, R. K. & Vo, X. V. (2020). COVID-19 and oil market crash: Revisiting the safe haven property of gold and Bitcoin. *Resources Policy*, *69*, 101816. https://doi.org/10.1016/j.resourpol.2020.101816

21. Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar – A GARCH volatility analysis. *Finance Research Letters*, *16*, 85-92. https://doi.org/10.1016/j.frl.2015.10.008

22. Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D. & Giaglis, G. M. (2015). Using time-series and sentiment analysis to detect the determinants of bitcoin prices. In Kokolakis, S. et al. (Eds.). *Proceedings of the 9th Mediterranean Conference on Information Systems (MCIS 2015)* (pp. 371-382). Samos: University of the Aegean. http://dx.doi.org/10.2139/ssrn.2607167

23. Giudici, P. & Abu-Hashish, I. (2019). What determines Bitcoin exchange prices? A network VAR approach. *Finance Research Letters*, *28*, 309-318. https://doi.org/10.1016/j.frl.2018.05.013

24. Gujarati, D. N. (1999). *Temel Ekonometri*. LiteratürYayıncılık.

25. Güleç, Ö. F., Çevik, E. & Bahadır, N. (2018). Bitcoin ile finansal göstergeler arasindaki ilişkinin incelenmesi. *Kırklareli Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *7*(2), 18-37.

26. Hanif, W., Ko, H. U., Pham, L. & Kang, S. H. (2023). Dynamic connectedness and network in the high moments of cryptocurrency, stock, and commodity markets. *Financial Innovation*, *9*(1), 84. https://doi.org/10.1186/s40854-023-00474-6

27. Henriques, I. & Sadorsky, P. (2018). Can bitcoin replace gold in an investment portfolio?. *Journal of Risk and Financial Management*, *11*(3), 48. https://doi.org/10.3390/jrfm11030048

28. Huang, Y., Duan, K. & Mishra, T. (2021). Is Bitcoin really more than a diversifier? A pre-and post-COVID-19 analysis. *Finance Research Letters*, *43*, 102016. https://doi.org/10.1016/j.frl.2021.102016

29. Işığıçok, E (1994). *Zaman Serilerinde Nedensellik Çözümlemesi*. Uludağ Üniversitesi Basımevi.

30. Johansen, S. & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, *52*(2), 169-210. https://doi.org/10.1111/j.1468-0084.1990.mp52002003.x

31. Kajtazi, A. & Moro, A. (2019). The role of bitcoin in well diversified portfolios: A comparative global study. *International Review of Financial Analysis*, *61*, 143-157. https://doi.org/10.1016/j.irfa.2018.10.003

32. Kang, S. H., Yoon, S. M., Bekiros, S. & Uddin, G. S. (2020). Bitcoin as hedge or safe haven: evidence from stock, currency, bond and derivatives markets. *Computational Economics*, *56*, 529-545. https://doi.org/10.1007/s10614-019-09935-6

33. Kliber, A., Marszałek, P., Musiałkowska, I. & Świerczyńska, K. (2019). Bitcoin: Safe haven, hedge or diversifier? Perception of bitcoin in the context of a country's economic situation – A stochastic volatility approach. *Physica A: Statistical Mechanics and Its Applications*, *524*, 246-257. https://doi.org/10.1016/j.physa.2019.04.145

34. Maghyereh, A. & Abdoh, H. (2022). COVID-19 and the volatility interlinkage between bitcoin and financial assets. *Empirical Economics*, *63*, 2875-2901. https://doi.org/10.1007/s00181-022-02223-7

35. Mushtaq, R. (2011). *Augmented Dickey Fuller Test.* Available at SSRN: https://ssrn.com/abstract=1911068 or http://dx.doi.org/10.2139/ssrn.1911068

36. Özbek, G. B. (2023). Bitcoin'in çeşitlendirme ve riski dengeleme kabiliyeti. *Ekonomi Politika ve Finans Araştırmaları Dergisi*, *8*(3), 519-538. https://doi.org/10.30784/epfad.1333841

37. Öztürk, M. B., Arslan, H., Kayhan, T. & Uysal, M. (2018). Yeni bir hedge enstrumanı olarak Bitcoin: Bitconomi. *Ömer Halisdemir Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *11*(2), 217-232. https://doi.org/10.25287/ohuiibf.415713

38. Qarni, M. O. & Gulzar, S. (2021). Portfolio diversification benefits of alternative currency investment in Bitcoin and foreign exchange markets. *Financial Innovation*, *7*, 17. https://doi.org/10.1186/s40854-021-00233-5

39. Tarı, R. (2015). *Ekonometri* (11th ed)*.* Umuttepe Yayınları.

40. Van Wijk, D. (2013). *What can be expected from the BitCoin*. Erasmus Universiteit Rotterdam. https://thesis.eur.nl/pub/14100/Final-version-Thesis-Dennis-van-Wijk.pdf

Nesrin Akbulut
*Alanya Alaaddin
Keykubat University
Department of Economics
07425 Alanya, Turkiye
nesrin.akbulut@alanya.edu.tr*

Yakup Ari
*Alanya Alaaddin
Keykubat University
Department of Economics
07425 Alanya, Turkiye
yakup.ari@alanya.edu.tr*

Berkay Aktürk
*Yıldız Technical University
34349 İstanbul, Türkiye
berkayakturk@gmail.com*

TVP-VAR FREQUENCY CONNECTEDNESS ANALYSIS ON CPI-BASED MONTHLY REAL RETURN VOLATILITY OF FINANCIAL INVESTMENT INSTRUMENTS

Abstract

Purpose: This study explores volatility transmission among the real returns of financial investment instruments, using the Diebold-Yilmaz approach and data from the Turkish Statistical Institute. The dataset includes monthly real return rates of instruments like Gross Interest Rate (GIR) for deposits, ingot Gold (GOLD), Istanbul Stock Exchange 100 Index (BIST-100), United States Dollar (USD), Euro (EUR), and Government Domestic Debt Instruments (GDDI) from January 2005 to April 2023.

Methodology: Real return rates were adjusted using the Consumer Price Index (CPI). Absolute values of real returns served as volatility proxies. To evaluate volatility spillover among these instruments, the Time-Varying Parameter Vector Autoregressive (TVP-VAR) frequency connectedness approach was utilized.

Results: The average of the Total Connectedness Index (TCI) suggests 40.37% of error variance in investment instruments is due to network connectedness, with short-term and long-term values at 33.95% and 6.41%, respectively. Dynamic TCI values spiked during events like the 2008 crisis, 2018 and 2021 exchange rate shocks, and COVID-19. USD and EUR consistently caused net volatility spillovers, GOLD in the long run, GDDI in the short run and aggregate. GIR was most impacted by network shocks. The study also examined the Net Pairwise Connectedness Index (NPCI) to identify dominant instruments in the network.

Conclusion: The findings show the interdependencies and significant roles of particular investment instruments in the transmission of volatility, offering insights for portfolio diversification and risk management.

Keywords: Real return, CPI, TVP-VAR frequency connectedness, volatility

1. Introduction

The concept of volatility is defined as the change in the price of a product in financial markets within a certain period of time. This concept can also be used in terms of volatility, mobility, fluctuation and similar meanings. Especially with the increasing integration among financial markets, the volatility in one market can also affect the other. This situation

is referred to as "volatility spillover". Since these volatilities on financial investment instruments lead to uncertainty, they emerge as a risk factor for investors. As a matter of fact, when making investment decisions, investors look at whether a shock in one market affects the other market. Therefore, estimating volatility spreads among financial investment instruments is of great importance for risk management and effective portfolio diversification for both investors and financial institutions (Poon & Granger, 2003; Verma & Jackson, 2012; Topaloğlu, 2020; Cao & Wen, 2019).

With the developing technology, today's investors are able to invest in global stock exchanges as well as in their own countries. Due to the increasing risks with these investments, investors prefer to invest in a wide range in order to minimize risks while forming their investment portfolios, and therefore they apply to different investment instruments. At this point, traditional investment instruments such as foreign currency and gold have been among the most preferred investment instruments by investors. Due to the high volatility of foreign exchange markets, risk-seeking investors who aim to earn higher returns often invest their money in these markets. Gold, on the other hand, is an investment instrument that is seen as a safe harbor by investors, especially during periods of increased uncertainty (Baur & Lucey, 2010). Moreover, investment instruments such as the Gross Interest Rate (GIR) for deposits, the Istanbul Stock Exchange 100 (BIST-100) Index, and Government Domestic Debt Instruments (GDDI) are among the alternative investment areas where savings are evaluated in Turkey. Such investment instruments are affected by many factors such as economic developments, market conditions, political factors and even developments in international markets, and may show high volatility over time. Therefore, it is important to understand the interaction of investment instruments with each other and the volatility spread mechanism among them for an effective portfolio diversification.

In the literature, there are many studies on the relationship between various financial investment instruments. However, in addition to the financial assets used in these studies, methods used and the time period (daily, weekly, monthly, etc.) vary. While some of the empirical studies focused on the effect between investment instruments, others investigated the causal relationship between these

instruments (Wang & Chueh, 2013; Bhunia, 2013; Başarır, 2019; Cingöz & Kendirli, 2019; Güney & Ilgın, 2019; Jain & Biswal, 2019). On the other hand, as financial markets become increasingly integrated, strong correlations emerge among financial assets, leading to the transmission of volatility from one market to another. These volatility spillovers also lead to the transmission of market risks. In line with these developments, the number of studies on volatility spillovers among financial investment instruments is increasing day by day.

For example, Sumner et al. (2010), used weekly data for the period from 1 October 1970 to 25 April 2009 to reveal the interrelationship between gold, stock and bond returns and volatilities with the spillover index approach of Diebold and Yilmaz (2012). According to the results, there was no evidence of return spillovers for the whole sample, but evidence of volatility spillovers was recorded. Return spreads were high in the early 1980s, mid-1990s, and the 2008 crisis, while volatility spreads were high in the late 1970s, early 1990s, and the 2008 crisis.

Badshah et al. (2013) revealed the spillover between stock, gold, and exchange rate volatility indices and the causality relationship between them, using daily data for the period from 3 June 2008 to 30 December 2011, defined by Rigobon's (2003) heteroscedasticity approach. The findings of the study showed that increased stock market volatility was associated with increased gold and exchange rate volatility. They also found a significant unidirectional spillover from the stock market to gold and exchange rates.

Shahrazi et al. (2014) used daily data from 2007 to 2013 to investigate the volatility and contagion between Iranian gold and foreign exchange markets using the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model. The findings showed that when a shock is transmitted from the gold market to the foreign exchange market, there is a two-way volatility spillover between the gold and foreign exchange markets. Similarly, Hein (2015) examined the connection and volatilities between the S&P 500 stock index, gold, crude oil and exchange rate (CHF/USD) returns between January 1999 and December 2013 using the GARCH model. The analysis showed that there is a significant contagion between the return volatilities of the variables. Moreover, a positive relationship was also found between gold and oil, while a negative relationship was found between gold and the exchange

rate. The study also revealed that gold was a safe investment instrument during the 2008 crisis.

By using weekly data from 6 January 1987 to 22 July 2014, Antonakakis and Kizys (2015) investigated the dynamic connectedness between the returns and volatility of commodities (gold, silver, platinum, palladium) price, crude oil price and exchange rate (EUR/USD, JPY/USD, GBP/USD, CHF) /USD) markets with the generalized VAR method of Diebold and Yilmaz (2012). They found that gold, silver, platinum, CHF/USD and GBP/USD exchange rates were net transmitters of return and volatility spillovers during the sampled period, while palladium, crude oil, EUR/USD and JPY/USD exchange rates were net receivers. It was also revealed that gold is the largest volatility transmitter, exhibiting a strong bidirectional correlation between gold and silver, platinum and palladium in terms of bilateral return spreads.

Roy and Roy (2017) investigated the extent of financial contagion in Indian asset markets using daily data from 3 April 2006 to 31 March 2016. They used the commodity future price index, bond price, exchange rate, gold price and stocks as variables and daily returns of assets was estimated by the DCC-MGARCH method, while volatility spillover estimation was performed by a generalized VAR approach. The stock market was found to have the highest financial contagion, while the gold market was observed to have the lowest financial contagion. Furthermore, it was found that whilst gold, bonds, and foreign exchange were net volatility receivers, commodities and equities were net volatility transmitters.

Şenol (2021) used the GARCH model to study the volatility correlations and spillovers among the BIST-100 index, currency rate, interest rate, and CDS premiums from 2 January 2010 to 10 April 2020. The findings show that there is a unidirectional volatility spillover from CDS premiums to the exchange rate and that there is a mutual volatility spillover between the exchange rate and the BIST 100 index and interest rate, and between CDS premiums and the interest rate. Additionally, it has been determined that there is a positive volatility relationship between CDS premiums and exchange rates and interest rates, and a negative volatility relationship between the BIST 100 index and these two factors.

Cihangir et al. (2020) investigated the interaction between returns by examining the dynamic effect between four investment instruments (gold, foreign exchange, stock market and interest return) using the Vector Autoregressive (VAR) method from January 2002 to November 2019. They found that due to a shock in one financial instrument, other financial instruments reacted in the same way in the first period and vice versa in subsequent periods. In the same way, Şeker (2021) analyzed the interaction between interest rates, USD, EUR, gold, the BIST 100 index and government domestic debt instrument returns for the period 2005 to 2021 using the VAR approach. The results of the impulse-response analysis revealed that all return variables responded positively to the shocks that occurred in them in the first period. In addition, it was also determined that there were both complementarity and substitution relationships between the return variables, and at the same time, government domestic debt instruments, interest and USD returns were more in interaction with each other.

Using daily data from February 2017 to February 2021, Önem (2021) revealed volatility interaction among gold and silver price returns and BIST Mining Index returns with the diagonal VECH GARCH approach. According to the results, intense volatility clusters were determined in gold and silver price returns and BIST Mining Index returns, and these volatilities were found to have permanent effects. Wen et al. (2021) examined the dynamic volatility spillovers between the Chinese stock market and commodity markets from May 2009 to June 2020, using the TVP-VAR approach. They revealed that there was a very high correlation between the stock market and commodity prices and that the stock market acts as a net shock receiver. On the other hand, it was found that the volatility between the stock market and the commodity market increases during crisis periods. Correspondingly, Ahmed and Huo (2021), investigated a dynamic relationship between the Chinese stock market, commodity markets and global oil prices with daily data for the period between July 2012 and June 2017, by using a three-variable VAR-BEKK-GARCH model. They found that there were a unidirectional return and volatility spillover effect from the Chinese stock market and the global oil market to the commodity market, as well as a unidirectional return spillover effect from the oil market to the Chinese stock market.

Yilmaz and Kılıç (2022) investigated the return and volatility interaction among interest rate, gold, USD and EUR investment instruments with weekly data from January 2010 to July 2021, using the VAR Exponential Generalized Autoregressive Conditional Heteroscedastic (EGARCH) method. They revealed that there was a bidirectional return interaction between interest rate-USD and USD-EUR, and a one-way return interaction between interest rate to EUR and gold to USD. They also determined that volatility interaction is bidirectional between USD-EUR, gold-USD, EUR-gold, and unidirectional from USD to gold.

Şak and Öcal Özkaya (2022) investigated the volatility spillovers between USD, EUR, gold and the BIST 100 index with daily data from January 2000 to August 2022, using the Diebold and Yilmaz (2012) approach. They found that EUR and USD were volatility transmitters, while gold and the BIST 100 index were volatility receivers. According to the bilateral connectedness, gold and the BIST 100 index were dominated by the USD, and in recent years, there has been a spillover from gold to the BIST 100 index.

In the post-COVID-19 period, a trend of rising inflation has emerged in the world economy. Developing markets, including Turkey, have been significantly affected by this situation. In Turkey, alongside dramatic increases in inflation, currency-related risks have escalated, leading to uncertainties in investors' portfolio formation efforts. The country's unique political risks, along with fluctuations in the currency, have increased these uncertainties. This situation in Turkey did not start only in the post-COVID-19 period; its roots go back further. For instance, the currency fluctuation experienced in August 2018 negatively impacted all markets and caused structural ruptures. Naturally, in such situations, the risk transmission and spread in portfolios formed among traditional investment instruments have become significant issues. When we examined the literature, we noticed a gap in the comprehensive examination of risk transmission among traditional investment instruments based in Turkey and, more importantly, in studies on inflation-based real returns. Additionally, the lack of application of the Diebold-Yılmaz connectedness approach emerged as a notable shortcoming.

Therefore, we examine the volatility pass-through between inflation-adjusted real returns announced monthly by the Turkish Statistical Institute (TSI). Our research applies the TVP-VAR frequency method based on the Diebold-Yılmaz connectedness approach. In this way, short- and long-term volatility transmissions between traditional investment instruments in Turkish financial markets are revealed. For this purpose, real return rates on gross deposit interest, gold bullion, the Borsa Istanbul 100 Index, the US dollar, EURO and Government Domestic Debt Securities, which are the most popular instruments by investors in Turkish financial markets, are defined as variables in connectivity analysis. Volatilities of real return rates were obtained from the absolute values of returns based on Poon's (2005) study. In other words, the absolute values of CPI-based monthly real returns were used as proxy values for volatility values. Summary statistics show that the volatilities of the financial instruments mentioned are flat and skewed to the right, and also stationary and have an ARCH effect. In addition, the Elliott, Rothenberg, and Stock (ERS) unit root test result shows that the volatility series are stationary at level. Thus, we demonstrate that the volatility connectedness in the network formed by these financial investment instruments can be analyzed with a Time-Varying Parameter Vector Autoregressive (TVP-VAR) based approach. In our study, we use the novel TVP-VAR frequency connectedness approach proposed by Chatziantoniou et al. (2021), which effectively benefits from the essence of the works of Baruník and Krehlík (2018) and Antonakakis et al. (2020). The paper is organized as follows: an overview of the TVP-VAR Connectedness Approach in the time and frequency domains comes first, then a discussion of the dataset. The findings are explained in the parts that follow. The study concludes with a summary of the findings and their consequences.

2. Data set

In this study, the volatility pass-through among the real returns of the investment instruments most preferred by the investors in the Turkish financial markets will be examined by means of the Diebold-Yilmaz (2012; 2014) approach. For this purpose, data from the Turkish Statistical Institute (TSI) database on monthly real return rates of investment instruments like the Gross Interest Rate (GIR) for deposits, ingot Gold (GOLD), the Istanbul Stock Exchange 100 Index (BIST-100), United States Dollar (USD), Euro (EUR) and Government Domestic Debt Instruments (GDDI) between January 2005 and April 2023 were calculated by reducing the

Consumer Price Index (CPI). Figure 1 illustrates the real return series of financial investment instruments. The data set can be downloaded from TSI's website[1]. TSI calculates the returns of assets included in the real return rates of financial investment instruments as follows.

- Deposit interest return rate calculations are made using the weighted average deposit interest rates applied to savings deposits actually opened in banks.

- The monthly average of the combined index of the 1st and 2nd session closing prices obtained from Borsa Istanbul (BIST 100) is used in the stock market index. The BIST 100 index is calculated from the stocks of the 100 companies with the largest market value and daily average trading volume among stocks that have

been traded on the stock exchange for at least 60 days.

- US Dollar and Euro are the 1-month average of the foreign exchange buying rate of the Central Bank of the Republic of Turkey for 1 US Dollar and 1 Euro.

- Istanbul Stock Exchange monthly average gold bullion prices (TL/gr) are used in gold prices.

- Real return rates of government domestic debt securities are calculated using the "BIST-KYD GDS All Index" within the scope of BIST-KYD GDS Indices published by Borsa Istanbul. This index reflects the yields of discounted and fixed-interest coupon government domestic debt securities traded in the debt securities market covering all maturities (TSI, 2024).

Figure 1 Real return series of financial investment instruments
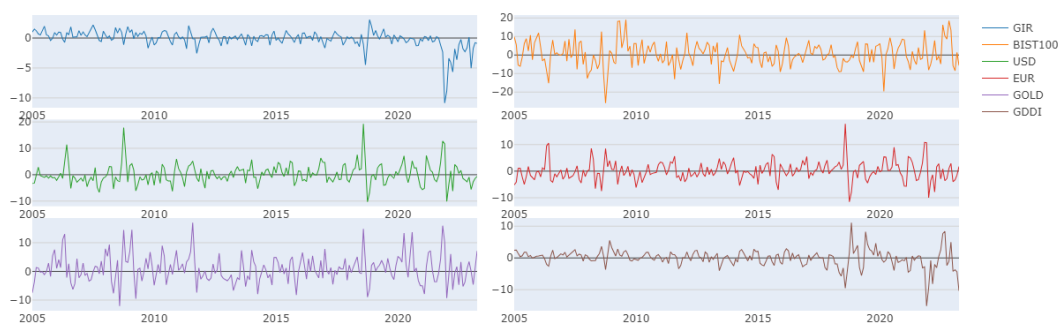
Source: Authors' own calculations

3. Volatility series

When working with a monthly frequency data set, choosing the appropriate volatility model becomes a bit complicated. Simply selecting the most appropriate model among historical or conditional models based on predictive power using error metrics is an insufficient approach because this approach neglects whether the model is reliable and valid. The assumptions of the applied model should be checked with diagnostic tests. Generally, meeting these assumptions is related to the number of observations. Generalized Autoregressive Conditional Heteroskedastic (GARCH) family models, which are widely used, have difficulties in their application

to low-frequency data sets due to both the number of observations and diagnostic tests.

Hwang and Pereira (2006) recommend using a minimum of 250 observations for ARCH(1) models and at least 500 observations for GARCH(1,1) models to reduce biases and convergence problems. GARCH estimates derived from low-frequency data also face the problem of temporal aggregation, as highlighted by Drost and Nijman (1993). In small samples, maximum likelihood estimates for the GARCH(1,1) model exhibit significant negative bias, and frequently the estimates do not comply with Bollerslev's non-negativity conditions, causing the estimated model to fail diagnos-

[1] https://data.tuik.gov.tr/Kategori/GetKategori?p=Enflation-ve-Fiyat-106

tic tests. Heteroskedastic Autoregressive (HAR) models, also from the GARCH family, are not a suitable approach for our study since they obtain monthly volatilities using daily data. Additionally, a commonly used approach to forecasting future volatility based on the average change in historical volatility is the exponentially weighted moving average (EWMA). According to Arı (2022), EWMA models showed better prediction results than both the GARCH family and Conditional Autoregressive Range (CARR) type models. Since CARR models are range-based and our data set does not contain the lowest and the highest values, they are not suitable for use in this study. In addition, the change of the lambda coefficient used in the EWMA model does not produce good results because it is sensitive to capturing volatility clouds.

For this reason, Open-High-Low-Close (OHLC) based volatility approaches, which do not have complex assumptions, are preferred, especially in volatility connectivity studies. Since the data we have does not include OHLC data, the most appropriate approach seems to be to accept the absolute values of the return data as volatility data.

It has been revealed in the literature that using monthly frequency data in volatility calculations has various advantages. For example, Figlewski (1997) found that the size of the forecast error doubled when daily data over a 24-month period was used to forecast volatility rather than monthly data. Because volatility mean reversion can be difficult to manipulate using high-frequency data, weekly or monthly volatility forecasts are sometimes preferable when applications extend beyond a 10-year time horizon. Current practice is to use the monthly absolute value as an indicator of macro volatility because many macroeconomic variables are only accessible in the monthly range (Poon, 2005).

The volatilities of real return rates of the series were obtained based on the study of Poon (2005) over the absolute values of returns. In other words, the absolute values of CPI-based monthly real returns were used as proxy values for volatility values. Figure 2 showcases the real return volatility series of these financial investment instruments. It is evident from the figure, the 2008 Global Financial Crisis, the currency shock in August 2018, the announcement of the COVID-19 pandemic in March 2020, the currency shock in December 2021, and the Russia-Ukraine war that started in late February 2022 caused an increase in volatility in almost all series.

Figure 2 Real return volatility series of financial investment instruments
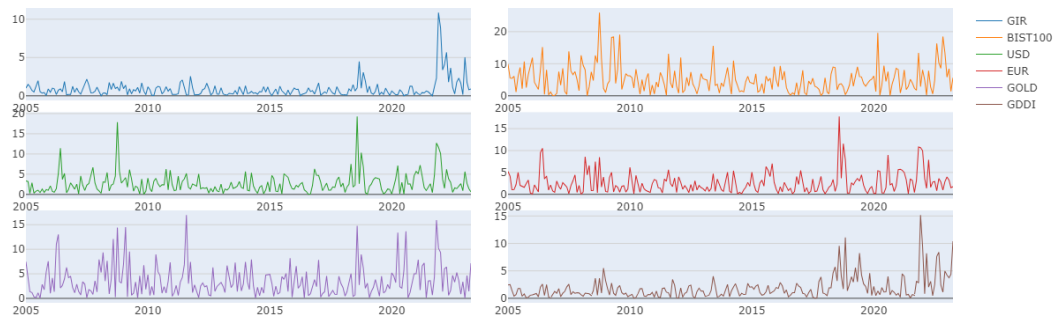
Source: Authors' own calculations

Summary statistics for the volatility series, as detailed in Table 1, reveal that the volatilities of the financial investment instruments exhibit skewness and excess kurtosis, and lack an ARCH effect. The Jarque-Bera (JB) test statistic further corroborates the non-normal distribution of these series at a 1% significance level. To assess stationarity, the Elliott-Rothenberg-Stock (ERS) unit root test, particularly suited for distributions with skewness and kurtosis, was employed. The results from the ERS test indicate that all series are stationary at their level. This confirms the appropriateness of the dataset for analysis using the TVP-VAR model.

Table 1 Summary statistics of volatility series

Statistics	GIR	BIST-100	USD	EUR	GOLD	GDDI
Mean	0.89	**5.159**	2,742	2,609	3,598	1,853
Variance	1.445	18,128	6,989	6,277	10,703	4,343
Skewness	4.742***	1.540***	2.823***	2.184***	1.653***	2.925***
Ex.Kurtosis	30.382***	3.362***	11.996***	7.120***	3.050***	11.245***
JB	9285.995***	190.556***	1611.339***	639.696***	185.509***	1472.901***
ERS	-4.594***	-2.839***	-5.928***	-3.211***	-2.788***	-3.259***
Q(20)	126.134***	17.411**	26.421***	19.954**	13.353	117.170***
Q²(20)	74.775***	22.470***	10.534	21.484***	11.410	55.156***
Spearman	GIR	BIST-100	USD	EUR	GOLD	GDDI
GIR	1.000					
BIST-100	0,097	1.000				
USD	0.162	0.258	1.000			
EUR	0.079	0.195	**0.546**	1.000		
GOLD	0.073	**0.049**	0.249	0.300	1.000	
GDDI	0.417	0.265	0.295	0.277	0.187	1.000

Note: ***, **, * denote significance levels at 1%, 5% and 10%.
Source: Authors' own calculations

Additionally, Table 1 elaborates the unconditional correlation matrix among the real return volatilities of financial investment instruments. The most substantial correlation is observed between USD and EUR with 0.546, followed by GDDI and GIR having a value of 0.417. With a value of 0.049, the correlation between the BIST-100 index and GOLD is the weakest.

3.1 Volatility connectedness via the TVP-VAR approach in the time domain

In their groundbreaking work, Diebold and Yilmaz (2012; 2014) introduced the connectedness approach, which uses both static and dynamic time-series network analysis to identify linkage and spillover within a given network. The dynamic method uses the rolling window VAR approach, while the static method uses a Vector Autoregression (VAR) model for the entire dataset. Thus, it reveals the return or volatility spreads within the network. This method has grown in popularity because it allows researchers to draw meaningful conclusions about networks. By using the TVP-VAR technique, Antonakakis et al. (2020) improved the connectedness approach that Diebold and Yilmaz (2014) had initially constructed. The TVP-VAR-based connectedness approach has several advantages, including insensitivity to outliers, no data loss, no require-

ment to specify arbitrary window size, and application to low-frequency datasets. This enhancement increases the variance-covariance matrix's adaptability by incorporating forgetting components and applying a Kalman filter. These developments make this method especially appropriate for researching the dynamic relationships between the volatilities of financial assets. The basic work of Koop and Korobilis (2013; 2014) on VAR models and the use of EWMA forgetting factors is incorporated into this methodology. The structure and features of the TVP-VAR(2) model, selected based on Bayesian Information Criteria (BIC), can be expressed as follows:

$$y_t = A_t z_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0, \Sigma_t) \qquad (2)$$

$$vec(A_t) = vec(A_{t-1}) + v_t \qquad v_t \sim N(0, S_t) \qquad (3)$$

In the above model, expressed in matrix form, the properties of vectors and related matrices are listed as follows. The vector z_{t-1} is a $2k \times 1$ vector, and the vector y_t has a dimension of $k \times 1$. A_t is a $k \times 2k$ matrix. Vectors ϵ_t and v_t are $k \times 1$ and $2k^2 \times 1$ dimensional vectors, respectively. Additionally, matrices Σ_t and S_t are time-varying variance-covariance matrices, with dimensions $k \times k$ and $2k^2 \times 2k^2$, respectively. The last vector is $vec(A_t)$ with a $2k^2 \times 1$ dimension. This structure of matrices and dimensions provides a thorough

framework for dissecting the time-varying relations in the model.

Generalized Forecast Error Variance Decomposition (GFEVD) values obtained from the vector model are the most basic element of the Diebold and Yılmaz methodology. This approach is very important to investigate the effects of shocks to the variables in a network on the dynamic relationships of the variables in a time series framework. Using the equation $y_t = \sum_{h=0}^{\infty} A_{ht}\epsilon_{t-h}$, the Time-Varying Parameter Vector Moving Average (TVP-VMA) model is obtained from the TVP-VAR model, where $A_0 = I_k$ and I_k is an identity matrix. As a result of this transformation, it becomes easier to evaluate the effect of the shock in variable j on the other variable i. The TVP-VMA transformation allows the impact of shocks on variable i to be measured by variable j in terms of both magnitude and direction at various time intervals. This assessment is made possible by a comprehensive calculation process described in the equation below, which gives a clear picture of the impact over time.

$$\widetilde{\phi}_{ij,t}^{g}(H) = \frac{\sum_{h=0}^{H-1}(\epsilon_i^T A_{ht}\Sigma_t\epsilon_j)^2}{(\epsilon_i^T\Sigma_t\epsilon_j)\sum_{h=0}^{H-1}(\epsilon_i^T A_h\Sigma_t A_{ht}^T\epsilon_i)} \quad (4)$$

One can obtain $\sum_{j=1}^{m}\widetilde{\phi}_{ij,t}^{g}(H) = 1$ and $\sum_{i,j=1}^{m}\widetilde{\phi}_{ij,t}^{g}(H) = k$ using the above equation. As mentioned before, Diebold and Yılmaz (2012; 2014) proposed connectedness indices based on the GFEVD technique. These indices are determined by computing the proportion of the forecast error change of a particular variable that can be attributed to shocks to other variables in the system. These calculations are made as follows:

Directional connectedness to others - TO

$$TO_{jt}(H) = \sum_{i=1,i\neq j}^{k} \widetilde{\phi}_{ij,t}^{g}(H) \quad (5)$$

Directional connectedness from others - FROM

$$FROM_{jt}(H) = \sum_{i=1,i\neq j}^{k} \widetilde{\phi}_{ji,t}^{g}(H) \quad (6)$$

Net total directional connectedness - NET

$$NET_{jt}(H) = \sum_{i=1,i\neq j}^{k} \widetilde{\phi}_{ij,t}^{g}(H) - \sum_{i=1,i\neq j}^{k} \widetilde{\phi}_{ji,t}^{g}(H) = TO_{jt} - FROM_{jt} \quad (7)$$

Total connectedness index – TCI

$$TCI_t(H) = k^{-1}\sum_{j=1}^{k} TO_{jt} \equiv k^{-1}\sum_{j=1}^{k} FROM_{jt} \quad (8)$$

Net pairwise directional connectedness – NPDC

$$NPDC_{ij,t}(H) = \widetilde{\phi}_{ij,t}^{g}(H) - \widetilde{\phi}_{ji,t}^{g}(H) \quad (9)$$

3.2 Volatility connectedness via the TVP-VAR approach in the frequency domain

We use the novel TVP-VAR frequency connectivity methodology introduced by Chatziantoniou et al. (2021). In fact, the novel approach skillfully combines the basic concepts proposed by Baruník and Krehlík (2018) and expanded by Antonakakis et al. (2020). Additionally, the methodology applied in this study is consistent with studies presented by Huang et al. (2023) and Akbulut et al. (2023), and it includes a comprehensive appendix that serves as a guide to put this methodology into practice. The TVP-VAR frequency connectivity model makes it possible to analyze volatility dependence over various periods by decomposing volatility dependence into short- and long-term components and accounting for changes in coefficients and the variance-covariance matrix over time. Interestingly, it does this without suffering the negative effects of adopting an arbitrary sliding window, including variable parameters, data loss, or outliers.

One can evaluate the spectral density of y_t at frequency ω by using the frequency response function given as $\Psi(e^{-i\omega}) = \sum_{h=0}^{\infty} e^{-i\omega h}\Psi_h$, where i is the imaginary unit of a complex number (the square root of -1) and ω is the frequency. A Fourier transform of the TVP-VMA(∞) effectively describes the spectral density of y_t across ω. Thus,

$$S_y(\omega) = \sum_{h=-\infty}^{\infty} E(y_t y_{t-h}')e^{-i\omega h} = \Psi(e^{-i\omega})\Sigma_t\Psi'(e^{i\omega h}). \quad (10)$$

As a result, we may calculate the GFEVD, which is frequency domain integration of the spectral density and the GFEVD. This calculation is performed as follows:

$$\phi_{ij,t}^{g}(\omega) = \frac{(\Sigma_t)_{jj}^{-1}\left|\sum_{h=0}^{\infty}(\Psi(e^{-i\omega h})\Sigma_t)_{ij,t}\right|^2}{\sum_{h=0}^{\infty}[\Psi(e^{-i\omega h})\Sigma_t\Psi'(e^{i\omega h})]_{ii}} \quad (11)$$

$$\widetilde{\phi}_{ij,t}^{g}(\omega) = \frac{\phi_{ij,t}^{g}(\omega)}{\sum_{i=1}^{k} \phi_{ij,t}^{g}(\omega)} \tag{12}$$

The equation for aggregating all frequencies within a specified range of interest is given by $\widetilde{\phi}_{ij,t}^{g}(d) = \int_{a}^{b} \widetilde{\phi}_{ij,t}^{g}(\omega)d\omega$, where d is defined as the interval (a, b), with both a and b falling within the spectrum (- π, π), ensuring that $a < b$. Thus, the frequency connectedness metrics can be outlined as follows:

$$TO_{jt}(d) = \sum_{i=1,i\neq j}^{k} \widetilde{\phi}_{ij,t}^{g}(d) \tag{13}$$

$$FROM_{jt}(d) = \sum_{i=1,i\neq j}^{k} \widetilde{\phi}_{ji,t}^{g}(d) \tag{14}$$

$$NET_{jt}(d) = \sum_{i=1,i\neq j}^{k} \widetilde{\phi}_{ij,t}^{g}(d) - \sum_{i=1,i\neq j}^{k} \widetilde{\phi}_{ji,t}^{g}(d) = \tag{15}$$

$$TO_{jt} - FROM_{jt}$$

$$TCI_{t}(d) = k^{-1} \sum_{j=1}^{k} TO_{jt} \equiv k^{-1} \sum_{j=1}^{k} FROM_{jt} \tag{16}$$

$$NPDC_{ij,t}(d) = \widetilde{\phi}_{ij,t}^{g}(H) - \widetilde{\phi}_{ji,t}^{g}(H) \tag{17}$$

Then all the frequency connectedness metrics may be computed, which gives information on spillovers in the particular frequency range denoted by d. This is stated in the following way:

$$\phi(H) = \sum_{d} \phi(d). \tag{18}$$

Within this context, a group of connectedness metrics [NPDC, TO, FROM, NET, TCI], each of which has been previously discussed, is represented by the symbol φ(·). This suggests that the totality of frequencies associated with the frequency connectedness measure is in agreement with the corresponding connectedness seen in the time domain.

4. Empirical findings

Total Connectedness Index (TCI): The average TCI results among financial investment instruments are presented in Table 2. The TCI shows that 40.37% of the variance in these financial investment instruments - specifically, the generalized estimated error variance - is due to the connectedness relationship in these networks. TCI values are 33.95% in the short term (1-12 months) and 6.41% in the long term (12-inf).

The diagonal entries of the 6x6 matrix depicted in Table 2 represent the variance in prediction errors attributed to the variables themselves. In contrast, the off-diagonal elements detail how the error variances are broken down among the variables. In accordance with the data, the BIST-100 index is the investment instrument with the highest self-induced volatility in total data with 72.91%. In the short run, the BIST-100 index accounts for 64.77%, whereas in the long run, it is realized as GDDI with 12.3%. As a result, it is possible to say that the BIST-100 index and GDDI are investment instruments least affected by the network. In contrast, investment instruments with the lowest self-induced volatility spillovers and thus most affected by network spillovers are USD (47.32% and 40.24%, respectively) in the total and short run, and EUR (6.43%) in the long run. USD is the investment instrument that is most affected by the network, with 52.68% (FROM), and with 63.13% (TO), it also has the most volatility spreads in the network. In the short and long run, USD is the investment instrument most affected by the network with 43.77% (FROM) and 8.9% (FROM), respectively, while USD (in the short run) and EUR (in the long run) exhibited the highest volatility spreads in the network with 53.96% (TO) and 9.52% (TO), respectively.

Net total directional connectedness: NET values are obtained from the difference between the volatility emitted by a variable in the network and the volatility affecting it. A positive NET value indicates that the variable is a net volatility transmitter in the network, while a negative NET value indicates that the variable is a net volatility receiver. According to the data given in Table 2, USD is the highest volatility transmitter in the network with 10.45%, followed by EUR (8.46%) and GDDI (1.12%), respectively.

The highest net volatility transmitter in the short term is USD with 10.19%, while in the long run, it is EUR with 2.85%. GIR acts as the highest net volatility receiver for all periods.

Table 2 Total, short-run, and long-run average TCI

Total	GIR.Total	BIST-100.Total	USD.Total	EUR. Total	GOLD. Total	GDDI. Total	FROM. Total
GIR	65.16	1.13	6.27	5.02	3.13	19.27	34.84
BIST-100	0.80	72.91	12.16	5.48	3.88	4.77	27.09
USD	2.23	8.54	47.32	22.57	11.61	7.72	52.68
EUR	1.72	4.21	22.04	50.57	13.66	7.8	49.43
GOLD	1.46	2.85	12.89	15.55	63.77	3.48	36.23
GDDI	15.09	4.11	9.76	9.27	3.7	58.08	41.92
TO	21.3	20.84	63.13	57.89	35.98	43.04	242,19
Inc.Own	86.47	93.75	110.45	108.46	99.76	101.12	**TCI**
Net	-13.53	-6.25	10.45	8.46	-0.24	1.12	**40.37**
Short-run	GIR.1-12	BIST-100.1-12	USD.1-12	EUR.1-12	GOLD.1-12	GDDI.1-12	FROM.1-12
GIR	55.34	1.04	4.92	3.83	2.52	15.17	27.48
BIST-100	0.69	64.77	11.2	4.95	3.29	4.25	24.38
USD	1.8	7.83	40.24	18.5	9.41	6.23	43.77
EUR	1.46	4.0	19.01	44.14	11.4	6.88	42.76
GOLD	1.23	2.71	11.37	13.91	57.27	3.05	32.26
GDDI	11.76	3.71	7.45	7.18	2.97	45.78	33.07
TO	16.94	19.29	53.96	48.37	29.59	35.57	203.72
Inc.Own	72.28	84.06	94.19	92.51	86.86	81.35	**TCI**
Net	-10.54	-5.08	10.19	5.62	-2.67	2.5	**33.95**
Long-run	GIR.12-Inf	BIST-100.12-Inf	USD.12-Inf	EUR.12-Inf	GOLD.12-Inf	GDDI.12-Inf	FROM.12-Inf
GIR	9.82	0.1	1.35	1.19	0.62	4.1	7.36
BIST-100	0.11	8.14	0.96	0.53	0.6	0.52	2.72
USD	0.43	0.71	7.09	4.08	2.2	1.49	8.9
EUR	0.26	0.21	3.02	6.43	2.26	0.93	6.67
GOLD	0.23	0.14	1.52	1.64	6.5	0.43	3.97
GDDI	3.32	0.4	2.31	2.09	0.73	12.3	8.85
TO	4.37	1.55	9.17	9.52	6.4	7.47	38.47
Inc.Own	14.19	9.69	16.26	15.95	12.9	19.77	**TCI**
Net	-2.99	-1.17	0.26	2.85	2.43	-1.38	**6.41**

Note: The findings of the study are derived from employing a TVP-VAR model with a two-lag order, selected on the basis of the Bayesian Information Criterion (BIC), and involve a forecast error variance decomposition looking 10 steps ahead. This analysis utilized the R programming language and the "ConnectednessApproach" package developed by Gabauer (2022). Important terminology used in the study includes "Inc.Own" to denote one's own contributions, "TCI" for the Total Connectedness Index, "NET" indicating Net Total Connectedness, and "NPT" representing Net Pairwise Total Connectedness. The analysis categorizes the short run as a period of 1-12 months and the long run as any period extending beyond 12 months.
Source: Authors' own calculations

Dynamic Total Connectedness Index: Dynamic TCI illustrates how the interconnections among the variables fluctuate over time. Figure 3 not only showcases the dynamic progression of the average total connectedness index but also its decomposition into short-term (1-12 months) and long-term (beyond 12 months) components. As depicted in Figure 3, dynamic TCIs undergo notable variations as time progresses. Accordingly, the highest connectedness between the series occurred in September 2018, with a value of 55.95%, which this points to the exchange rate shock experienced at

that time. Connectedness, which started to decline after 2018, rose again (47.08%) in March 2020, the date when the COVID-19 pandemic was declared. Similarly, with the impact of the Russia-Ukraine war that started at the end of February 2022, the interconnectedness between the series increased to a serious level of 54.42%. In addition, the dynamic connectedness between the series increased significantly during the 2008 global financial crisis and the exchange rate shock in December 2021.

Figure 3 Dynamic TCI plots
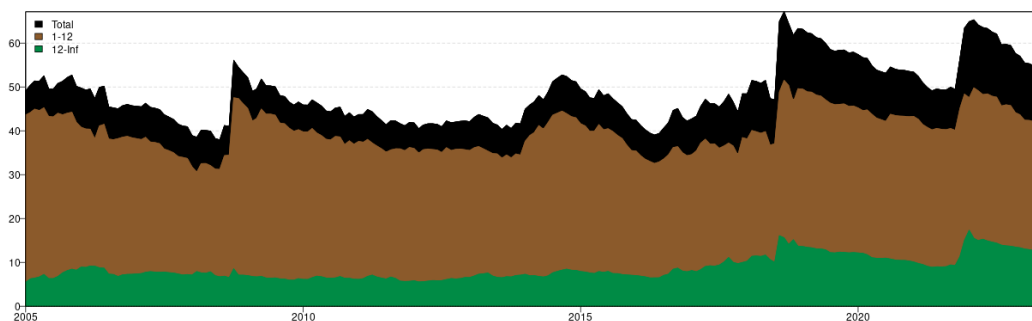
Source: Authors' own calculations

Robustness analysis: We followed two ways to examine the sensitivity of our results: First, we examined return connectivity of the variables because in spillover analysis, return transmission and connectedness are used as an indicator of robustness. The second is to test the approach using different forecasting horizons. We make an alternative h-step forward estimation for the forecasting error variance decomposition obtained from the TVP-VAR model. We apply a 20-steps ahead forecasting horizon instead of 10. According to the average TCI tables given in Appendix A, the spillover index appears to have a comparable structure and pattern. This shows that connectivity analysis is not always sensitive to returns and forecast horizons. Similar approaches have also been commonly used as robustness checks in numerous previous studies (Diebold & Yilmaz, 2012; 2014; Billah et al., 2022).

Net Pairwise Directional Connectedness (NPDC): NPDC indicates bilateral volatility transfer between two variables, with positive NPDC values signifying dominance of one investment instrument over another, and negative values indicating the opposite. The Pairwise Connectedness Index (PCI) quantifies the intensity of the relationship between variables, with values spanning from zero (indicating a weak connection) to one (signifying a strong connection). Examining the pairwise volatility connectedness is essential for grasping how volatility shocks propagate across different invest-

ment instruments, particularly in times of crises or epidemics, and observing how these interactions change over time.

Table 3 displays NPDCI and PCI values, illustrating the connectedness between various investment instruments. The highest level of the pairwise connectedness is found between USD and EUR across all timeframes, with PCI values reaching 62.99% overall, 52.95% for the short run, and 10.06% for the long run. With an NPDCI of 0.53% overall and 1.05% in the long term, USD acts as a net transmitter, showing its influence over EUR. Conversely, in the short term, EUR becomes the net transmitter with an NPDCI of -0.51%, indicating its dominance over USD. In other words, in the total and long run, USD is the net volatility transmitter, EUR is the net volatility receiver, while in the short term, EUR is the net volatility transmitter and USD is the net volatility receiver. According to PCI values, the lowest pairwise connectedness for all periods (2.73% (total), 2.44% (short term), 0.29% (long term)) is found between GIR and the BIST-100 index. Based on an NPDCI value, as GIR is a net volatility transmitter in total (0.33%) and short term (0.34%), it acts as a net volatility receiver in the long term (-0.01%). The minimal NPDCI and PCI values between the two investment instruments indicate that these two investment instruments can be used effectively for portfolio diversification. As a matter of fact, according to the data in Table 3, it is observed that

the lowest bilateral relationship emerged between GOLD and GDDI with an NPDCI value of 0.08%. Therefore, we have demonstrated that these two investment instruments are preferable for portfolio diversification.

Table 3 NPDCI and PCI

Time Domain	BIST-100	USD	EUR	GOLD	GDDI
GIR	0.33 (2.73)	4.04 (14.98)	3.30 (12.18)	1.67 (7.73)	4.18 (42.75)
BIST-100		3.62 (28.82)	1.26 (14.54)	1.03 (9.60)	0.66 (12.77)
USD			0.53 (62.99)	-1.28 (37.10)	-2.04 (28.57)
EUR				-1.88 (41.81)	-1.46 (27.52)
GOLD					-0.21 (11.61)
Frequency Domain Short-run (1 - 12 months)					
GIR	0.34 (2.44)	3.11 (11.81)	2.37 (9.49)	1.29 (6.20)	3.40 (33.42)
BIST-100		3.36 (26.48)	0.94 (13.44)	0.58 (8.53)	0.53 (11.47)
USD			-0.51 (52.93)	-1.95 (31.47)	-1.22 (22.36)
EUR				-2.50 (36.20)	-0.30 (22.58)
GOLD					0.08 (9.68)
Frequency Domain Long-run (12 – inf months)					
GIR	-0.01 (0.29)	0.92 (3.16)	0.92 (2.69)	0.38 (1.52)	0.77 (9.33)
BIST-100		0.25 (2.33)	0.32 (1.10)	0.45 (1.06)	0.12 (1.30)
USD			1.05 (10.06)	0.67 (5.63)	-0.81 (6.21)
EUR				-0.61 (5.61)	-1.16 (4.94)
GOLD					-0.92 (1.92)

Note: PCIs are presented in parentheses.
Source: Authors' own calculations

The net pairwise directional dynamic connectedness between the two variables is presented visually in Figure 4. GIR was a net receiver of volatility against other investment instruments throughout the whole period, and the impact of shocks received from USD, EUR, GOLD and GDDI has increased significantly after 2018. Similarly, the BIST-100 index acted a net receiver of volatility against other investment instruments. On the other hand, USD and EUR were net receivers of volatility against GOLD until 2008, while GOLD acted as a net receiver of volatility against these two investment instruments after 2008. Being a net volatility receiver against USD and EUR, the shocks received by the GDDI investment instrument increased after 2018. In addition, while EUR was a net volatility receiver against GDDI in the short term between 2008 and 2015, it became a net volatility transmitter investment instrument after 2015. Finally, while GOLD was a net volatility receiver against GDDI in the 2008-2018 period, it became a net volatility transmitter in other periods.

Considering the studies using the connectivity approach, the findings overlap with the study of Şak and Öcal Özkaya (2022). Similar to that study, USD and EURO financial instruments were found to be transmitters, and GOLD and BIST were found to be receivers. Although Şak and Öcal Özkaya (2022) used the rolling window VAR approach in their study and the period of the data set was longer, obtaining similar results shows that the approach is robust and that the relationship between financial assets in Turkey has a characteristic structure.

Figure 4 Net pairwise directional connectedness
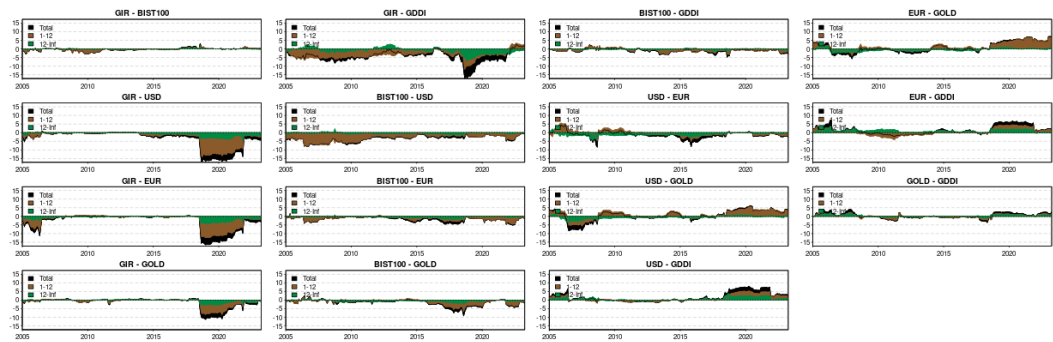
Source: Authors' own calculations

Network plots: Figure 5 visually presents the bilateral relationships between investment instruments. This visual representation allows us to see the dynamic interaction between variables in a TVP-VAR based frequency connectedness analysis. Blue and yellow nodes indicate that the variable is a net volatility transmitter and a net volatility receiver, respectively. The direction and thickness of the arrows determine the direction and strength of the volatility spread between the variables. In both the overall and short-run network plots, the investment instruments represented by USD, EUR, and GDDI emerge as having dominant volatility within the network. They are the primary drivers of network dynamics, influencing the flow and transmission of volatility among the various investment instruments included in the analysis. Moreover, for both periods, the strongest volatility spillover is from GDDI to GIR. In the long term, USD, EUR and GOLD are the most dominant investment instruments in the network. Although Yılmaz and Kılıç (2022) examined the interaction between variables with a different methodology, they obtained results parallel to our findings. For example, the fact that USD and EURO are more dominant, in other words, a transmitter, as opposed to the GIR variable, can be seen as similar results. Another similarity in their studies is that the interaction between USD and EURO is an indication of a strong bilateral connectedness.

Figure 5 Network plots
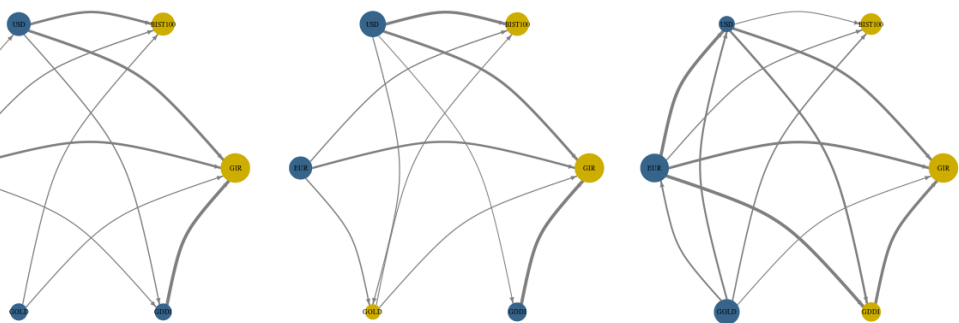
Note: A time-domain network (left), a short-run network (middle), a long-run network (right).
Source: Authors' own calculations

5. Conclusion

In this study, the volatility pass-through between the real returns of selected financial investment instruments is analyzed using the innovative TVP-VAR connectedness approach in the frequency domain. For this purpose, firstly, the real return

volatilities of financial investment instruments are obtained based on the study of Poon (2005), and then the time and frequency (short and long terms) connectedness between these investment instruments is revealed by using the method based on the approach of Diebold and Yilmaz (2012; 2014). According to the findings, while the total connectedness index (TCI) was 40.37%, it was 33.95% in the short run and 6.41% in the long run. The BIST-100 index is the investment instrument with the highest self-induced volatility in total and in the short term, which means that the BIST-100 index is the investment instrument that is least affected by the network. In the long run, the investment instrument with the highest self-driven volatility is GDDI. The investment instrument with the lowest self-induced volatility and thus most affected by the network is USD both in total and in the short term, while EUR is observed in the long term. This is an expected situation. As a matter of fact, since USD and EUR have become global investment instruments, shocks in other investment instruments can spread to these investment instruments in a short time. On the other hand, USD is an investment instrument that both transmits the highest volatility in the network and receives the highest volatility from the network. When we look at the dynamic total connectedness, we observe that the connectedness between these investment instruments increased significantly during periods of turbulence at both global and local levels, such as the 2008 global financial crisis, the currency shock experienced in our country in 2018 and 2021, the COVID-19 outbreak announced in March 2020, and the Russia-Ukraine war that started in February 2022. According to the NET values of the variables, USD (in total and in the short term) and EUR (in the long term) is the highest net volatility transmitter, while GIR is the highest net volatility receiver for all periods. When it comes to results of the net bilateral connectedness between variables, the highest net bilateral connectedness for all periods is between USD and EUR, while the lowest net bilateral connectedness is between GIR and BIST-100 index. A low degree of connectedness among investment instruments indicates that they can be used effectively in portfolio diversification. Therefore, this study reveals important results for investors, portfolio managers, risk managers and foreign exchange trading companies.

Another interesting result is that the spread through the network is higher in the short term, i.e. 84% of the total risk spread occurs in the short term. This is because USD and EUR are the most effective investment instruments in the network and have high volatility. This is an expected result. In an emerging market like Turkey, inflation is mostly caused by the effect of USD and EUR, the reserve currencies in the world, on raw material prices. Furthermore, this situation is affecting all financial markets, especially the low deposit interest rate and high inflation that emerged with the economic policies implemented in the last five years led to a high depreciation of the Turkish Lira. As a result, negative real deposit interest drove investors away from the local currency and increased demand in foreign exchange markets. Moreover, it led investors to stay away from local financial instruments such as BIST-100, GDDI, and GIR, and it impacted the connectedness index.

In addition, interest in the cryptocurrency market in Turkey is increasing dramatically. Considering the number of wallets in the country, it ranks among the top five countries in the world with the highest number of investors in the cryptocurrency market. This situation makes it necessary to reveal the relationship between cryptocurrency markets and traditional investment instruments. For example, the inflation-discounted real Bitcoin return, or the risk spread between a crypto stock market index and Turkish investment instruments can be revealed with the connectivity approach.

Another effective factor is the impact of emerging geopolitical risks on Turkish financial markets. Türkiye is affected by the conflicts occurring around it due to its geographical location. To reveal this situation, the interaction between geopolitical risk indices and markets can be examined with the same methods. It is expected that asymmetric connectivity analyses, especially those applied to reveal asymmetric effects, will provide more detailed results. After asymmetric return and volatility modeling, the difference between positive and negative volatility spreads can be analyzed using dynamic quantile connectedness and extended joint connectedness approaches.

We can summarize our recommendations for future studies.

1. Investigate the relationship between cryptocurrency markets and traditional investment instruments in Turkey. Provide in-

sights into risk management and portfolio diversification by determining how traditional markets impact or affect cryptocurrencies. Use connectedness approaches and cointegration approaches as a method when revealing this relationship.

2. Additionally, unlike the above suggestion, examining the asymmetric effects of cryptocurrency volatility on the dynamic total connectedness index of traditional investment instruments obtained in this study with Johansen cointegration or ARDL/NARDL bounds test methods can be the subject of a future study.

3. Examine the impact of geopolitical risks specific to Türkiye on Turkish financial markets. Another study suggestion is to reveal the dynamic interconnectedness between geopolitical risk indices and traditional investment instruments. This relationship can also be investigated with methods that reveal symmetric and asymmetric effects.

4. The difference in connectivity that occurs before and after exchange rate shocks is a separate research topic that needs to be addressed. Thus, it will be revealed what investment instrument groups will provide optimum portfolio diversity in the pre- and post-crisis periods.

5. Finally, examining the risk transfer with these conventional investment instruments, especially with banking and technology sectors, will help develop sector-oriented investment and risk strategies.

References

1. Ahmed, A. D. & Huo, R. (2021). Volatility transmissions across international oil market, commodity futures and stock markets: Empirical evidence from China. *Energy Economics*, *93*, 104741. https://doi.org/10.1016/j.eneco.2020.104741

2. Akbulut, N. & Ari, Y. (2023). TVP-VAR Frequency Connectedness between the Foreign Exchange Rates of Non-Euro Area Member Countries. *Folia Oeconomica Stetinensia*, *23*(2), 1-23. https://doi.org/10.2478/foli-2023-0016

3. Antonakakis, N. & Kizys, R. (2015). Dynamic spillovers between commodity and currency markets. *International Review of Financial Analysis*, *41*, 303-319. https://doi.org/10.1016/j.irfa.2015.01.016

4. 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*, *13*(4), 84. https://doi.org/10.3390/jrfm13040084

5. Arı, Y. (2022). Chasing Volatility of USD/TRY Foreign Exchange Rate: The Comparison of CARR, EWMA, and GARCH Models. *Ekoist: Journal of Econometrics and Statistics*, (37), 107-127. https://doi.org/10.26650/ekoist.2022.37.1113670

6. Badshah, I. U., Frijns, B. & Tourani-Rad, A. (2013). Contemporaneous spill-over among equity, gold, and exchange rate implied volatility indices. *Journal of Futures Markets*, *33*(6), 555-572. https://doi.org/10.1002/fut.21600

7. Baruník, J. & Krehlík, T. (2018). Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics*, *16*(2), 271-296. https://doi.org/10.1093/jjfinec/nby001

8. Başarır, Ç. (2019). Altın ve Hisse Senedi Getirileri Arasındaki Nedensellik İlişkisi: Türkiye Örneği. *Trakya Üniversitesi Sosyal Bilimler Dergisi*, *21*(2), 475-490. https://doi.org/10.26468/trakyasobed.472190

9. Baur, D. G. & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review*, *45*(2), 217-229. https://doi.org/10.1111/j.1540-6288.2010.00244.x

10. Bhunia, A. (2013). Cointegration and Causal Relationship Among Crude Price, Domestic Gold Price and Financial Variables: An Evidence of BSE And NSE. *Journal of Contemporary Issues in Business Research*, *2*(1), 1-10.

11. Billah, M., Karim, S., Naeem, M. A. & Vigne, S. A. (2022). Return and volatility spillovers between energy and BRIC markets: Evidence from quantile connectedness. *Research in International Business and Finance*, *62*, 101680. https://doi.org/10.1016/j.ribaf.2022.101680

12. Cao, J. & Wen, F. (2019). The impact of the cross-shareholding network on extreme price movements: Evidence from China. *Journal of Risk*, *22*(2), 79-102. https://doi.org/10.21314/JOR.2019.423

13. Cihangir, M., Polat, M. A. & Çalışkan, U. (2020). Türkiye'de alternatif finansal yatırım araçları arasındaki dinamik etkileşim: uygulamalı bir analiz. *Yaşar Üniversitesi E-Dergisi*, *15*(60), 920-940.

14. Cingoz, F. & Kendirli, S. (2019). Altın Fiyatları, Döviz Kuru ve Borsa İstanbul Arasındaki İlişki. *Finans Ekonomi ve Sosyal Araştırmalar Dergisi*, *4*(4), 545-554. https://doi.org/10.29106/fesa.649254

15. Chatziantoniou, I., Gabauer, D. & Gupta, R. (2021). *Integration and Risk Transmission in the Market for Crude Oil: A Time-Varying Parameter Frequency Connectedness Approach* (Department of Economics Working Paper No. 2021-47). Pretoria: University of Pretoria, Department of Economics.

16. Diebold, F. X. & Yilmaz, K. (2012). Better to Give than to Receive: Predictive Measurement of Volatility Spillovers. *International Journal of Forecasting*, *28*, 57-66. https://doi.org/10.1016/j.ijforecast.2011.02.006

17. Diebold, F. X. & Yılmaz, K. (2014). On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms. *Journal of Econometrics*, *182*, 119-134. https://doi.org/10.1016/j.jeconom.2014.04.012

18. Drost, F. C. & Nijman, T. E. (1993). Temporal Aggregation of GARCH Processes. *Econometrica*, *61*(4), 909-927. https://doi.org/10.2307/2951767

19. Figlewski, S. (1997). Forecasting Volatility. *Financial Markets, Institutions and Instruments*, *6*(1), 1-88. https://doi.org/10.1111/1468-0416.00009

20. Gabauer, D. (2022). Connectednessapproach: Connectedness approach (Version 1.1) [R package]. Retrieved from https://cran.r-project.org/package=ConnectednessApproach

21. Güney, S. & Ilgin, K. S. (2019). Yatırım araçlarının BİST-100 endeksi üzerindeki etkisinin değerlendirilmesi. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, (53), 226-245. https://doi.org/10.18070/erciyesiibd.449164

22. Hein, L. U. (2015). Investigating correlation and volatility transmission among equity, gold, oil and foreign exchange. *MaRBLe*, *2*. https://doi.org/10.26481/marble.2015.v2.99

23. Huang, J., Chen, B., Xu, Y. & Xia, X. (2023). Time-frequency volatility transmission among energy commodities and financial markets during the COVID-19 pandemic: A Novel TVP-VAR frequency connectedness approach. *Finance Research Letters*, *53*. https://doi.org/10.1016/j.frl.2023.103634

24. Hwang, S. & Valls Pereira, P. L. (2006). Small sample properties of GARCH estimates and persistence. *The European Journal of Finance*, *12*(6-7), 473-494. https://doi.org/10.1080/13518470500039436

25. Jain, A. & Biswal, P. C. (2019). Does internet search interest for gold move the gold spot, stock and exchange rate markets? A study from India. *Resources Policy*, *61*, 501-507. https://doi.org/10.1016/j.resourpol.2018.04.016

26. Koop, G. & Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics*, *177*(2), 185-198. https://doi.org/10.1016/j.jeconom.2013.04.007

27. Koop, G. & Korobilis, D. (2014). A new index of financial conditions. *European Economic Review*, *71*, 101-116. https://doi.org/10.1016/j.euroecorev.2014.07.002

28. Önem, H. B. (2021). Altın, Gümüş ile BİST Madencilik Endeksi Getirileri Arasındaki Volatilite Etkileşiminin Diagonal VECH GARCH Modeliyle Analizi. *OPUS International Journal of Society Researches*, *18*(43), 6220-6240. https://doi.org/10.26466/opus.905547

29. Poon, S. H. & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*(2), 478-539. https://doi.org/10.1257/.41.2.478

30. Poon, S. H. (2005). *A Practical Guide for Forecasting Financial Market Volatility*. John Wiley & Sons.

31. Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics*, *85*(4), 777-792. https://doi.org/10.1162/003465303772815727

32. Roy, R. P. & Roy, S. S. (2017). Financial contagion and volatility spillover: An exploration into Indian commodity derivative market. *Economic Modelling*, *67*, 368-380. https://doi.org/10.1016/j.econmod.2017.02.019

33. Sumner, S., Johnson, R. & Soenen, L. (2010). Spillover effects among gold, stocks, and bonds. *Journal of Centrum Cathedra*, *3*(2), 106-120. https://doi.org/10.7835/jcc-berj-2010-0041

34. Shahrazi, M. M., Elmi, Z. M., Abounoori, E. & Rasekhi, S. (2014). The influence of structural changes in volatility on shock transmission and volatility spillover among Iranian gold and foreign exchange markets. *Iranian Economic Review*, *18*(2), 73-86.

35. Şak, N. & Öcal Özkaya, H. G. (2022). Türkiye'de 2000 yılı sonrasında altın, borsa, döviz kuru piyasaları etkileşimi ve volatilite yayılım etkisi. *EKOIST Journal of Econometrics and Statistics*, *37*, 237-256. https://doi.org/10.26650/ekoist.2022.37.1199285

36. Şeker, K. (2021). Türkiye'de Finansal Sistemden Sağlanan Çeşitli Getirilerin VAR Modeli ile Etkileşimlerinin Analizi. *Muhasebe ve Finansman Dergisi*, (92), 183-206. https://doi.org/10.25095/mufad.947821

37. Şenol, Z. (2021). Borsa endeksi, döviz kuru, faiz oranları ve CDS primleri arasındaki oynaklık yayılımları: Türkiye örneği. *Business and Economics Research Journal*, *12*(1), 111-126. https://doi.org/10.20409/berj.2021.313

38. Topaloğlu, E. E. (2020). Borsa İstanbul pay endekslerinin volatilite yapısı ve volatilite yayılımı: GARCH ve MGARCH modelleri ile BİST Sınai ve Mali Endeksleri örneği. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, (63), 17-38.

39. TSI: Turkish Statistical Institute (2024). *Data Portal for Statistics - Finansal Yatırım Araçları İstatistikleri Metaverisi.* https://data.tuik.gov.tr/Kategori/GetKategori?p=Enflasyon-ve-Fiyat-106

40. Verma, P. & Jackson, D. (2012). The dynamic relationship between adrs, interest rates, exchange rates and their spillover effects. *North American Journal of Finance and Banking Research*, *6*(6), 1-26.

41. Yılmaz, Y. & Kılıç, E. (2022). Geleneksel Yatırım Araçları Arasında Getiri ve Volatilite Etkileşiminin Var-Egarch Modeli ile Analizi. *Dicle Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (30), 1-14.

42. Wang, Y. S. & Chueh, Y. L. (2013). Dynamic transmission effects between the interest rate, the US dollar, and gold and crude oil prices. *Economic Modelling*, *30*, 792-798. https://doi.org/10.1016/j.econmod.2012.09.052

43. Wen, F., Cao, J., Liu, Z. & Wang, X. (2021). Dynamic volatility spillovers and investment strategies between the Chinese stock market and commodity markets. *International Review of Financial Analysis*, *76*, 101772. https://doi.org/10.1016/j.irfa.2021.101772

Appendix A: Findings of robustness analysis

Table A1 Total, short-run, and long-run average TCI (return connectedness)

Total	GIR.Total	BIST100.Total	USD.Total	EUR.Total	GOLD.Total	GDDI.Total	FROM.Total
GIR	55.22	1.29	6.23	6.75	4.09	26.41	44.78
BIST-100	1.08	53.79	16.17	10.87	5.38	12.72	46.21
USD	1.1	13.56	39.77	23.71	12.75	9.11	60.23
EUR	0.96	8.99	23.87	41.07	16.47	8.63	58.93
GOLD	1.29	6.67	16.19	17.24	54.57	4.04	45.43
GDDI	18.77	8.91	11.44	11.32	4.04	45.52	54.48
TO	23.2	39.43	73.9	69.89	42.73	60.91	310.05
Inc. Own	78.42	93.22	113.67	110.96	97.3	106.43	cTCI/TCI
Net	-21.58	-6.78	13.67	10.96	-2.7	6.43	62.01/51.68
Short-run	GIR.1-12	BIST100.1-12	USD.1-12	EUR.1-12	GOLD.1-12	GDDI.1-12	FROM.1-12
GIR	44.09	1.08	4.72	5.18	3.07	20.03	34.08
BIST-100	0.82	45.1	13.34	8.95	4.55	10.12	37.78
USD	0.91	10.79	32.78	19.67	10.56	7.74	49.66
EUR	0.79	7.45	20.27	34.98	13.58	7.41	49.49
GOLD	1.13	5.42	13.96	15.08	46.89	3.65	39.25
GDDI	14.44	7.51	8.89	8.51	3.09	35.84	42.43
TO	18.08	32.25	61.17	57.38	34.85	48.95	252.69
Inc. Own	62.17	77.35	93.95	92.36	81.74	84.8	cTCI/TCI
Net	-16	-5.52	11.51	7.89	-4.4	6.52	50.54/42.11
Long-run	GIR.12-Inf	BIST100.12-Inf	USD.12-Inf	EUR.12-Inf	GOLD.12-Inf	GDDI.12-Inf	FROM.12-Inf
GIR	11.14	0.22	1.51	1.57	1.02	6.38	10.69
BIST-100	0.26	8.7	2.84	1.92	0.83	2.59	8.43
USD	0.18	2.77	6.99	4.05	2.2	1.37	10.57
EUR	0.17	1.54	3.6	6.09	2.89	1.23	9.43
GOLD	0.16	1.25	2.23	2.15	7.68	0.39	6.18
GDDI	4.33	1.4	2.55	2.82	0.95	9.68	12.05
TO	5.11	7.18	12.73	12.5	7.89	11.95	57.36
Inc. Own	16.25	15.87	19.72	18.6	15.57	21.63	cTCI/TCI
Net	-5.58	-1.26	2.16	3.07	1.7	-0.09	11.47/9.56

Source: Authors' own calculations

Table A2 Total, short-run, and long-run average TCI (volatility connectedness with forecast horizon 20)

Total	GIR.Total	BIST100.Total	USD.Total	EUR. Total	GOLD. Total	GDDI. Total	FROM.Total
GIR	63.58	0.51	6.9	5.96	2.42	20.61	36.42
BIST-100	0.45	76.07	11.9	4.79	3.6	3.18	23.93
USD	2.75	7.12	46.55	24.57	11.04	7.97	53.45
EUR	2.6	2.95	23.89	49.15	13.25	8.17	50.85
GOLD	1.36	2.3	13.02	15.49	64.24	3.59	35.76
GDDI	16.72	2.19	10.49	10.09	3.34	57.18	42.82
TO	23.87	15.07	66.21	60.91	33.65	43.53	243.23
Inc.Own	87.45	91.14	112.76	110.06	97.89	100.7	cTCI/TCI
Net	-12.55	-8.86	12.76	10.06	-2.11	0.7	48.65/40.54
Short-run	GIR.1-12	BIST100.1-12	USD.1-12	EUR.1-12	GOLD.1-12	GDDI.1-12	FROM.1-12
GIR	58.03	0.5	5.97	5.12	2.1	18.25	31.94
BIST-100	0.42	71.13	11.37	4.58	3.29	3.02	22.69
USD	2.44	6.75	42.55	22.15	9.89	7.15	48.39
EUR	2.42	2.86	22.09	45.62	12.09	7.66	47.11
GOLD	1.23	2.24	12.14	14.61	60.64	3.34	33.56
GDDI	14.66	2.08	9.08	8.79	2.95	50.63	37.56
TO	21.17	14.42	60.65	55.25	30.34	39.42	221.25
Inc.Own	79.2	85.55	103.2	100.87	90.97	90.05	cTCI/TCI
Net	-10.77	-8.27	12.26	8.14	-3.23	1.87	44.25/36.87
Long-run	GIR.12-Inf	BIST100.12-Inf	USD.12-Inf	EUR.12-Inf	GOLD.12-Inf	GDDI.12-Inf	FROM.12-Inf
GIR	5.55	0.02	0.94	0.84	0.32	2.36	4.48
BIST-100	0.03	4.94	0.54	0.21	0.31	0.16	1.24
USD	0.3	0.37	4	2.42	1.14	0.82	5.06
EUR	0.18	0.09	1.8	3.53	1.16	0.52	3.74
GOLD	0.13	0.06	0.88	0.88	3.6	0.25	2.2
GDDI	2.06	0.11	1.41	1.3	0.38	6.55	5.27
TO	2.7	0.65	5.56	5.66	3.31	4.1	21.98
Inc.Own	8.26	5.59	9.56	9.18	6.91	10.65	cTCI/TCI
Net	-1.78	-0.59	0.5	1.91	1.11	-1.16	4.40/3.66

Source: Authors' own calculations

Lorena Škuflić
University of Zagreb
Faculty of Economics and Business
10000 Zagreb, Croatia
lskuflic@efzg.hr

Dora Walter
University of Zagreb
Faculty of Economics and Business
10000 Zagreb, Croatia
dwalter@net.efzg.hr

Valentina Vučković
University of Zagreb
Faculty of Economics and Business
10000 Zagreb, Croatia
vvuckovic@efzg.hr

THE COST OF FISCAL AUSTERITY: A SYNTHETIC CONTROL APPROACH

Abstract

Purpose: This paper analyses economic and social impact of fiscal austerity policies on economic growth and income distribution. In response to the European public debt crisis, austerity measures were implemented in 2010 to decrease the budget deficit and avoid the default of the government debt, but have also caused negative effects on the whole economy.

Methodology: In order to evaluate the effectiveness of fiscal austerity, the synthetic control method (SCM) is applied by creating a synthetic counterfactual from European countries. Greece is used as an example to assess the impact of the aforementioned policy due to having experienced fiscal consolidation to a much larger extent than other crisis-affected countries.

Results: Fiscal austerity causes a decline in real GDP per capita compared to its pre-austerity level. Additionally, it results in higher unemployment and a more unequal distribution of income in the initial years following the treatment.

Conclusion: The objective of fiscal austerity, i.e. the reduction of the debt-to-GDP ratio, is frequently not achieved due to negative effects of these measures on GDP. Fiscal austerity may occasionally be unavoidable, but even in these cases, deliberate measure-taking is required to prevent the increase in unemployment and income inequality, as witnessed after the global financial crisis.

Keywords*: Debt crisis, European countries, fiscal austerity, income inequality*

1. Introduction

In order to balance the state budget, the practice of fiscal austerity is defined by a decrease in government spending and/or an increase in taxation, which typically takes place during times of crisis. Regarding the application of fiscal austerity measures, their efficacy, as well as their repercussions, there is no agreement in the political and scientific community. Lowering state spending and raising tax revenues is assumed to reduce the budget deficit, restore investor confidence, and promote long-term economic recovery (Anderson, 2010). However, fiscal austerity measures have historically not reduced the budget deficit due to decreased output and tax revenue (Okeke et al., 2021).

The main goal of this paper is to analyse the empirical impact of fiscal austerity on economic growth, while also considering its effects on some variables that affect the social status of citizens. It is believed that tightening the state's budget primarily affects the lowest socioeconomic groups, thereby widening the income gap (Campoy-Muñoz et al., 2022). Government spending cuts cause the labour income share to contract, which causes income to be redistributed away from the working class. Additionally, since these measures are typically implemented during recessions, they also affect unemployment, directly through the loss of jobs in the public sector and indirectly through a decline in private-sector employment. Upon examining the historical fiscal austerity measures, it can be inferred that the primary focus of these measures was on increasing regressive taxation and cutting social expenditures, which served as the impetus for the rise in both unemployment and inequality. Mattei (2022) views fiscal austerity as a logical political agenda to protect capitalism since it increases worker precariousness and dependence on the market.

The example of the 2008 global financial crisis often illustrates the failure of fiscal austerity policies. Due to the crisis, most EU members implemented fiscal austerity measures to balance their budgets and prevent sovereign debt default (Ray et al., 2020). However, compared to other EU members, Greece had to implement far more austerity measures during the crisis. This was due to an austerity programme imposed by the EU and the European Central Bank, which aimed to bring Greece's finances under control in exchange for bailouts (Economides et al., 2021). The programme included a tax reform, a pension reform, and pay cuts in the public sector. In relation to the tax reform, the administration was expected to present a fresh strategy aimed at enhancing revenue collection, curbing capital flight, and combating tax evasion. According to the IMF (2013), it was observed that no substantial progress has been made towards curbing tax evasion, and both the wealthy and the self-employed are still not paying their required taxes. To make matters worse, the value-added tax (VAT), which is a more regressive way of collecting taxes, was raised by 10 percentage points, putting a burden on low-income households. This programme led to a reduction in public spending and an increase in taxes, which affected the public employees of Greece. Greece's deficit has significantly reduced since then. However, the nation's austerity programme in 2010 had only sporadic positive effects on the economy. This is because the financial situation was not significantly impacted by the austerity measures, as the nation had already experienced a deficiency in aggregate demand. To escape the austerity trap, it is crucial to increase exports to potentially offset the decline in aggregate demand (Stiglitz, 2014). However, Greece's export performance has been poor, and the country has been carrying a massive current account deficit, which exceeds 10% of the GDP (Pagoulatos, 2018). The Greek economy shrank by 25% in 7 years, with unemployment above 20%. Austerity failed to achieve sustainable debt, and a 2015 referendum rejected the second bailout due to its negative impact on the public (Ifanti et al., 2013).

In this paper, the synthetic control method is employed to examine the effects of fiscal austerity on real GDP per capita, unemployment, and inequality. This approach allows for testing the causal relationship between the variables. Considering the severity of the measures, the effects of fiscal austerity are investigated using Greece as an example. Since the fiscal austerity measures started to be implemented in 2010, when the first Economic Adjustment Programme (EAP) was approved, 2010 is considered the year of treatment. Changes in real GDP per capita, unemployment, and income inequality are examined from 1995 to 2018.

The remainder of the paper is structured as follows. Section 2 provides a literature review on the relationship between fiscal consolidation and economic and social variables of interest. Section 3 brings the methodology and the used data. The results are presented and discussed in Section 4. Finally, the main conclusions of the research are provided in Section 5.

2. Literature review

The 2008 crisis reignited the long-running controversy about the efficiency of fiscal multipliers and the impact of fiscal austerity on GDP, and scholars still disagree on that matter. They divide between those who consider that fiscal austerity harms the economy (Blanchard & Cottarelli, 2010; Blanchard & Leigh, 2013; Alexiou & Nellis, 2016; Brinca et al., 2020) and those who claim that fiscal austerity can have expansionary effects (Giavazzi & Pagano,

1990; Alesina & Ardagna, 2010; Mirdala, 2016; Alesina et al., 2015; 2019), mainly when based on spending cuts.

A variety of empirical research supports the neoclassical theory. Giavazzi and Pagano (1990) demonstrate a positive correlation between fiscal consolidation and an increase in private consumption. Alesina and Perotti (1997) discovered that there is occasionally a connection between fiscal consolidation and rapid output growth, mainly when the consolidation is carried out by reducing government spending instead of raising taxes. Further research, such as Alesina and Ardagna's (2010) paper, has confirmed these findings by analysing larger samples of countries and years.

The economic impacts of budgetary reforms in OECD economies are examined by Alesina et al. (2015). They show that fiscal consolidation based on spending has relatively low economic costs, whereas consolidation based on taxes has far higher costs. Jordà and Taylor (2015) use a counterfactual analysis of the effects of the UK government's transition to austerity measures, which took place in 2010. They demonstrate that austerity is always detrimental to growth, particularly in times of recession. In a downturn, fiscal consolidation amounting to 1% of GDP measures a four per cent loss of real GDP over five years, as opposed to just one per cent in a boom. Similar conclusions are drawn by House et al. (2019), who claim that counterfactuals demonstrate that removing austerity would have significantly decreased output losses in Europe. Moreover, in the wake of the crisis, several European nations experienced an increase in debt-to-GDP ratios due to endogenous decreases in GDP and tax income caused by contractionary austerity shocks.

Using the synthetic control method, Rayl (2020) shows that GDP per capita in Greece, Spain, and Italy would have been higher if fiscal consolidation measures had not been implemented. Similarly, using the same method, Revuelta (2021) examines the effects of the EAPs in Greece and finds that their implementation has had a negative impact of 35.3% of the country's GDP per capita.

Within the framework of fiscal austerity, most research focuses on how the policy mix and economic output interact. However, only a few have examined how these budgetary adjustments affect unemployment and income distribution. In contrast to neoclassical perspectives, it asserts that austerity causes output loss and raises long-term unemployment, which sets off an economic hysteresis effect (Alexiou & Nellis, 2016). The IMF (2014) examined specific cases where policy actions were meant to lower budget deficits. They show that reducing the budget deficit had increased unemployment and caused domestic demand to contract for both kinds of adjustment instruments, tax-based instruments and spending-based instruments. Using counterfactual simulations, Lama and Medina (2019) demonstrate that improving the fiscal balance might come at the cost of a higher unemployment rate.

The relation between unemployment and the impact of fiscal austerity on income inequality is quite strong. Woo et al. (2013) point out that the unemployment channel accounts for roughly 15-20% of increases in income inequality. Research demonstrates that spending-based adjustments have a significant and negative impact on the labour share of national income and increase income disparities, mainly because they cause long-lasting unemployment (Okeke et al., 2021), whereby the impact on the labour's share of national income is more significant than that on capital income (Ball et al., 2013).

3. Methodology and data

The synthetic control method (SCM) enables the evaluation of the effect of an intervention that affects one unit, which can be, for example, a region or a country. The synthetic control method is a generalisation of the difference-in-differences model that shows what would have happened if an intervention or policy had not occurred. Essentially, the SCM proposes that a group of units is a more reliable "imitation" for the unit receiving the treatment than any unit by itself. The SCM is an innovative way to detect causality, presented first in the work of Abadie and Gardeazabal (2003). It has several benefits: 1) prevents endogeneity bias; 2) useful for small samples, and 3) produces more accurate results compared to standard regression (Abadie et al., 2010; Abadie, 2020).

The SCM searches for a set of control unit weights w_j such that:

$$Y_{1,\,t<T_0} = \sum_{j=2}^{J+1} w_j * Y_{j,\,t<T_0}. \tag{1}$$

There are outcome variables (real GDP per capita, unemployment rate and GINI) for a treated unit (Greece) and J control units (27 European coun-

tries). The outcome variable in country i at time t is $Y_{i,t}$; $i = 1$ is the treated unit; $i = 2,3, ... , J + 1$ are the control units; T_0 refers to the treatment year (fiscal austerity measures in 2010); the time periods where $t < T_0$ are 'pre-treatment' (1995-2009), and $t \geq T_0$ is 'post-treatment' (2010-2018).

During the pre-treatment period, the treated unit is equivalent to the weighted total of the control units. But since it is often not possible to obtain a precise match, the following equation is used instead (Abadie and L'hour, 2021):

$$Y_{1,\, t<T_0} \approx \sum_{j=2}^{J+1} w_j * Y_{j,\, t<T_0}, \qquad (2)$$

such that the mean squared error prior to treatment is minimised:

$$w_j = \min_{w} \sqrt{\frac{1}{T_0}\sum_{t=1}^{T_0}\left(Y_{1,t} - \sum_{j=2}^{J+1} w_j Y_{j,t}\right)^2}. \qquad (3)$$

By averaging pre-intervention results of chosen donor states, the SCM constructs a "synthetic" counterfactual region. A pool of possible candidates is used to pick the donor countries that come together to generate synthetic control. The choice of donor countries and weights is determined by predictor variables that impact the result.[1] The resulting synthetic is a control for the affected region after policy adoption and closely resembles the outcome of the affected region before policy enactment.

Using the w_j weights, the causal effect is estimated such that:

$$Y_{1,t\geq T_0} = Y_{1,\, t\geq T_0}^N + \alpha_{1,t}. \qquad (4)$$

The fundamental problem of causal inference is that $Y_{1,\, t\geq T_n}^N$ cannot be observed, thus the weighted control units are used instead. The difference between the treated time series and the weighted average of control units which represents the treatment's effect, also known as the causal effect, can be written as (Abadie et al., 2010):

$$\alpha_{1,t} = Y_{1,t\geq T_0} - Y_{1,\, t\geq T_0}^N. \qquad (5)$$

For the approach to be used effectively, three assumptions must be valid. First, for every year during the pre-treatment period that was utilised to generate synthetic control, the policy change only affects the treated country. Second, the policy change has an effect only when it is implemented. In other words, the treatment is not anticipated and does not cause an effect before the year of implementation. Third, a fixed combination of donor states can approximately represent the counterfactual outcome for the treated country. This means that in order to prevent any potential interpolation bias, the values of the variables used to construct the weights must be comparable between the donor pool countries and the affected country. The values of the treated country cannot be outside any linear combination of the values from the donor pool countries. The assumptions of the model are explained in the following paragraphs.

Given that in this paper we want to examine the impact of fiscal austerity on real GDP per capita, unemployment and inequality, it is necessary to construct three models. In order to be able to observe the causality in all three models, to begin with, it is necessary to select a country that has introduced fiscal austerity measures to a greater extent than the other countries that make up the donor pool[2] (McClelland and Gault, 2017). In the selection of the treated unit, the methodology according to Rayl (2020) was employed, where the degree of strength of fiscal austerity is determined from IMF data in such a way that the average increase and decrease of the budget deficit in the period from 2010-2013, but also the percentage year-to-year increase in the government structural balance, are observed (presented in Table 1). It can be concluded that Greece and Spain experienced heavy austerity measures, but Greece faced these measures to a much larger extent.

---

1 Thus, different predictor variables imply different selection of weights and countries in the "synthetic" counterfactual country.
2 The donor pool is constructed out of potential candidates that combine to form synthetic control.

Table 1 Evaluating the magnitude of fiscal austerity in European countries, 2010-2013

Type of Austerity	Definition	Countries
No Austerity	> 0 avg, no more than 1% increase in YTY GSB[3] in any year during 2010 - 2013	Denmark, Estonia, Finland, Norway, Sweden, Switzerland
Very Light Austerity	> 0.5, no more than 2% increase in YTY GSB in any year during 2010 - 2013	Austria, Belgium, Germany, Luxembourg, Malta
Light Austerity	> 0.85, no more than 2.5% increase in YTY GSB in any year during 2010 - 2013	Croatia, France, Netherlands, Poland, Slovenia
Moderate Austerity	> 0.85 avg GSB (% GDP)	Czech Republic, Island, Ireland, Italy, Latvia, Lithuania, Portugal, Romania, Slovak Republic, United Kingdom
Heavy Austerity (Treatment)	>2 avg, greater than 4% increase in YTY GSB in a year during 2010 - 2013	Greece, Spain

Source: International Monetary Fund, World Economic Outlook Database, October 2019
General Government Structural Balance

This paper uses Greece as the treated unit, and the year 2010 as the year of the intervention, which is related to the introduction of the first Economic Adjustment Programme[4]. The idea of the synthetic control method is to construct a "synthetic" Greece based on averaging the predictor variables of the countries from the donor pool. Therefore, the first step is the selection of predictor variables, i.e. variables that influence outcomes in countries from the donor pool both before and after treatment.

Table 2 presents the selected predictor variables, as well as the variables of interest. Different predictor variables are used in the three models. The lagged values of the outcome variables from 1998, 2002, and 2009 are included, as the lagged outcome variable is one potentially significant predictor of the outcome variable. It addresses the problem of omitting important predictor effects since it includes the impacts of all predictor variables, regardless of whether the analyst collects them. Indeed, according to Athey and Imbens (2006), it is customary to include the lagged outcome variable for a few pre-treatment years, since other predictor variables are rarely significant. In implementing the synthetic control method, it is necessary to define a pre-treatment period, i.e. a period before the intervention over which the predictors will be averaged and a post-treatment period, i.e. a period in which the effect of the intervention will be examined. In this

paper, the pre-treatment period refers to the period from 1995 to 2009, while the post-treatment period refers to the period from 2010 to 2018. The next step is the choice of countries for the so-called donor pool. In order to acquire values for the donor pool countries that are comparable to those for Greece, EU countries are selected in the donor pool. Ideally, units that have not been the subject of treatment should be included in the donor pool. Given that all EU countries introduced certain fiscal austerity measures during the financial crisis, the goal is to exclude those countries where these measures were more pronounced. Accordingly, all EU countries are included in the donor pool, with the exception of Spain. Since the assumption that the treatment was present only in Greece is not entirely satisfied, this could cause the effect to be relatively underestimated in Greece, which is not as concerning as an overestimation of the effect, but it is still important to consider this when analysing the results.

Furthermore, each country is assigned a certain share in the "synthetic" control group with the aim that it closely matches Greece in the pre-treatment period. In the period of intervention, an alternative scenario can be seen, i.e. what would have happened if fiscal austerity measures had not been introduced. Thus, the causal effect can be determined as the difference between Greece and "synthetic" Greece.

3   Year-to-year change in the general government structural balance.

4   See Revuelta (2021).

Table 2 Description of variables

Outcome variables	
gdppcap (Model 1, M1)	GDP per capita, based on purchasing power parity (PPP) (in constant 2011 international dollars)
unem (Model 2, M2)	Unemployment rate
ginim (Model 3, M3)	The Gini index on market income—sometimes referred to as a pre-tax Gini index—measures inequality in income without considering the effect of taxes and social spending already in place in a country. It is shown on a scale from 0 to 100%.
Predictor variables	
popg (M1)	Population growth (annual %)
inv (M3)	Net investment in government nonfinancial assets (% GDP) includes fixed assets, inventories, valuables, and non-produced assets. Nonfinancial assets are stores of value and provide benefits either through their use in the production of goods and services or in the form of property income and holding gains. Net investment in nonfinancial assets also includes consumption of fixed capital.
gdppcap (M2, M3)	GDP per capita based on purchasing power parity (PPP)
terenr (M2)	School enrolment, tertiary (% gross)
ind (M1, M2)	Industry, value added (% of GDP)
agr (M1, M2)	Agriculture, value added (% of GDP)
debtgdp (M1, M3)	General government gross debt (% of GDP)
trade (M2)	Trade (% of GDP); the sum of exports and imports of goods and services measured as a share of gross domestic product
unem (M3)	Unemployment rate

Note: All variables are gathered from the World Bank Database; except for the variable Ginim, which is gathered from the Standardized World Income Inequality Database.
Source: World Bank Database, Standardized World Income Inequality Database

To ensure that the findings are not the result of pure chance, placebo tests are employed (in-space and in-time placebo). In order to conduct the in-space placebo test, the synthetic control approach is applied to untreated countries to quantify a degree to which the impact size in untreated countries deviates from the effect found for treated countries. There is reason to question the significance of the estimations for the treated unit if impacts of comparable magnitudes are frequently detected. It can be claimed that there is strong evidence of a significant impact of fiscal austerity on Greece's economic and social variables if placebo testing reveals that the difference between the synthetic outcome and the actual outcome is unusually large for Greece in comparison to other nations. Throughout the course of the treatment year, in-time placebo tests are also conducted. To further verify that the outcomes are not the product of pure chance, the intervention should be reassigned to years when there was none (Chen & Yan, 2023).

4. Results and discussion

Estimating the impact of fiscal austerity in Greece requires determining the optimal counterfactual.

Synthetic control weights and the optimal counterfactual composition have been optimised using the process outlined in the preceding section. The STATA18 synth2 package was used to perform the estimations.

Three models were constructed. The first model refers to the impact of fiscal austerity on real GDP per capita. Predictor variables used in the first model are popg, debtgdp, agr, ind and lagged outcome variables for 1998, 2002 and 2009. In constructing the counterfactual, the following shares were allocated to the countries: Hungary 47.1%, Iceland 20.2%, Latvia 11.3%, Belgium 11%, Luxembourg 6.5%, and Lithuania 4%. The result of our baseline specifications is given in Figure 1. In the pre-treatment period, real GDP per capita in "synthetic" Greece adequately captures the dynamics of real GDP per capita in Greece (R-squared = 0.98). This allows

for a more certain interpretation of the outcomes produced by the synthetic control method. The difference between the real GDP per capita of Greece and "synthetic" Greece between 2010 and 2018 was used to quantify the effect of the introduction of fiscal austerity measures. Although the EAP measures were introduced in 2010, the effect was already visible in 2009, which can be explained by the possible anticipation of such measures and the slowdown of the economy. The method showed a negative effect of fiscal austerity measures on real GDP per capita in Greece. The red (dashed) line indicates what real GDP per capita in Greece would have been if fiscal austerity measures had not been introduced. As can be seen, real GDP per capita would then be significantly higher, on average (through the post-treatment period) 6,916.96 international dollars. This finding supports the heterodox view of fiscal austerity presented in Section 2.

Figure 1 The effect of EAP on real GDP per capita (gdppcap) in Greece
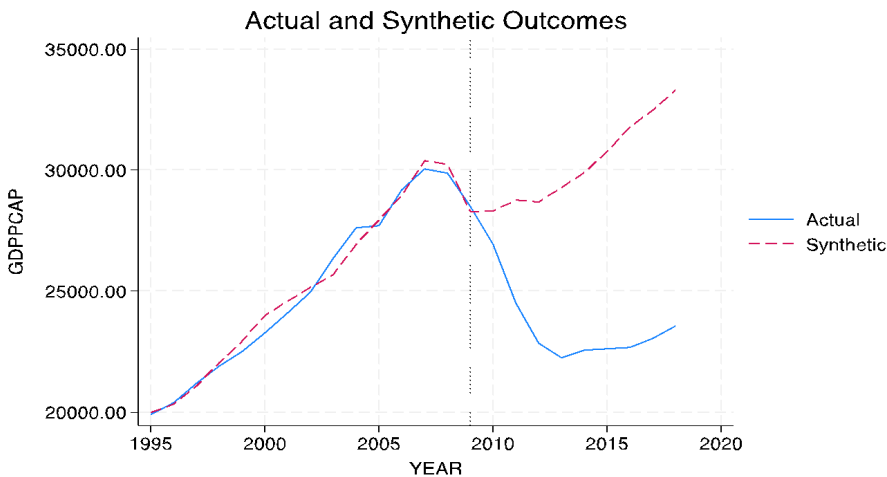
Source: Authors' calculations

The second model shows the impact of fiscal austerity on the unemployment rate (see Figure 2). Predictor variables included in the model are gdppcap, terenr, agr, ind, trade and lagged values of the dependent variable (same as in the first model). In this model, "synthetic" Greece is constructed from France (41.3%), Portugal (22.8%), Poland (19.2%), and Italy (16.6%). In this model, the influence of the effect can also be seen since 2009; however, the synthetic version does not match the actual values for Greece (prior to the treatment) as pre-

cisely as in Model 1. The validity of the model will be demonstrated by the placebo tests conducted in the following stage. Figure 2 shows that Greece has significantly higher unemployment rates compared to "synthetic" Greece, which is in line with the scientific literature. The average treatment effect, estimated in the post-treatment period, is 12 percentage points. The effect itself is most pronounced in 2013 (15.68 percentage points) and thereafter it starts to decrease gradually.

Figure 2 The effect of EAP on the unemployment rate (unem) in Greece

Actual and Synthetic Outcomes
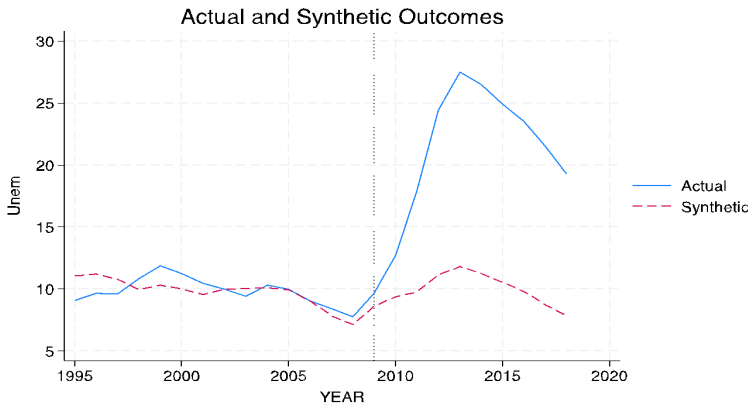
Source: Authors' calculations

The impact of the introduction of the Economic Adjustment Programme on income inequality (measured by the Gini index before taxes and transfers) is examined by the third model, and it is shown in Figure 3. Predictor variables included in this model are unem, gdppcap, debtgdp and inv, and lagged values of the dependent variable. "Synthetic" Greece is constructed from Sweden (47.3%), Austria (27.3%), Portugal (23.6%), and Ireland (1.6%), and it closely resembles the actual values for Greece prior to the treatment (R-Squared = 0.95). The impact of fiscal austerity on income inequality varies in comparison to the impact on real GDP per capita and the unemployment rate. Namely, at the beginning of the post-treatment period, income inequality in Greece is higher than in "synthetic" Greece, and the peak was reached in 2013, when the Gini index in

Greece was 1.357 percentage points higher than it would have been if Greece had not started the EAP. However, after 2017, the effect turns negative, i.e. the Gini index in Greece is 1.1 percentage points lower in 2017, and 2.02 percentage points lower in 2018 than it would have been if Greece had not introduced fiscal austerity measures. Thus, from the aspect of income inequality, fiscal austerity initially increased income inequality, and later it resulted in the reduction of income disparities. Thus, the level of inequality in Greece was lower compared to that of "synthetic" Greece. Fiscal austerity is expected to increase income inequality, while the opposite effect can be explained by the reduction of public debt and the indirect effect it has on income distribution (see Hager, 2016).

Figure 3 The effect of EAP on income inequality (ginim) in Greece


Actual and Synthetic Outcomes
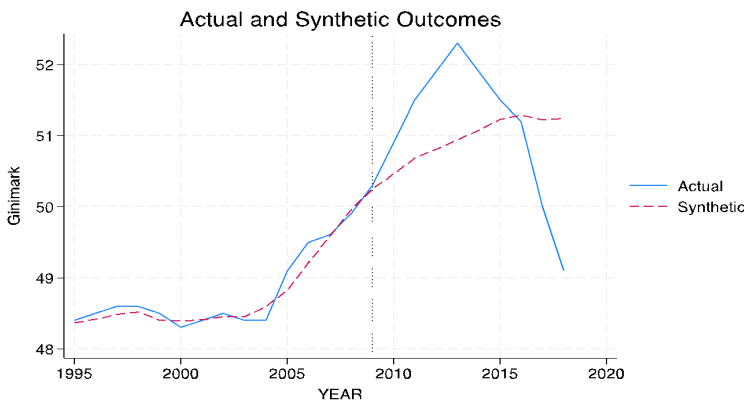
Source: Authors' calculations

To test the reliability of the synthetic control method results, two types of placebo tests were conducted – in-space placebo and in-time placebo. The results of these tests are displayed in the paragraphs that follow.

First, an in-space placebo test was performed to analyse the significance of the estimates. The test requires that all of the donor countries are subject to the 2010 intervention.[5] Figures 4, 5 and 6 show the results of the placebo test for the three models, respectively, by graphing the distribution of placebo effects. The mean squared prediction error (MSPE) and the p-value derived from the placebo test provide detailed information on the robustness of the model, as can be seen in Figure 4. In the first model, the expected impact for Greece differed markedly from the estimates of the placebo tests for the

countries in the donor pool. The red line shows the impact on Greece, whereas the grey lines illustrate the effects when the treatment is applied to different donor pool countries. Greece's influence can be deemed significant if the red line falls below the grey lines during the post-treatment period. This is because, in comparison to the placebo countries, Greece encounters the greatest negative effect from the treatment. Since that is the case in Figure 4, the effect was caused by the EAP in Greece, as opposed to a common shock that the EU suffered as a result of the financial crisis and a deep recession. The results indicate that Greece has the largest MSPE ratio among all donor countries, with a general p-value of $1/16^{6} = 0.0625$ (significant at the 10% level). Additionally, the pointwise p-values (two-sided and left-sided) present that for most post-treatment years the effects are significant at the 10% level.

Figure 4 Robustness of results for Model 1 – placebo gaps in control countries
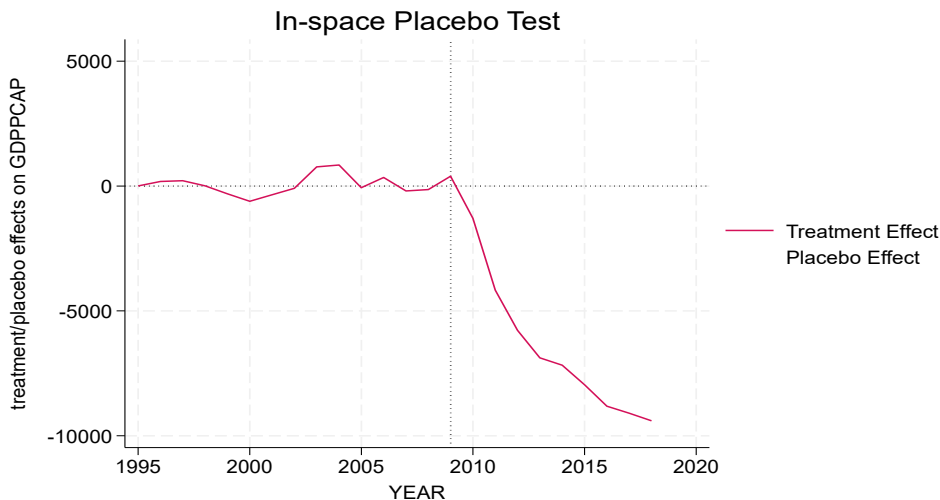
Note: The red line represents Greece, while the grey lines show other countries from the donor pool with a pre-treatment MSPE two times larger than the treated unit.
Source: Authors' calculations

In the second model, the placebo test yields an overall p-value of 0.05, while the pointwise p-values (either two-sided or right-sided) are significant at 5% for all post-treatment years. Figure 5 illustrates how Greece's effect on unemployment differs significantly from other countries included in the placebo test. The red line in this figure indicates that

the effect on Greece is the most significant as the red line is above the grey lines. In other words, Greece's unemployment rate increased more than that of any other donor country as a result of the fiscal austerity measures. Therefore, we can validate the accuracy of this model.

5   This analysis excludes those with a pre-treatment MSPE two times larger than that of the treated unit (option (unit cutoff(2)).

6   Number 16 is calculated by extracting the number of units with a pre-treatment MSPE two times larger than that of the treated unit (which are 8 units) from the total number of donor states (24).

Figure 5 Robustness of results for Model 2 – placebo gaps in control countries

Note: The red line represents Greece, while the grey lines show other countries from the donor pool with a pre-treatment MSPE two times larger than the treated unit.
Source: Authors' calculations

Figure 6 presents the placebo test results for the third model. The figure shows that the outcome for Greece does not significantly differ from the estimates of the placebo tests for the donor pool nations. This can also be confirmed by the overall p-value of 0.5[7], as well as by the pointwise p-values that are not significant at the common levels of significance. The outcome shows that the 2010 treatment impact is not specific to Greece; hence, it is not possible to link it to the implementation of EAPs with certainty.

Figure 6 Robustness of results for Model 3 – placebo gaps in control countries

Note: The green line represents Greece, while the grey lines show other countries from the donor pool with a pre-treatment MSPE two times larger than the treated unit.
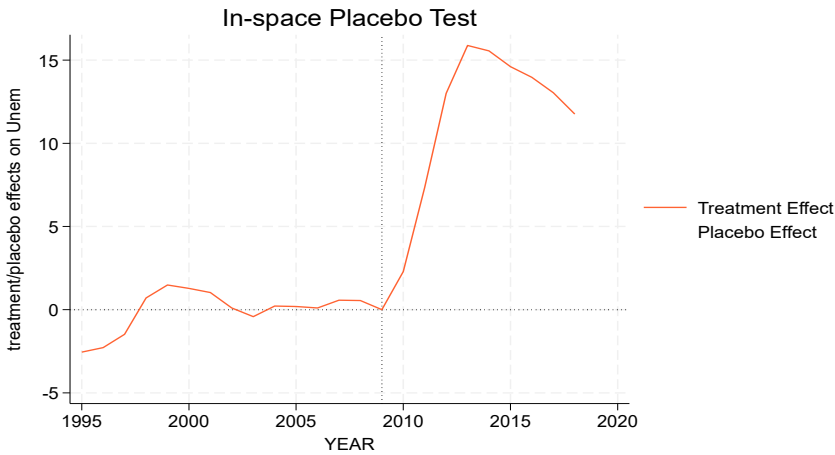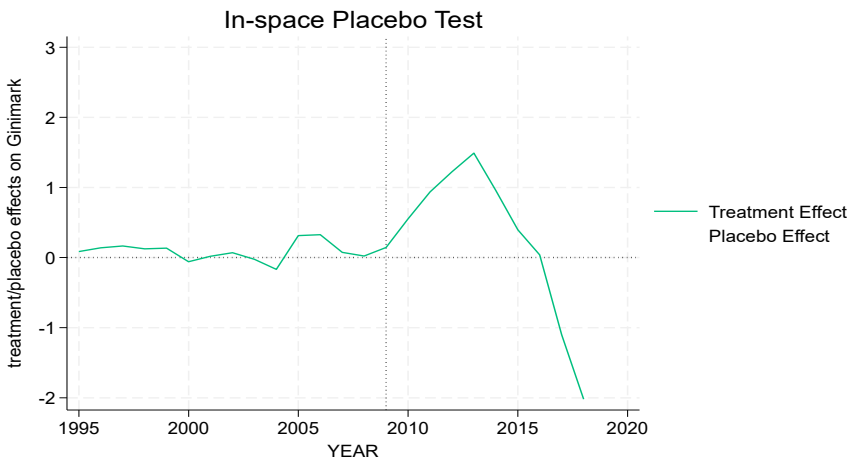*Source: Authors' calculations*

---

7   There are a total of 22 units with a pre-treatment MSPE two times larger than the treated unit; therefore, the p-value is calculated as ½.

Secondly, to further test the robustness of the results, an in-time placebo test was conducted. We use the synthetic control approach to establish an alternative intervention year for the treatment; rather than 2010, we consider the treatment to begin in 2005, five years prior to the first actual implementation of EAPs.

Figure 7 Robustness of results for Model 1 – in-time placebo test (treatment year = 2005)
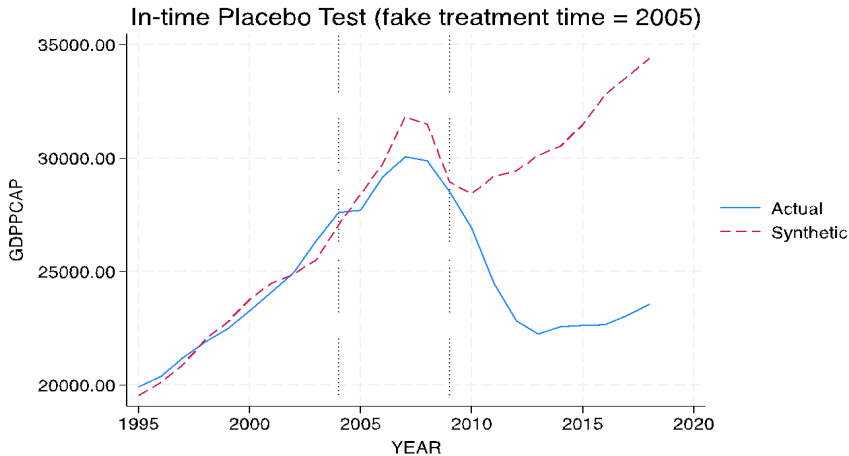
In the first (see Figure 7) and the second model (see Figure 8), the EAP effect became apparent in 2010. Thus, the results of the first two models are robust. Figures 7 and 8 illustrate that, even when a different treatment year is selected, the impact of the treatment is most evident from 2010 onward. This indicates that there is no bias in these models regarding the treatment year. Therefore, the effect shown by the synthetic control method can be considered a consequence of the implementation of EAPs in 2010.

Figure 8 Robustness of results for Model 2 – in-time placebo test (treatment year = 2005)
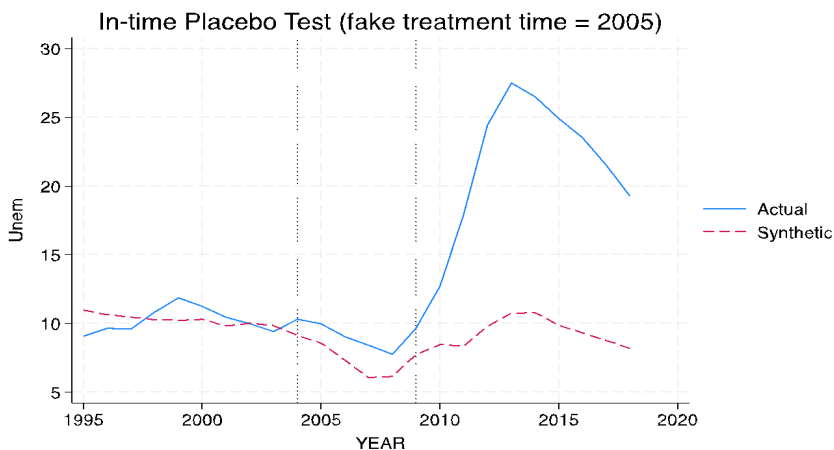
However, the in-time placebo test result of Model 3 indicates that the model is not robust, given that the effect appears to have begun in 2004 (see Figure 9).

As a result, the effect determined by the synthetic approach cannot solely be attributed to Greece's fiscal austerity measures.

Figure 9 Robustness of results for Model 3 – in-time placebo test (treatment year = 2005)

Source: Authors' calculations

The in-time placebo test reassigns the treatment to start in a year during the pre-treatment period. If the placebo estimate significantly differs from the actual pre-treatment, the robustness of the model is questionable. While Models 1 and 2 can be considered robust, the third model is biased due to the discrepancy between the actual treatment year and the year when the effect occurred in the in-time placebo test.

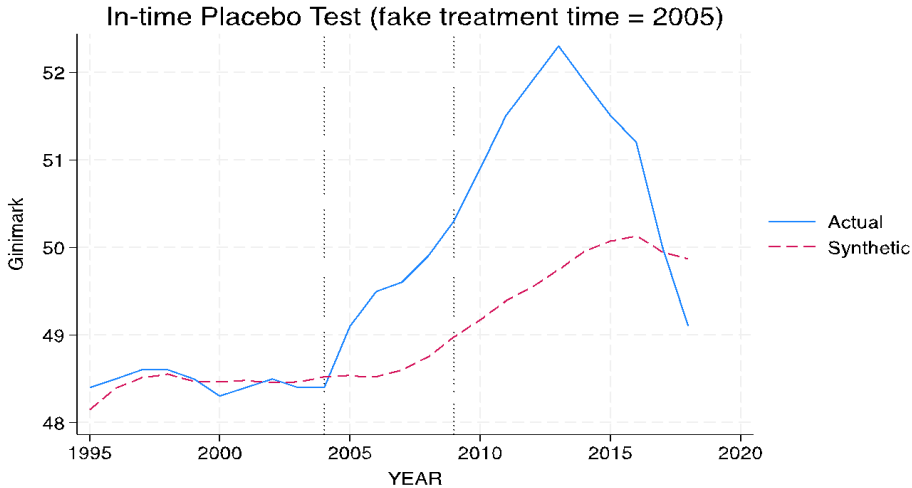Furthermore, reasons for decreasing GDP per capita in Greece are external (a financial crisis), but also internal, i.e. orientation towards tourism, which is related to structural weaknesses of the economy and relative incompatibility with the structure of developed countries, low resilience, low productivity and a low level of competitiveness. The methodology used in this paper enables the identification of the causal effect of "fiscal austerity", given that the donor pool includes European countries which were also affected by the financial crisis. The main cause of the decline was observed uniquely in all countries, whereby internal factors are to a greater or lesser extent represented at least in one or more EU member states used in the donor pool. For instance, low productivity and a low level of competitiveness were also observed in Portugal, Croatia,

Romania, etc. Therefore, with this method, the effects of the introduction of EAP measures and their causality with respect to GDP, unemployment and inequality can be determined quite precisely, abstracting to a greater extent the influence of other factors relevant to the phenomenon itself.

5. Conclusion

In this paper, the effect of fiscal austerity measures introduced in 2010 in Greece on the variables of real GDP per capita, unemployment rate and income inequality, is evaluated using the synthetic control method. The causal effect is estimated by creating a counterfactual from a group of European countries, which depicts Greece as closely as possible in the years prior to the intervention. In the post-treatment period, the difference between the synthetically created Greece and the actual Greece can be observed, which shows the effect of the intervention. The period in which the movement of the variables is observed is from 1995 to 2018.

The paper shows that the fiscal austerity measures introduced during the global financial crisis in Greece had a negative impact on the movement of real GDP per capita, which is 42.7% lower in Greece

in 2018 in comparison to the real GDP per capita of the "synthetic" Greece. Precisely, in the post-intervention period (from 2010 to 2018), real GDP per capita is on average 6,916.96 international dollars lower than it would have been if fiscal austerity measures had not been introduced. As for the unemployment rate, the average treatment effect after the intervention is 12 percentage points. The unemployment rate in 2018 in Greece is 61.54% higher when compared to the "synthetic" Greece. It should be emphasised that since 2013 the effect has been decreasing.

The findings are not consistent when it comes to how fiscal austerity affects income inequality. First, the implementation of EAPs causes Greece's Gini coefficient to rise relative to the "synthetic Greece"; subsequently, the Gini coefficient decreases, and since the beginning of 2016, Greece's inequality has decreased relative to what it would have been in the absence of fiscal austerity measures. Still, the results may only be partially interpreted since the robustness of the third model was not verified by the in-time and the in-space placebo test.

It can be concluded that fiscal austerity in Greece reduced real GDP per capita and caused an increase in the unemployment rate. Because fiscal austerity measures were linked to the global financial crisis, this work is limited in that it is impossible to determine whether fiscal austerity is always harmful (the heterodox perspective) or only problematic when implemented during recessions (the Keynesian view). This research, however, confirmed the statements of those who criticised the strict austerity policy in Greece. The empirical results of this paper are in line with the findings of Jordà and Taylor (2015) and House et al. (2019) (see Section 2), which suggest that fiscal consolidation is detrimental to economic growth. Considering the effect on unemployment, this paper is consistent with Lama and Medina (2019), who also examine a rise in the unemployment rate due to fiscal austerity. Since the aforementioned papers examine the effects of fiscal austerity in different countries, this enables a comprehensive view of the consequences of such policies. Rayl (2020), Revuelta (2021) and Alogoskoufis (2019), who also concentrate on the case of Greece, likewise conclude that the 2010 strategy in Greece came at a significant cost in terms of a reduction in output and employment. The strong fiscal adjustments required from Greece had depressing effects by sinking domestic demand and thereby lowering real GDP per capita while exaggerating unemployment. As a consequence, Greece's economy was unable to grow or achieve fiscal sustainability, which led to a vicious cycle.

The obtained results also bring policy implications. In all instances, the findings of our research indicate that the potential distributional effects of fiscal austerity should be weighed against its benefits. Governments are frequently able to introduce tax increases or spending reductions so that the overall impact on distribution is minimised. It is crucial to pay attention to how fiscal policy affects inequality, especially now that it is rising and could be a cause of social unrest. When evaluating the causal effects of comprehensive policies, the synthetic control method has the advantage of being able to resolve the endogeneity issue. Its limitation, however, is that it depends on several assumptions, some of which are challenging to verify. One example of these assumptions in this paper is that the policy was not anticipated and only implemented in the observed country. Future research could consider different methods in addition to the synthetic approach to further evaluate the consequences of fiscal austerity.

REFERENCES

1. Abadie, A. (2020). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, *59*(2), 391-245. https://www.doi.org/10.1257/jel.20191450

2. Abadie, A. & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *The American Economic Review*, *93*(1), 113-132. https://doi.org/10.1257/000282803321455188

3. Abadie, A. & L'hour, J. (2021). A Penalized Synthetic Control Estimator for Disaggregated Data. *Journal of the American Statistical Association*, *116*(536), 1817-1834. https://doi.org/10.1080/01621459.2021.1971535

4. Abadie, A., Diamond, A. & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, *105*(490), 493-505. https://doi.org/10.1198/jasa.2009.ap08746

5. Alesina, A. & Ardagna, S. (2010). Large Changes in Fiscal Policy: Taxes versus Spending. *Tax Policy and Economics*, *24*(1), 35-68. https://doi.org/10.1086/649828

6. Alesina, A. & Perotti, R. (1997). Fiscal Adjustments in OECD Countries: Composition and Macroeconomic Effects. *IMF Economic Review*, *44*(2), 210-248. https://doi.org/10.2307/3867543

7. Alesina, A., Favero, C. & Giavazzi, F. (2015). The output effect of fiscal consolidation plans. *Journal of International Economics*, *96*(S1), S19-S42. https://doi.org/10.1016/j.jinteco.2014.11.003

8. Alesina, A., Favero, C. & Giavazzi, F. (2019). Effects of austerity: expenditure- and tax- based approaches. *Journal of Economic Perspectives*, *33*(2), 141-162. https://doi.org/10.1257/jep.33.2.141

9. Alexiou, C. & Nellis, J. G. (2016). A post-mortem of austerity: The Greek experience. *Economic Issues*, *21*(2), 1-32.

10. Alogoskoufis, G. (2019). *Greece and the Euro: A Mundellian Tragedy* (GreeSE Paper No. 136). London: London School of Economics and Political Science.

11. Anderson, B. (2010). Restoring Fiscal Sustainability: Lessons for the Public Sector. *Public Policy Review*, *6*(4), 611-624.

12. Athey, S. & Imbens, G. W. (2006). Identification and Inference in Nonlinear Difference-in-Difference Models. *Econometrica*, *74*(2), 431-497. https://doi.org/10.1111/j.1468-0262.2006.00668.x

13. Ball, L. M., Furceri, D., Leigh, D. & Loungani, P. (2013). *The Distributional Effects of Fiscal Consolidation* (IMF Working Paper No. 13/151). Washington, DC: International Monetary Fund. https://doi.org/10.5089/9781475551945.001

14. Blanchard, O. & Cottarelli, C. (2010). *Ten commandments for fiscal adjustment in advanced economies.* VoxEU 28. https://voxeu.org/article/ten-commandments-fiscal-adjustment-advanced-economies

15. Blanchard, O. J. & Leigh, D. (2013). *Growth forecast errors and fiscal multipliers* (NBER Working Paper No. w18779). Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w18779

16. Brinca, P., Ferreira, M., Franco, F., Holter, H. & Malafry, L. (2020). Fiscal consolidation programs and income inequality. *International Economic Review*, *62*(1), 405-460. https://doi.org/10.1111/iere.12482

17. Campoy-Muñoz, P., Cardenete, M. A., De Miguel-Vélez, F. J. & Pérez-Mayo, J. (2022). How does fiscal austerity impact on poverty and inequality? The Spanish case. *Economia Politica: Journal of Analytical and Institutional Economics*, *39*(3), 715-737. https://doi.org/10.1007/s40888-021-00253-5

18. Chen, Q. & Yan, G. (2023). A mixed placebo test for synthetic control method. *Economics Letters*, *224*, 111004. https://doi.org/10.1016/j.econlet.2023.111004

19. Economides, G., Papageorgiou, D. & Philippopoulos, A. (2021). Austerity, Assistance and Institutions: Lessons from the Greek Sovereign Debt Crisis. *Open Economies Review*, *32*, 435-478. https://doi.org/10.1007/s11079-020-09613-3

20. Giavazzi, F. & Pagano, M. (1990). *Can severe fiscal contractions be expansionary? Tales of two small European countries* (NBER Working Paper No. 3372). Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w3372

21. Hager, S. B. (2016). *Public Debt, Inequality, and Power: The Making of a Modern Debt State.* University of California Press. https://doi.org/10.1525/luminos.14

22. House, C. L., Proebsting, C. & Tesar, L. L. (2019). Austerity in the aftermath of the Great Recession. *Journal of Monetary Economics*, *115*, 37-63. https://doi.org/10.1016/j.jmoneco.2019.05.004

23. Ifanti, A. A., Argyriou, A. A., Kalofonou F. H. & Kalofonos, H. P. (2013). Financial crisis and austerity measures in Greece: Their impact on health promotion policies and public health care. *Health Policy*, *113*(1-2), 8-12. https://doi.org/10.1016/j.healthpol.2013.05.017

24. IMF (2013). *Greece: 2013 Article IV Consultation Concluding Statement of the IMF Mission.* http://www.imf.org/external/np/ms/2013/050313.htm

25. IMF (2014). *Fiscal Policy and Income Inequality* (IMF Policy Paper No. 40). Washington, DC: International Monetary Fund. https://doi.org/10.5089/9781498343671.007

26. Jordà, Ò. & Taylor, A. M. (2015). The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy. *Economic Journal*, *126*(590), 219-255. https://doi.org/10.1111/ecoj.12332

27. Lama, R. & Medina, J. P. (2019). Fiscal Deficits and Unemployment Dynamics: The Role of Productivity Gains and Wage Rigidities. *Review of Economic Dynamics*, *34*, 121-140. https://doi.org/10.1016/j.red.2019.02.007

28. Mattei, C. (2022). *The Capital Order: How Economists Invented Austerity and Paved the Way to Fascism.* University of Chicago Press. https://doi.org/10.7208/chicago/9780226818405.001.0001

29. McClelland, R. & Gault, S. (2017). *The Synthetic Control Method as a Tool to Understand State Policy.* Urban Institute. https://www.urban.org/sites/default/files/publication/89246/the_synthetic_control_method_as_a_tool_0.pdf

30. Mirdala, R. (2016). Effects of fiscal policy shocks in the euro area (lessons learned from fiscal consolidation). *Journal of Advanced Research in Law and Economics*, *7*(8), 2236-2277.

31. Okeke, A., Alexiou, C. & Nellis, J. (2021). Austerity Effects on National Output and Income Inequality: A Systematic Literature Review. *Review of Economics*, *72*(2), 97-148. https://doi.org/10.1515/roe-2021-0012

32. Pagoulatos, G. (2018). *Greece after the Bailouts: Assessment of a Qualified Failure* (GreeSE Paper No. 130). London: London School of Economics and Political Science.

33. Ray, R., Gallagher, K. & Kring, W. (2020). *IMF Austerity Since the Global Financial Crisis: New Data, Same Trend, and Similar Determinants* (GEGI Working Paper No. 11/2020). Boston, MA: Global Development Policy Center.

34. Rayl, N. (2020). *Cost of Austerity: Effect of Fiscal Consolidation in Europe Post 2010.* Available at: https://ssrn.com/abstract=3596470 or http://dx.doi.org/10.2139/ssrn.3596470

35. Revuelta, J. (2021). The Effects of the Economic Adjustment Programmes for Greece: A Quasi-Experimental Approach. *Sustainability*, *13*(9), 4970. https://doi.org/10.3390/su13094970

36. Stiglitz, J. (2014). *Austerity has been an utter disaster for the eurozone.* The Guardian. https://www.theguardian.com/business/2014/oct/01/austerity-eurozone-disaster-joseph-stiglitz

37. Woo, J., Bova, E., Kinda, T. & Zhang, Y. (2013). *Distributional Consequences of Fiscal Consolidation and the Role of Fiscal Policy: What Do the Data Say?* (IMF Working Paper No. 13/195). Washington, DC: International Monetary Fund.

Alka Obadić
University of Zagreb
Faculty of Economics
and Business
10000 Zagreb, Croatia
aobadic@efzg.hr

Viktor Viljevac
University of Zagreb
Faculty of Economics
and Business
10000 Zagreb, Croatia
vviljevac@net.efzg.hr

OCCUPATIONAL MISMATCH IN THE LABOUR MARKET OF SELECTED EU COUNTRIES

Abstract

Purpose: This paper aims to examine the occupational structures of selected EU member countries and assess their compatibility with the labour market demands, aiming to identify potential structural mismatches. The search evaluates the alignment between existing knowledge and skills in specific qualifications and those required by employers, utilizing disaggregated data from registered employment offices in Austria, Croatia, Slovenia, and Spain, spanning from 2010 to 2022.

Methodology: Methodologically, the study uses Beveridge curves, labour market tightness, and matching efficiency estimates to measure the matching needs of employers and unemployed job seekers within various occupation groups. The analysis focuses on the impact of economic downturns and fluctuations in unemployment rates on different occupation groups, with a particular emphasis on more complex and better-paid occupations.

Results: Results indicate that workers in higher-skilled occupations may experience a more resilient position within the aggregate labour market trends. However, the study reveals that deviations from these trends among occupation groups are relatively minor, underscoring the substantial influence of overall labour market conditions on all segments.

Conclusion: The research finds that differences in occupation groups have a limited impact on the broader labour market trends. Regardless of occupation complexity, improvements and deteriorations in labour market conditions affect all groups uniformly. These findings suggest a nuanced interplay between occupational structures and aggregate labour market dynamics, emphasising the need for comprehensive policy considerations to address potential mismatches and promote overall labour market efficiency.

Keywords: Mismatch, occupational groups, Beveridge curve, matching efficiency, labour market tightness

1. Introduction

The current divide between the prevailing educational and occupational framework, the skills acquired in schools and universities, and the skills demanded in the professional sphere poses a significant challenge. This discrepancy becomes increasingly challenging to navigate in the face of rapid technological advancements, presenting a substantial threat to economic growth and development. Over the long term, such a misalignment has

the potential to markedly contribute to structural unemployment in the economy. It is crucial to acknowledge that the efficacy of the matching process also depends on business cycles.

This research primarily focuses on the fundamental aspect of the matching process, specifically addressing the alignment of employers' needs with those of unemployed job seekers to fill available positions. The overall efficiency of this matching dynamics fluctuates throughout the economic cycle due to variations in the average characteristics of the labour market. Regardless of the examined dimension (i.e., education mismatch or skill mismatch), occupational mismatch always highlights some inefficiency of a country's educational system and labour market (Flisi et al., 2017). In this part of the research, the main theoretical assumptions and existing empirical findings regarding the compatibility of the existing occupational group structure and labour market needs within the European Union will be elaborated. Most labour markets are tighter than they were before COVID-19. According to IMF research (Duval et al., 2022), the main reason why employment remains restrained, particularly compared to the pre-crisis trend, is that disadvantaged groups – including low-skilled workers, older workers, or women with young children – have yet to fully return to the labour market. To investigate the labour market more deeply, we apply the labour market matching model to various occupation groups, focusing on the link between unemployment and vacancies (new job posts). As the dynamics of job matching evolve with shifting business cycles, it becomes crucial to assess the real-time correlation. The most effective visual representation of the matching process in the labour market is the Beveridge curve, which illustrates the empirical trade-off between job vacancies and unemployment.

The Beveridge curve serves as an instrument for assessing the efficiency of labour market operations. Its negative slope signifies a tendency for vacancy and unemployment rates to exhibit inverse movements throughout the business cycle. Movements in the vacancy-unemployment space are usually related to labour market tightness and labour market efficiency (Consolo & da Silva, 2019). We examine the process of demand and supply matching by estimating labour market tightness and matching efficiency, by using the traditional aggregate matching function (Cobb-Douglas form), which relates the flow of new hires to the stock of vacancies and unemployment. To study trends in different occupation groups in the labour market instead, we focus on disaggregated data. For instance, economic downturns, causing spikes in unemployment, may disproportionately impact workers in occupations demanding lower levels of knowledge and skills.

This research adds to the current prevailing literature by utilising registered data from national employment offices in four chosen EU countries (Austria, Croatia, Slovenia, and Spain). The available data is disaggregated based on different occupations, providing a valuable perspective in this regard. Previous research mainly used Labour Force Survey data that are not disaggregated into ten ISCO-88 classification of occupations[1].

Based on the conceptual framework of the existing available literature, we develop the main research question: Do workers in different occupation groups experience similar movements in labour market tightness and matching efficiency?

Our methodological approach is divided into two parts. Firstly, we construct the Beveridge curves for different occupation groups of the four countries in our sample – Austria, Croatia, Slovenia, and Spain, and secondly, we estimate labour market tightness and matching efficiency for different occupation groups for each country. The paper is divided into six parts. After introduction, the second part provides a theoretical and conceptual background regarding different aspects of the labour market outcomes for different occupation groups, as well as focus on both historical and recent empirical evidence of labour market developments in different countries. The third part explains data and variables used, provides summary statistics, and describes the implemented model. The fourth part presents the results of disaggregated Beveridge curves and the estimates of labour market tightness and matching efficiency. In the fifth part, we discuss the results and explain the main limitations of our findings, while the sixth and final chapter concludes the paper.

---

1   ISCO-88 (International Standard Classification of Occupations – ISCO) outlines a broad structure of ten major occupational groups. 2, 3 and 4 digits and levels at the aggregate level are: 1. Legislators, senior officials and managers, 2. Professionals, 3. Technicians and associate professionals, 4. Clerks, 5. Service workers and shop and market sales workers, 6. Skill agricultural and fishery workers, 7. Craft and related workers, 8. Plant and machine operators and assemblers, 9. Elementary occupations, 10. Armed forces (Europa.eu, 2022).

2. Literature review

Meeting the demands of the labour market involves establishing a successful match that emphasises the interplay between unemployment and job creation. Enhanced productivity amplifies the rewards associated with job creation, leading to an elevated rate of job formation. Consequently, a higher job creation rate facilitates the process for unemployed individuals to secure employment, ultimately decreasing unemployment. This accounts for the observed counter-cyclical (pro-cyclical) nature of unemployment (job creation), (Hornstein et al., 2005). The balance between unemployment and job vacancies fluctuates based on the robustness of the labour market requirements. In a strong labour market characterised by low unemployment and high job vacancies, increases in job openings are less likely to significantly impact unemployment. This characteristic is reflected in the steep slope of the Beveridge curve[2]. Intuitively, when lots of employers are looking to hire workers but few active job seekers are available, the process of filling job openings is slowed down by the relative scarcity of available workers (Bok et al., 2022) and the efficiency of the functioning of the labour market decreases.

Beyond its slope, the shifts of the Beveridge curve (when vacancies rise and unemployment does not fall or falls too slowly) may signal the existence of structural characteristics in the labour market (Obadić, 2016) that determine how quickly job matches occur and how long they last. The ease with which job matches are established reflects the efficiency of the matching process. A decrease in matching, indicating reduced efficiency in connecting the unemployed with available positions, corresponds to a simultaneous rise in both unemployment and job vacancies. This results in an outward shift of the Beveridge curve. Conversely, an inward shift of the Beveridge curve signals an improvement in matching efficiency. Movements along the curve itself when unemployment and vacancies move in opposite directions indicate cyclical fluctuations in economic activity (Obadić, 2005).

The Beveridge curve tends to undergo shifts over time. For instance, outward shifts in the Beveridge curve were observed across various regions in Europe during the early 1970s. One of the reasons for this is an increase in the number of unemployed individuals with an unchanged number of vacancies due to the beginning of the recession (reduced aggregate demand), and the other resulted in reduced efficiency of the adjustment process due to structural factors, such as the existence of a more rigid labour market (Obadić, 2016). During the transition period in many new EU member states, the Beveridge curve exhibited an outward shift. This indicates an increase in the number of unemployed individuals relative to vacancies, even though there were instances of a rise in job openings. Such shifts of the Beveridge curve outwards with a simultaneous increase in supply and demand indicate reduced matching efficiency, i.e. an increase in the share of structural unemployment, or may be an indication of problems of structural mismatch.

In their analysis of the United States between January 2001 and December 2017, Lange & Papageorgiou (2020) found that the Beveridge curve shifted during the Great Recession and this shift was accompanied by a decline in matching efficiency (Lange & Papageorgiou, 2020). Barrero et al. (2021) investigated the onset of the COVID-19 pandemic. They asserted that the COVID-19 recession and subsequent recovery induced a significant reallocation shock, leading to unusually large shifts of jobs and workers across various industries. These changes were pushed by stable shifts in demand, such as a move from in-person services to delivered goods and an increase in industries and occupations favourable to remote work. The pandemic consistently displaced low-skilled and older workers from employment, but the transformation of labour markets was less extensive than initially anticipated after the initial wave (Duval et al., 2022). The labour markets have tightened, as evident in the outstanding flow in unfilled job vacancies (Duval et al., 2002). Similar findings were proposed by Pizzinelli and Shibata (2023). They measured the US and the UK misallocation between job seekers and vacancies across sectors until the fourth quarter of 2021, and found that total loss in employment caused by the rise in mismatch was smaller during the COVID-19 crisis than in the aftermath of the global financial crisis. During the COVID-19 recession, both countries experienced a sharp but short-lived rise in mismatch in the second and third quarters of 2020, because employment recovery started in the second half of 2020 for the US and in early

---

2   The negative relationship between unemployment and job vacancies was first identified by William Beveridge in the 1940s, and therefore the current curve bears his name. With this curve, he wanted to determine the extent to which the economy deviates from a state of full employment (Bleakly & Fuhrer, 1997).

2021 for the UK, after the abolishment of the lock-down measures (Pizzinelli & Shibata, 2023). This poses challenges for both employers and workers, hindering the job-matching process and leading to an outward shift of the Beveridge curve.

The findings from *LinkedIn's Economic graph* data suggest that the current outward shift in the U.S. Beveridge curve has to do primarily with cyclic factors driven by an overheated economy rather than structural problems in the labour market stemming from a decrease in matching efficiency. These cyclic factors will likely diminish soon as the economy slows, suggesting that the outward shift in the Beveridge curve should largely move backward as aggregate demand relaxes (Ghayad & Dickens, 2012; Ghayad, 2022). More precisely, the COVID-19 pandemic resulted in a strong outward shift of the Beveridge curve, marked first with a substantial increase in unemployment and followed by increasing job vacancies even as the unemployment rate returned to pre-pandemic levels. As shown by Forsythe et al. (2022), by spring 2022, the U.S. labour market had largely recovered and was characterised by extremely tight markets and a slightly depressed employment-to-population ratio driven largely by retirements. The COVID-19 pandemic has dramatically and permanently changed the way we live and work. We do see that employment has reallocated somewhat away from low-skilled service jobs, and, considering the job vacancy patterns, conclude that worker preferences or changes in job amenities are driving this shift.

The data for the Netherlands analysed in the Cabus and Somers' (2018) paper show that mismatch rates, which measure employers' view on the match between employees' skills and the job requirements, are lower in those sectors in which the average years spent in formal education by workers is lower. For example, sectors such as "Construction", and "Trade, catering, repair" reported relatively low mismatch rates both in 1991 and 2011, around 12-13%. The average number of years spent in formal education was relatively low in these two sectors as well, around 12.6 in 2011. On the other hand, the "Education" sector reported a mismatch rate of 35.5% in 2011, with an average of 16.5 years spent in formal education for workers in this sector. This clearly indicates that mismatch rates increase as job complexity increases, and sectors with relatively simpler (which, of course, does not mean easier since many of the low-skill jobs are physically

very demanding) jobs have fewer problems finding workers who fit the position. However, putting these differences in sectors aside, the authors find that increases in the average schooling level of the workforce result in lower mismatch rates (Cabus & Somers, 2018).

In line with this theoretical background and the analysis of previous empirical studies, we evaluate the labour market developments in different occupation groups, as well as the relationship between newly created hires and current labour market conditions, i.e. unemployment and vacancies. The construction of the Beveridge curves allows us to compare the movements in the labour market among different occupation groups. The computation of labour market tightness enables us to examine variations in tightness dynamics across occupation groups and estimate matching functions. Through the estimation of various matching functions, we estimate the effectiveness of the matching process (matching efficiency) within different occupation groups in the four EU countries.

Therefore, we expect that the differences in occupation groups do not have a significant influence on labour market movements. We anticipate that economic downturns, which lead to increased unemployment and lower vacancies, will be felt in a similar way regardless of the differences in occupation groups, and the same outcome is expected during expansions. Moreover, we expect that labour market segments in different occupations experience similar movements in labour market tightness and matching efficiency over time.

3. Methodology

3.1 Data source and key variables

Our analysis includes four EU countries - Austria, Croatia, Slovenia, and Spain, for which disaggregated data according to occupation were available to us. The data are monthly, from January 2010 to October 2022, and were collected by national employment offices. The dataset includes three variables – Employed, Unemployed and Vacancies. Employed represents new hires, flows from the stock of the unemployed people into employment based on a new employment relationship or the start of other business activities by the previously unemployed person. Unemployed is a stock variable which represents the number of unemployed persons according to the situation on the last day of the month.

The variable Vacancies represents the stock of demanded workers that employers reported to the Employment Service during the reporting period.

We use the data disaggregated by 10 International Standard Classification of Occupations (ISCO-88) groups - managers, professionals, technicians and associate professionals, clerical support workers, service and sales workers, skilled agricultural, forestry and fishery workers, craft and related trades workers, plant and machine operators and assemblers, elementary occupations and armed forces. The occupation groups differ somewhat for Austria and are not in line with the ISCO classification, as outlined in the results section.

In constructing the Beveridge curve, the standard approach involves defining the unemployment rate as the proportion of unemployed workers to the total of employed and unemployed workers. Usually, the textbook measure of the job vacancy rate relates the number of vacancies to the size of the labour force (Obadić, 2005), while statistical databases (for example, Eurostat) often provide slightly different measures and define it as the ratio of job openings to the sum of employed workers plus job openings (Shimer, 2005). Both metrics are widely utilised, but maintaining consistency is crucial when comparing job vacancy rates across different occupation groups and time periods. Our approach to constructing the Beveridge curves differs slightly. Since we obtained the data on vacancies, unemployment and newly employed workers from different national employment offices, we were unable to obtain the data on the stock of currently employed workers

needed to calculate both the unemployment and vacancy rates. To the best of our knowledge, these data disaggregated by occupation categories do not exist, i.e. they are not collected.

However, this does not present an issue for constructing Beveridge curves. According to earlier definitions, both the unemployment and vacancy rates share the same denominator - either the sum of employed and unemployed workers or the sum of employed workers and job openings. Dividing the numerator by the same number, therefore, does not change the shape of the Beveridge curves, but only expresses (in the case of Beveridge curves) values as percentages. Such practice can be found in different papers (Gomez-Salvador & Soudan, 2022; Lange & Papageorgiou, 2020, etc.), and we also follow this approach.

3.2 Descriptive statistics

In this part of the paper, descriptive statistics for four examined countries included in our analysis are presented. We use three variables in our analysis – new flows into employment, the stock of unemployed workers and vacant positions. With these variables, we can construct the Beveridge curves, as well as estimate the matching functions. Summary descriptive statistics are presented for different occupation groups. Each time series contains a total of 154 observations, from January 2010 to October 2022. The tables (see Tables 1-4) present the mean, standard deviation, minimum value and maximum value for the aforementioned variables and countries we use in the empirical estimations.

Table 1 Summary statistics for different occupation groups, Austria

	Agricultural	Industry and small trade – 1st subgroup	Industry and small trade – 2nd subgroup	Industry and small trade – 3rd subgroup	Goods and services, sales personnel, transport	Services	Trained technicians	Administrative and clerical	Health service, teaching and cultural occupations	Total
Employed mean	1056	7948	3479	5708	6097	11237	1489	4685	3325	45213
Employed standard deviation	1098	7318	1411	1666	1259	6210	324	849	1336	12784
Employed minimum value	314	855	961	2176	3481	4509	669	2660	1899	17925
Employed maximum value	4426	32010	8162	10388	9413	28551	2269	7126	8676	82280
Unemployed mean	6052	34039	20950	49379	50034	72498	11359	41063	23086	310674

	Agricultural	Industry and small trade - 1st subgroup	Industry and small trade - 2nd subgroup	Industry and small trade - 3rd subgroup	Goods and services, sales personnel, transport	Services	Trained technicians	Administrative and clerical	Health service, teaching and cultural occupations	Total
Unemployed standard deviation	2903	18885	5014	8986	9106	20216	1748	5895	4780	62465
Unemployed minimum value	2594	15655	13183	33831	35535	40722	8284	31704	13932	207944
Unemployed maximum value	12286	76675	33067	74021	81909	167936	15601	60713	37246	522253
Vacancies mean	609	5933	8109	5973	8116	10219	5138	5280	4022	53400
Vacancies standard deviation	354	3264	3909	4093	4715	5346	3342	3676	2731	30439
Vacancies minimum value	134	1378	2907	1771	3316	4661	1353	1834	1540	21763
Vacancies maximum value	1572	13231	17079	18736	21730	30397	13555	15545	11324	141139

Source: Authors' calculations based on the Public Employment Service Austria (2022) data

Table 2 Summary statistics for different occupation groups, Croatia

	ISCO 0	ISCO 1	ISCO 2	ISCO 3	ISCO 4	ISCO 5	ISCO 6	ISCO 7	ISCO 8	ISCO 9	Total
Employed mean	1	2	1518	2491	1709	3188	109	1903	769	2258	13949
Employed standard deviation	3	2	609	855	678	1863	52	914	355	1131	5523
Employed minimum value	0	0	414	765	535	919	22	550	249	788	4760
Employed maximum value	23	10	4121	5088	3438	8215	249	4420	1753	5803	28764
Unemployed mean	15	47	16700	32817	31202	42626	2308	33033	12568	61948	233265
Unemployed standard deviation	9	22	3956	11770	10742	18356	651	19008	6592	16942	86291
Unemployed minimum value	0	15	9542	15214	14636	14068	1322	9260	4041	36489	105796
Unemployed maximum value	41	201	24406	54891	49105	76755	3481	63198	22849	92013	384376
Vacancies mean	9	17	2352	2453	1161	3417	68	2384	823	3183	15869
Vacancies standard deviation	57	7	989	942	486	1770	41	1019	441	1676	5825
Vacancies minimum value	0	4	371	547	237	665	5	515	177	508	5035
Vacancies maximum value	550	46	5430	4903	2229	7722	227	4976	2525	9120	30241

Source: Authors' calculations based on the Croatian Employment Services (2022) data

Table 3 Summary statistics for different occupation groups, Slovenia

	ISCO 0	ISCO 1	ISCO 2	ISCO 3	ISCO 4	ISCO 5	ISCO 6	ISCO 7	ISCO 8	ISCO 9	Total
Employed mean	5	92	524	475	364	810	33	930	386	879	4497
Employed standard deviation	3	24	229	109	94	232	24	478	121	333	1285
Employed minimum value	0	45	148	179	148	275	5	285	145	356	1986
Employed maximum value	23	152	1393	744	667	1767	136	2737	683	2263	8730
Unemployed mean	57	1807	6658	8386	7748	13477	603	13327	7174	18303	77539
Unemployed standard deviation	13	342	1366	1918	1090	2225	116	3928	2159	3240	15532
Unemployed minimum value	32	1140	3990	4528	4967	8319	300	5760	3520	9854	42412
Unemployed maximum value	94	2435	9129	11392	9404	17375	874	20790	10802	23619	103987
Vacancies mean	7	176	2059	1088	599	1662	46	2825	1050	2040	11553
Vacancies standard deviation	23	84	843	365	242	530	24	884	348	877	3404
Vacancies minimum value	0	60	632	313	115	560	10	1057	305	584	4336
Vacancies maximum value	160	444	4950	2139	1183	2880	133	5622	1917	4476	19527

Source: Authors' calculations based on the Employment Service of Slovenia (2022) data

Table 4 Summary statistics for different occupation groups, Spain

	ISCO 0	ISCO 1	ISCO 2	ISCO 3	ISCO 4	ISCO 5	ISCO 6	ISCO 7	ISCO 8	ISCO 9
Employed mean	174	2392	38782	35096	37325	107208	27801	67612	25906	115304
Employed standard deviation	37	640	15587	9639	8338	30827	4641	11034	4008	23681
Employed minimum value	95	1045	13689	19029	19103	53236	17901	40766	16945	73780
Employed maximum value	330	4593	128981	74718	65984	219764	40432	90578	35419	214191
Unemployed mean	1663	33900	295335	284524	410547	931465	81305	559203	220217	1097998
Unemployed standard deviation	474	5650	46485	44795	58458	97415	10275	186134	63350	131373
Unemployed minimum value	990	24199	213785	211542	314003	747077	60160	289992	125768	860665
Unemployed maximum value	2383	43698	411043	371974	521021	1127461	99669	866547	328344	1326683
Vacancies mean	7	107	3355	3072	2137	5672	7372	5874	1424	14156
Vacancies standard deviation	23	50	1234	1345	779	2247	2787	2039	1541	4725

	ISCO 0	ISCO 1	ISCO 2	ISCO 3	ISCO 4	ISCO 5	ISCO 6	ISCO 7	ISCO 8	ISCO 9
Vacancies minimum value	0	10	859	384	275	2100	1280	2078	235	5319
Vacancies maximum value	156	293	7103	10891	5512	20813	19529	15797	18167	38464

Source: Authors' calculations based on the Spanish Public Employment Service (2022) data

As pointed out earlier, the data is disaggregated according to different ISCO occupation groups for Croatia, Slovenia and Spain. Austria is an exception and uses a different classification methodology, as analysed further in the results section. It should also be emphasised that different national legal regulations exist regarding the obligation to report labour market needs by employers. That, for example, explains the relatively low number of vacancies compared to the number of unemployed workers for Spain.

To better explain possible compatibility between the existing offers and needs on the labour market, we estimate different matching functions for each observed country according to occupational groups.

3.3 Model

In the majority of macroeconomic models incorporating search and matching friction, the dynamics of new hires moving into the pool of job openings and the level of unemployment are typically represented through the aggregate matching function (Petrongolo & Pissarides, 2001; Pissarides, 2000; Bernstein et al., 2022). In the analysis of the labour market, the matching function is employed to grasp the interconnection between the number of job vacancies and unemployed workers, as well as to separate how alterations in one variable impact the other. One of the most common aggregate matching function models used in the labour market is the Cobb-Douglas matching function[3]. The function is typically represented as (Blanchard & Diamond, 1992; Kohlbrecher et al., 2014; Barnichon & Figura, 2015, Lange & Papageorgiou, 2020):

$$M_t = \text{ß}U_t^\alpha V_t^{1-\alpha}, \qquad (1)$$

where M is the number of matches or the number of outflows from the unemployed to the employed or hires, U is the number of unemployed workers, V is the number of vacancies, $ß$ indicates the efficiency of the labour market, exponents α and $1-\alpha$ are parameters that reflect the responsiveness of matches to changes in vacancies and unemployment, respectively, and t stands for linear time trend. The matching function is strictly increasing, strictly concave, twice differentiable in both arguments, and exhibits constant returns to scale (Petrongolo & Pissarides, 2001). The Cobb-Douglas matching function is universal in search and matching models, even though it imposes a constant[4] elasticity of matches with respect to vacancies that is unlikely to hold empirically (Kohlbrecher et al., 2014; Bernstein et al., 2022).

Following Barnichon and Figura (2015) and Consolo and da Silva (2019), we first define the job finding rate f_t as the ratio of new hires to the stock of the unemployed, $f_t = \frac{M_t}{U_t}$, so that

$$f_t = \beta\theta_t^{1-\alpha}. \qquad (2)$$

$\theta = \frac{V}{U}$ defines labour market tightness, and then we estimate the matching function in the log-linear form

$$lnf_{i,t} = \beta_0 + (1-\alpha) * ln\theta_{i,t} + \varepsilon_{i,t} \qquad (3)$$

The variable M *(Employed)* represents new hires, outflows from the stock of unemployment into employment. The *U (Unemployed)* variable represents the number of unemployed persons in the records according to the situation on the last day of the month and *V (Vacancies)* represents the stock of demanded workers that employers reported to the

---

3   It is named after economists Paul H. Douglas and Charles W. Cobb, who first proposed it in the 1950s.

4   The specification in log form imposes constant returns to scale so the coefficients sum to one (Lange & Papageorgiou, 2020).

national employment offices during the reporting period. As already mentioned, f_t is the job finding rate, θ_t is labour market tightness, and ε_t denotes regression residuals. The subscript i refers to different countries for which we estimate separate regression equations, i = Austria, Croatia, Slovenia, and Spain. Subscript t refers to monthly data from February 2010 to October 2022. The equation is estimated by using OLS.

The job finding rate f_t is related to a quantitative margin and a qualitative margin. The quantitative margin is the level of market tightness (vacancy-unemployment ratio), while the qualitative margin is related to the efficiency of the matching process (Consolo & da Silva, 2019). The regression residuals ε_t from Equation 3 capture the efficiency of the matching process or movements in matching efficiency for a particular occupation group in a specific country. The theoretical correlation between the job finding rate and labour market tightness is positive, indicating that increased tightness is expected to lead to a higher job finding rate. The question arises as to why we assess matching efficiency by utilising regression residuals. Let us assume that regression residuals are negative for a specific period. This means that the difference between the real (observed, empirical) job finding rate and the job finding rate predicted by the estimated matching function is negative. In other words, the observed job finding rate is lower than what one would expect based on the corresponding labour market tightness (the explanatory variable in regression) level and the estimated matching function. This means that, for some reason independent of the current labour market tightness level, the job finding rate

decreased, and this is interpreted as a decrease in matching efficiency. Positive residuals derived from the matching function estimates are construed in a similar manner, signifying an enhancement in matching efficiency. This implies a higher observed job finding rate than what would be anticipated based on the corresponding level of labour market tightness during that specific period.

Prior to computing labour market tightness and deriving estimates for the matching functions and matching efficiency, we create Beveridge curves using data on vacancies and unemployed individuals. As detailed in the Data section, we form the Beveridge curves by employing the overall counts of vacancies and unemployed workers, rather than presenting them as vacancy and unemployment rates. This approach does not alter the shapes of the Beveridge curves, enabling us to analyse both the shifts along the Beveridge curve and the inward and outward movements of the curve.

4. Results

Our research results section is divided into two parts. We present first the Beveridge curves disaggregated by occupation and then the estimates of labour market tightness and matching efficiency for different occupation groups for each country.

4.1 Beveridge curves disaggregated by occupation

In our further analysis, the disaggregated Beveridge curves are derived according to different ISCO-88 occupations for each country. Figure 1 shows disaggregated Beveridge curves for Austria.

Figure 1 Disaggregated Beveridge curves for different occupation groups, Austria
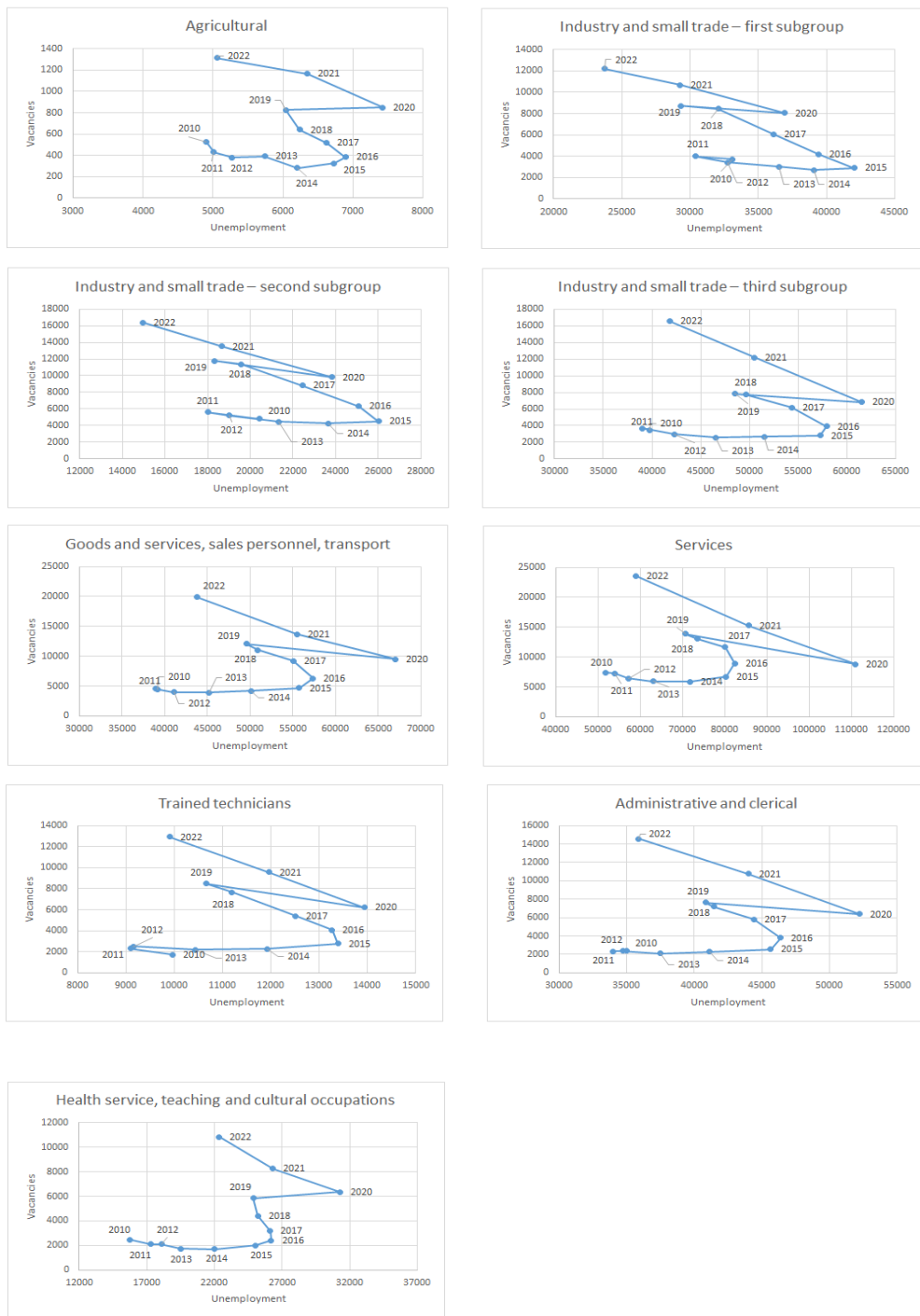
Source: Authors' calculations based on the Public Employment Service Austria (2022) data

All Beveridge curves for different occupation groups in Austria show relatively similar behaviour – the early years of the period, from 2010 to around 2016, are marked by an outward shift of the Beveridge curve, i.e. an increase in unemployment for roughly the same level of vacancies. The period from 2016 to 2019 is then marked by improving labour market conditions for workers, with unemployment decreasing and vacancies increasing for all occupation groups, with the exception of "Health service, teaching and cultural occupations". In this group, there is only a slight decrease in unemployment with an identical increase in vacancies as in other groups, which cannot indicate an improvement in matching in that group of classifications. As already mentioned, according to the aggregate

Beveridge curve for Austria, the 2020 pandemic resulted in a completely different trend in Austria, which was not present in any of the other countries in our group. Austria faced a significant increase in recorded unemployment – a strong increase in the number of unemployed workers and roughly similar levels of vacancies as in 2019. The worsening of labour market conditions was short-term, and 2021 and 2022 saw the continuation of the labour market tightening, with unemployment decreasing and vacancies increasing. The Beveridge curves disaggregated by occupation have similar shapes to the aggregate Beveridge curve, indicating rather similar developments in all areas of the Austrian labour market.

Figure 2 Disaggregated Beveridge curves for different occupation groups, Croatia
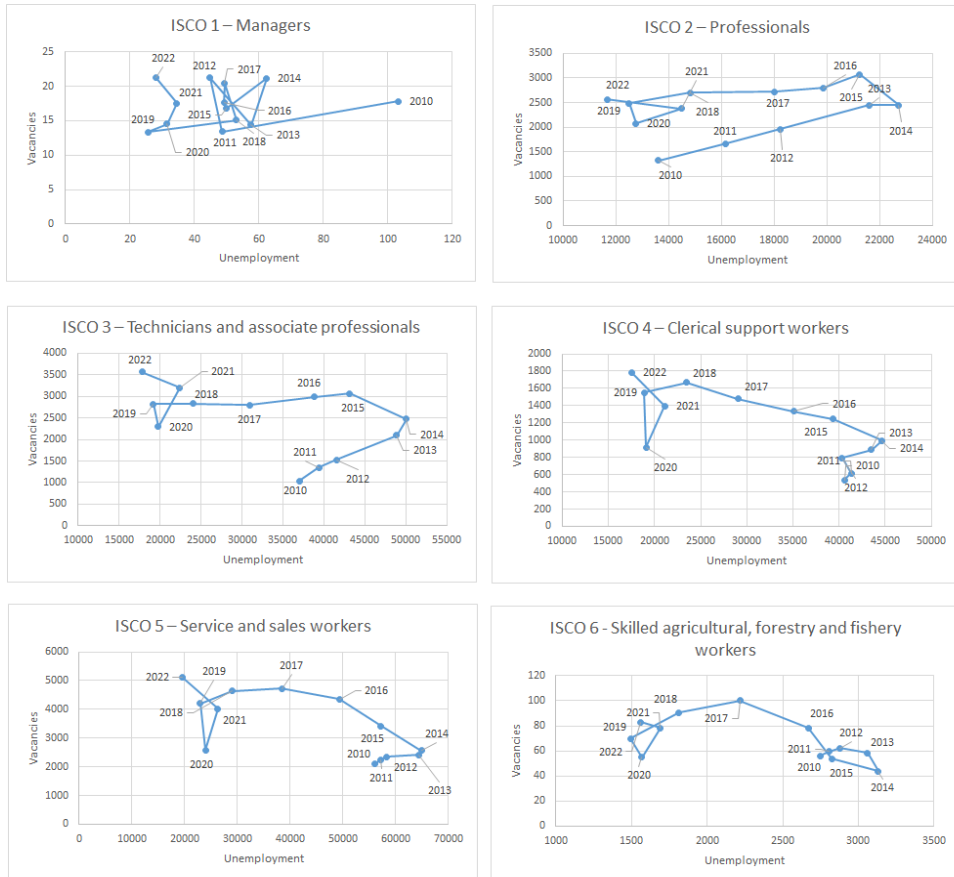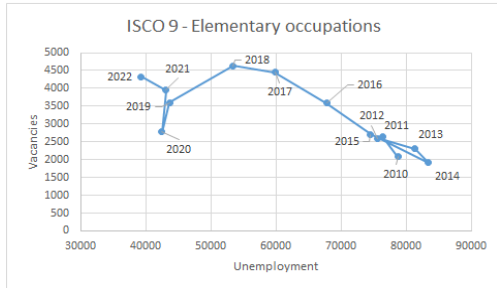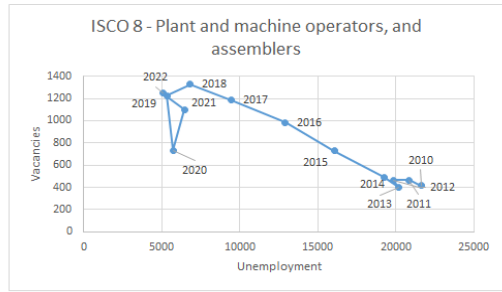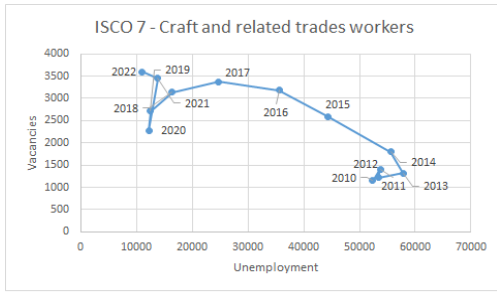
Source: Authors' calculations based on the Croatian Employment Services (2022) data

ISCO 2, 3, 4, 5 and 7 occupation groups show relatively similar behaviour. Firstly, the period from 2010 to 2014 was marked by increased unemployment, but also somewhat higher vacancies. The increases in unemployment vary from mild (ISCO 7, Craft and related trades workers) to severe (ISCO 2, Professionals), moving the Beveridge curve outwards. The period from 2014 to 2022 shows comparable movements for all but the ISCO 1 group. As the labour market conditions improved, unemployment decreased and vacancies increased, while as expected, 2020 was characterised by a drop in vacancy numbers. Unemployment did not rise noticeably in 2020 due to government measures to preserve jobs (wage subsidy measures for the private sector) in order to avoid increases in unemployment. Due to the significant recovery of aggregate demand, the year 2022 was marked by a shortage of workers in all occupation groups, which indicates increasing tightness in the Croatian labour market.

Figure 3 Disaggregated Beveridge curves for different occupation groups, Slovenia
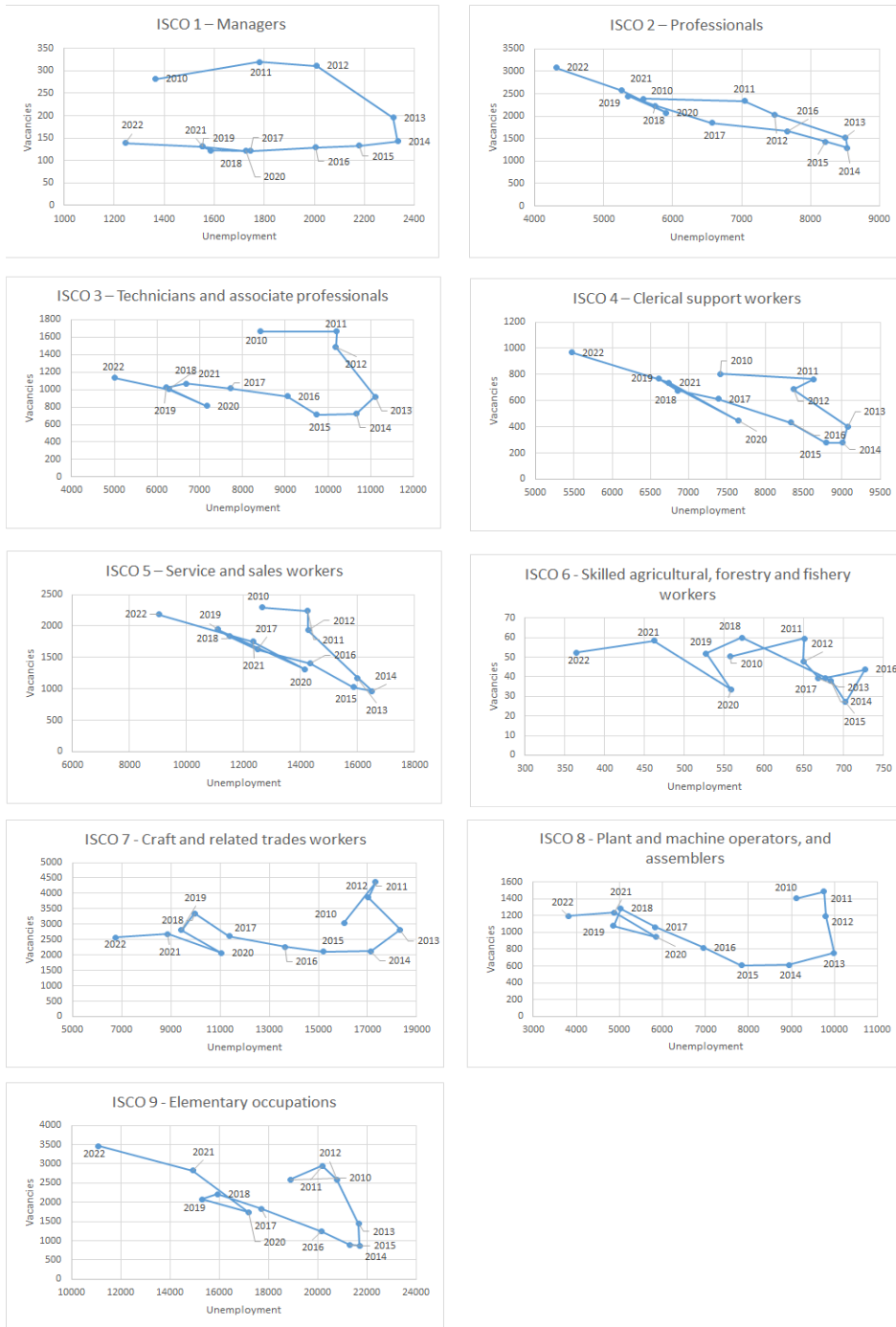
With the exception of ISCO 8, all occupation groups for Slovenia recorded increased unemployment and decreasing vacancies from 2010 to 2014, a worsening of labour market conditions. However, the subsequent period showed major improvements in labour market conditions – decreasing unemployment and increasing vacancies. As was the case in Spain and Austria, 2020 deviated from these positive developments, but the labour market continued to strengthen in 2021 and 2022. ISCO 1 (Managers) and ISCO 7 (Craft and related trades workers) groups show major improvements from 2014 to 2022, with unemployment decreasing for a roughly constant level of vacancies. The largest

post-pandemic increase in labour demand is present in the ISCO 2 (Professionals) and ISCO 9 (Elementary occupations) groups. This is in line with Obadić's (2020) findings that changes in employment shares of different occupation groups in EU-28 indicate present "job polarisation" - high-paid professionals, but also low-paid service and sales workers raise their shares in overall employment considerably. Medium-paid occupations, such as clerical support workers or craft and related trades workers and machine operators, suffered the largest losses in terms of employment share (Obadić, 2020).

Figure 4 Disaggregated Beveridge curves for different occupation groups, Spain
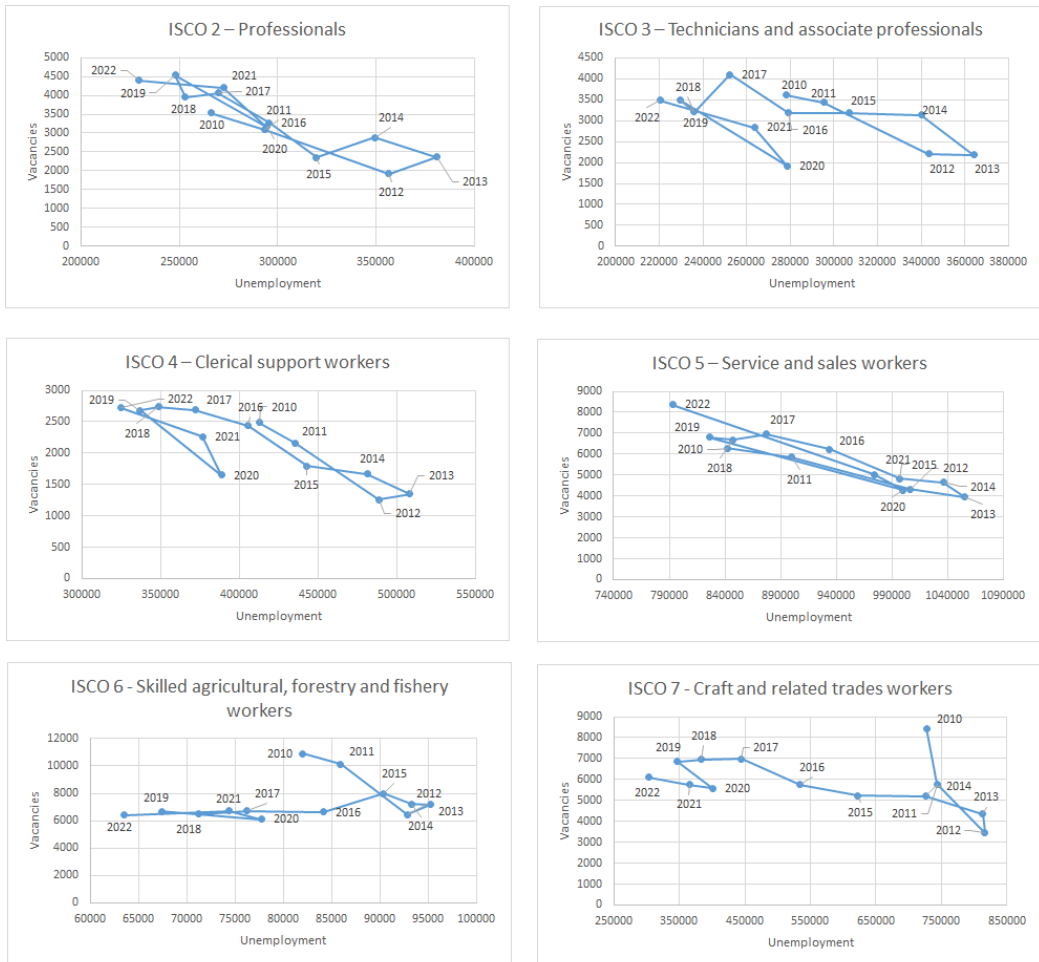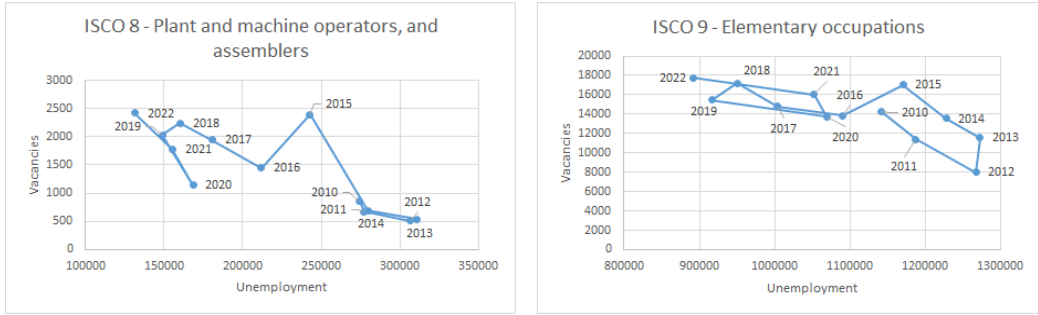
Source: Authors' calculations based on the Spanish Public Employment Service (2022) data

Disaggregated Beveridge curves for different occupation groups for Spain vary for different occupation groups but also show similar general patterns. The 2010-2013 period was marked by the worsening of the labour market conditions – unemployment increased, and the number of vacant positions decreased. The later period shows improvements in the labour market conditions – an inward move along the negatively sloped Beveridge curve (higher vacancies and lower unemployment) for ISCO 2, 3, 4, 5, 8 and 9 levels, as well as an inward straightforward shift (lower unemployment for roughly similar levels of vacancies) for ISCO 6 and 7 groups. All groups show short-term negative developments in 2020 – lower vacancies and increased unemployment, but also a subsequent recovery in 2021 and 2022. A similar conclusion as in the case of Austria applies to Spain – all occupation groups exhibit trends similar to the Beveridge curve for aggregate unemployment and vacancies.

In the next section, we present labour market tightness and our estimates of matching efficiency for different occupation groups for each country.

4.2 Empirical matching process – labour market tightness and matching efficiency by occupation groups

In this part of our analysis, we present labour market tightness and matching efficiency for different occupation groups in Austria, Croatia, Slovenia, and Spain. The corresponding estimates of the matching function (Equation 3) are shown in the Appendix.

Figure 5 Tightness by occupation groups, Austria, January 2010 – October 2022
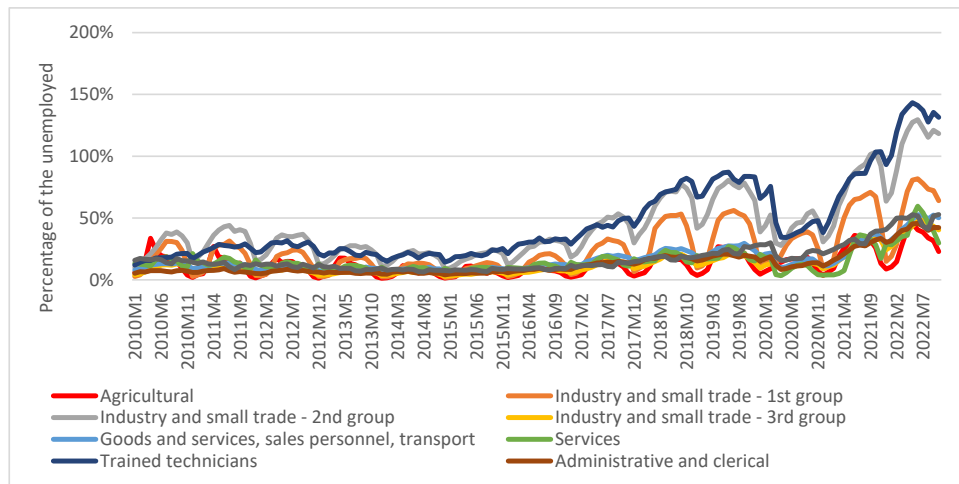
Source: Authors' calculations based on the Public Employment Service Austria (2022) data

Figure 6 Matching efficiency by occupation groups, Austria, February 2010 – October 2022
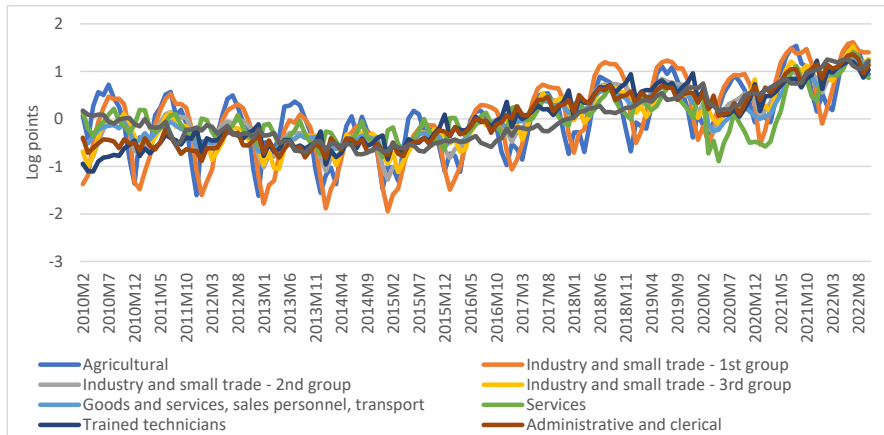
Source: Authors' calculations based on the Public Employment Service Austria (2022) data

Disaggregated by occupation, different groups of workers in Austria recorded an increase in labour market tightness during the ending years of the period. This increase was the strongest for Trained technicians and workers in the 2nd group of industry and small trade (woodworking occupations, leather producers and textile occupations). Regardless of the strength of the increase, a tight labour market is evident in 2022 for all occupation groups. Along with labour market tightening, matching efficiency recorded a steady increase from the beginning to the last years of the period, indicating that the education and skills of all occupation groups are in line with the needs of the labour market in Austria. Matching efficiency was highest in 2021 and 2022, the years also marked by the highest tightness, indicating highly aligned skills and education of the unemployed with the labour market needs in all occupation groups.

Figure 7 Tightness by occupation groups, Croatia, January 2010 – October 2022

Source: Authors' calculations based on the Croatian Employment Services (2022) data

Figure 8 Matching efficiency by occupation groups, Croatia, February 2010 – October 2022
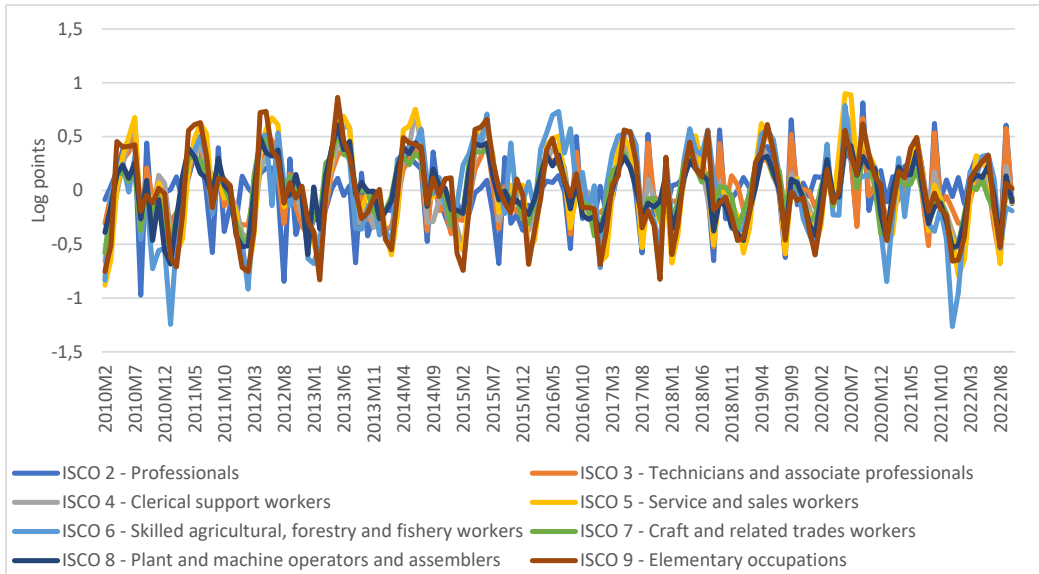
Source: Authors' calculations based on the Croatian Employment Services (2022) data

An increase in labour market tightness at the end of the period (2021 and 2022) is noticeable in all occupation groups but with considerable differences in magnitude. The increase was strongest for occupation groups such as service and sales workers, craft and related workers and professionals, and weakest for skilled agricultural, forestry and fishery workers. Labour market efficiency remained relatively similar during the entire period for all groups of workers, though the 2010-2012 period recorded somewhat lower levels of matching efficiency compared with the remaining part of the period. Since matching efficiency did not decrease along with increased tightness at the end of the period, this points to the conclusion that the skills of workers in different occupation groups are in line with the needs of the labour market. This conclusion holds more strongly for groups that experienced larger increases in tightness in 2021 and 2022 (craft and related trades, service and sales, professionals, plant and machine operations and assemblers, and technicians and associate professionals), which means that increases in demand for these workers did not result in fewer matches, or less successful job finding, compared to what one would expect based on the estimate of the matching function.

Figure 9 Tightness by occupation groups, Slovenia, January 2010 – October 2022

Source: Authors' calculations based on the Employment Service of Slovenia (2022) data

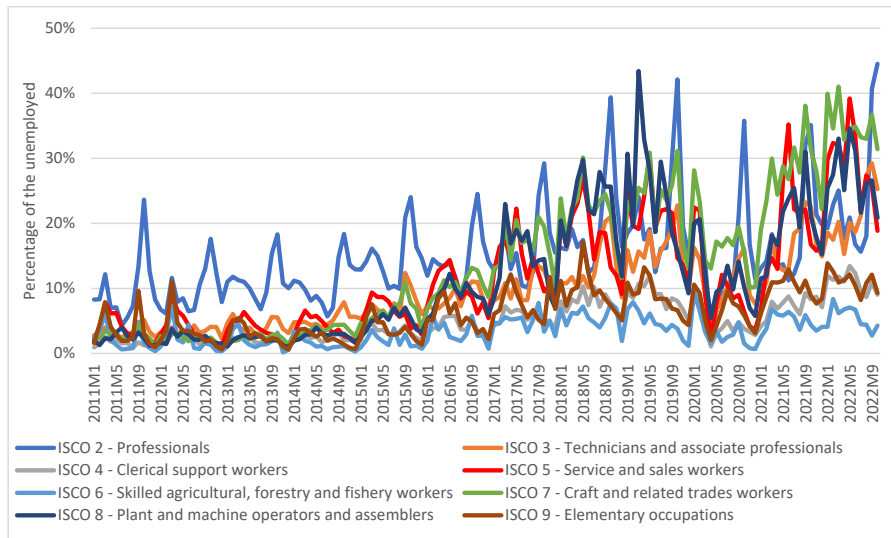**Figure 10 Matching efficiency by occupation groups, Slovenia, February 2010 – October 2022**
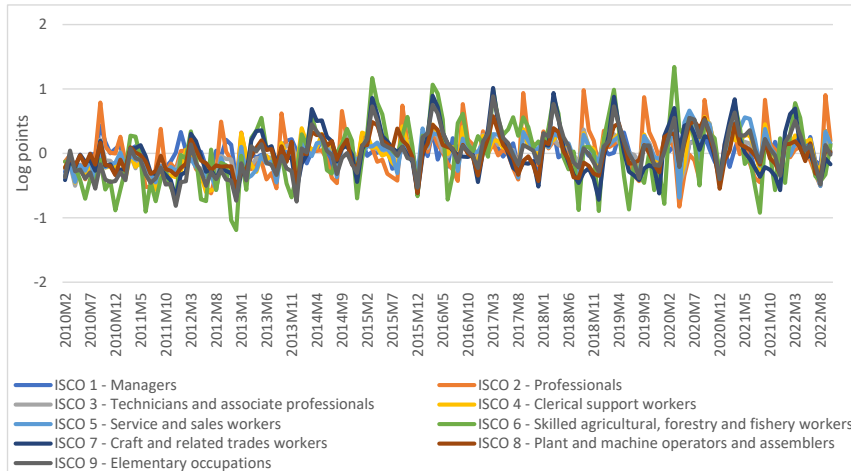
Source: Authors' calculations based on the Employment Service of Slovenia (2022) data

Matching efficiency in Slovenia was lowest during the early years of the period (2010-2013), increasing afterward. Matching efficiency remained relatively stable during the later years of the period, reaching relatively high levels during the period from 2015 to 2017. Interestingly enough, matching efficiency actually increased during 2020, the year which also recorded a drop in labour market tightness. Tightness increased in 2021 and 2022 compared to 2020,

especially for ISCO 2 - Professionals, and matching efficiency dropped only slightly compared to 2020 and the 2015-2017 period. This indicates that higher demand for workers (tightness) in Slovenia translated into more matches between the unemployed and employers without considerable losses in matching efficiency in 2021 and 2022. Therefore, the needs of the labour market are well adjusted with the education and skills of workers among dif-

ferent ISCO occupation groups. The only exception to this general trend is the ISCO 6 (skilled agricultural, forestry and fishery workers) group, which did not record considerable increases in tightness in 2021 and 2022, but did record a minor drop in matching efficiency.

Figure 11 Tightness by occupation groups, Spain, January 2010 – October 2022
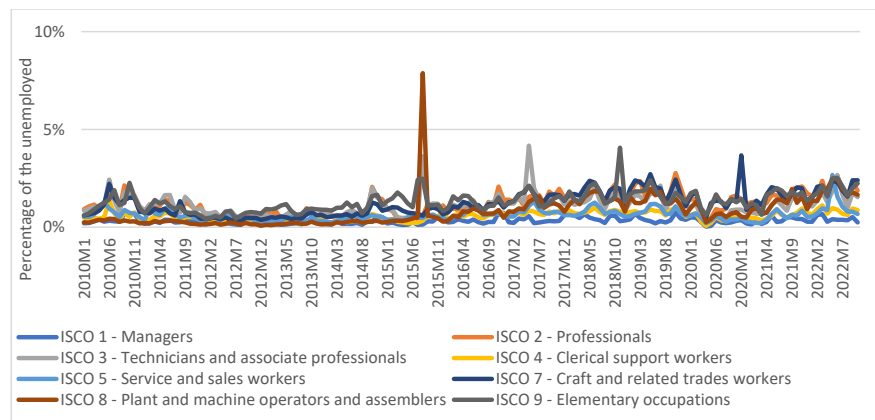
Source: Authors' calculations based on the Spanish Public Employment Service (2022) data

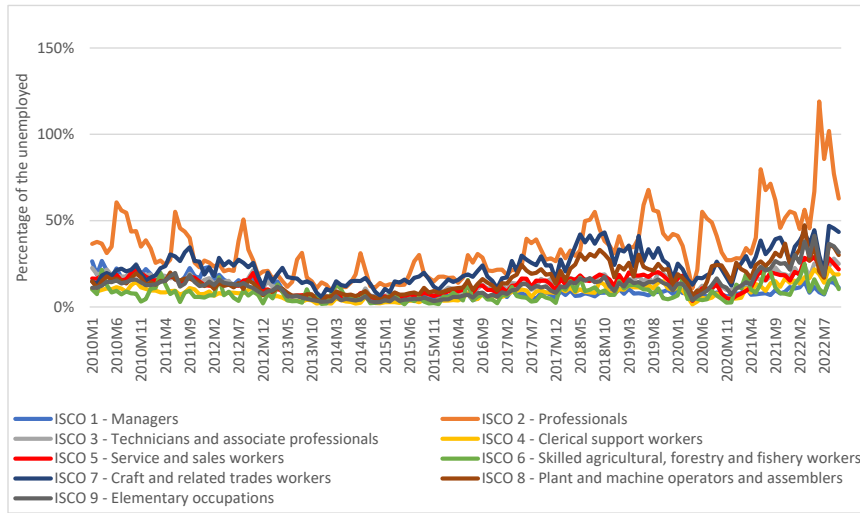*Figure 12 Matching efficiency by occupation groups, Spain, February 2010 – October 2022*
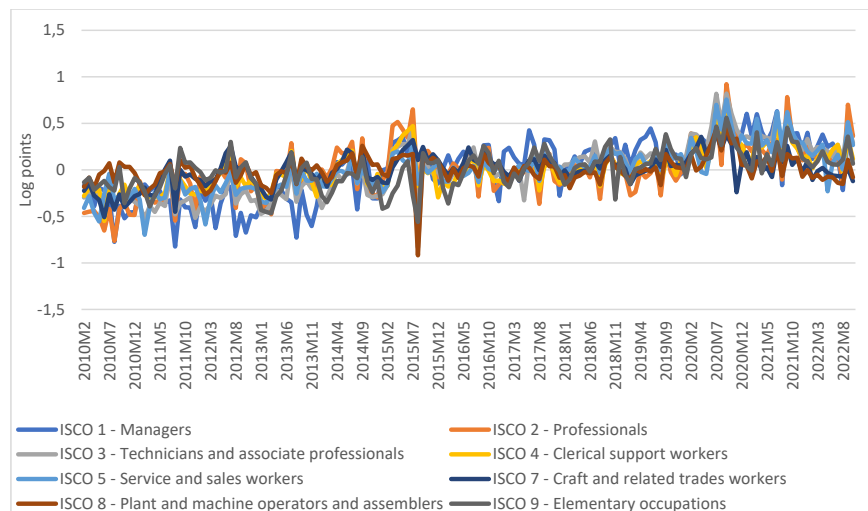
Source: Authors' calculations based on the Spanish Public Employment Service (2022) data

Matching efficiency in different occupation groups in Spain gradually increased over the 2010-2022 period, with lower efficiency in the first and higher efficiency in the second half of the period. Labour market tightness was relatively high during 2021 and 2022 in most occupation groups but with several exceptions such as ISCO 1 (managers), ISCO 4 (clerical supports workers) and ISCO 5 (service and sales workers). Overall, the results indicate that different occupation groups in Spain follow very similar trends when it comes to matching efficiency movements over time.

5. Discussion

Worker groups in different occupations in the selected group of countries experience similar trends in the labour market. However, there are exceptions to this general pattern in some countries and some occupation groups.

The Austrian labour market, disaggregated by occupation, shows very similar movements in the Beveridge curves. The Croatian labour market groups also follow similar trends, though with exceptions such as the ISCO 1 group, and slightly different shapes of the Beveridge curves for workers with higher levels of education. ISCO groups in Slovenia follow similar general patterns as well, but certain groups show their own peculiarities. For example, we found a huge increase in labour demand for ISCO 2 and ISCO 9 groups. The Beveridge curves for different labour market groups in Spain resemble the aggregate Beveridge curve but with their own peculiarities in groups like ISCO 6 and ISCO 7. Despite these exceptions, we believe it can justifiably be concluded that during the analysed period in the selected group of countries, different occupation groups in the labour market followed broadly similar trends in movements of vacancies and unemployment. In some countries, this co-movement is very strong (Austria), and in others, it is weaker (Spain, though the results for Spain need to be interpreted with caution due to the relatively low number of reported vacancies, i.e. missing data).

When it comes to our research question regarding the similarities in movements in labour market tightness and matching efficiency among the different occupation groups, similar conclusions hold – different occupation groups experienced relatively similar trends in Austria, Croatia, Slovenia, and Spain. However, this does not hold for every occupation group or during every time frame. Notable exceptions are, for example, ISCO 6 and ISCO 9 groups in Croatia regarding tightness – other occupation groups experienced an increase in tightness at the end of the period compared to the period before the pandemic, while tightness in these two groups remained like at the pre-pandemic levels.

Despite the differences in the levels of tightness and their volatility at different points in time, the general trends in tightness are similar in almost all occupation groups in the countries we have analysed. For instance, all occupation groups in Austria first

show a decrease, and then an increase in matching efficiency. The trend of increasing matching efficiency is noticeable in all occupation groups in Slovenia and Spain. All occupation groups in Croatia recorded relatively unchanging matching efficiency in the 2010-2022 period.

The prediction of future labour market trends and needs is often ungrateful and difficult. However, the COVID-19 pandemic and the ongoing technological progress in the labour market and the production processes (more frequent use of advanced robots, AI, etc.) continues to drive labour market changes. This, though, is not the only driver of the changes in the labour markets of the European Union member states. Other drivers include demand and supply factors such as population ageing and the consequent labour market shortages, and the developments in labour market institutions and policies. Well-designed active labour market policies could speed up job matching, for example through short-term training programmes that help detached (and employed) lower-skilled workers build the skills required for new fast-growing occupations or more traditional jobs that have experienced acute shortages. To accommodate shifting worker preferences, labour laws and regulations also need to facilitate telework. Immigration, whose sharp reduction slightly amplified labour shortages in some cases, could also help "grease the wheels" of the labour market (Duval, et al., 2022). As demonstrated by recent research exploring spatial variations in overall worker movements (Kuhn et al., 2021), the growing inclination towards remote work might have resulted in a geographical mismatch, as job seekers relocated away from densely populated areas where job opportunities are still concentrated. The COVID-19 pandemic has profoundly and eternally altered our lifestyle and work habits. We observe a notable shift in employment, particularly away from low-skilled service jobs. Considering the patterns in job vacancies, we can suppose that this change is driven by worker preferences or adjustments in job-related benefits.

6. Conclusion

In this paper, we have analysed the alignment of education and skills in different occupation groups with the labour market needs in Austria, Croatia, Slovenia, and Spain. Our research has two main findings. First, worker groups in different occupa-

tions in the selected group of countries experience similar trends in the labour market. This means that, for instance, during periods of decreasing unemployment and increasing vacancies in the aggregate labour market, workers in different occupation groups also record the same positive developments in a relatively homogeneous fashion in terms of the direction of these movements, though with different magnitudes. On the other hand, economic downturns, marked with increasing unemployment and decreases in vacancies are also felt relatively homogeneously among workers in different occupation groups.

The second main finding is that workers in different occupation groups experience similar movements in labour market tightness and matching efficiency. For example, during periods of increases in the aggregate labour market tightness, defined as the ratio of vacant job positions to the stock of unemployed workers, labour market tightness

also increases in different occupation groups and vice versa. A similar conclusion holds for matching efficiency, i.e. the market's ability to match unemployed workers to vacant job positions. It should also be emphasised that, though these two conclusions hold in general, exceptions to both conclusions exist in some occupation groups and in some periods of time. Future policy decisions should focus on tackling the issue of occupational mismatch and skill gaps to eliminate gaps between human capital in different occupation groups and increase their capacity regardless of the state of the business cycle. Further studies should take a step further and examine more nuanced dimensions through which mismatch may play a role in some countries. For example, examining the difference in occupational mismatch among the regions of individual member countries is important, as job seekers are increasingly encouraged to work from home and are moving away from high-density areas where vacancies are still primarily located.

References

1. Barnichon, R. & Figura, A. (2015). Labor Market Heterogeneity and the Aggregate Matching Function. *American Economic Journal: Macroeconomics*, *7*(4), 222-249. http://dx.doi.org/10.1257/mac.20140116

2. Barrero, J. M., Bloom, N., Davis, S. J. & Meyer, B. H. (2021). COVID-19 is a persistent reallocation shock. *AEA Papers and Proceedings*, *111*, 287-291. https://doi.org/10.1257/pandp.20211110

3. Bernstein, J., Richter, A. W. & Throckmorton, N. A. (2022). *The Matching Function and Nonlinear Business Cycles* (Working Paper No. 2201). Dallas, TX: Federal Reserve Bank of Dallas. https://doi.org/10.24149/wp2201

4. Bleakly, H. & Fuhrer, J. C. (1997). Shifts in the Beveridge curve, job matching and labor dynamics. *New England Economic Review*, *28*(Sept), 3-19.

5. Blanchard, O. J. & Diamond, P. (1992). The Flow Approach to Labor Markets. *The American Economic Review*, *82*(2), 354-359. https://doi.org/10.3386/w4000

6. Bok, B., Petrosky-Nadeau, N., Valetta, R. G. & Yilma, M. (2022). *Finding a Soft Landing along the Beveridge Curve* (FRBSF Economic Letter No. 2022-24). San Francisco, CA: Federal Reserve Bank of San Francisco.

7. Cabus, S. J. & Somers, M. A. (2018). Mismatch between education and the labour market in the Netherlands: is it a reality or a myth? The employers' perspective. *Studies in Higher Education*, *43*(11), 1854-1867. https://doi.org/10.1080/03075079.2017.1284195

8. Consolo, A. & da Silva, A. D. (2019), *The euro area labour market through the lens of the Beveridge curve* (ECB Economics Bulletin No. 4/2019). Frankfurt am Main: European Central Bank.

9. Croatian Employment Services (2022). *Statistics*. https://www.hzz.hr/en/statistics

10. Duval, R., Longji Li, Y. L., Oikonomou, M., Pizzinelli, C., Shibata, I., Sozzi, A. & Tavares, M. M. (2022). *Labor Market Tightness in Advanced Economies* (IMF Staff Discussion Notes No. 2022/001). Washington, DC: International Monetary Fund. https://doi.org/10.5089/9798400204340.006

11. Employment Service of Slovenia (2022). *About ESS*. https:// www.ess.gov.si/en/about-ess

12. Europa.eu (2022). *International Standard classification of occupations 2, 3 and 4 digits (ISCO-88 (com))*. https://ec.europa.eu/eurostat/documents/1978984/6037342/ISCO-88-COM.pdf

13. Flisi, S., Goglio, V., Meroni, E. C., Rodrigues, M. & Vera-Toscano, E. (2017). Measuring Occupational Mismatch: Overeducation and Overskill in Europe—Evidence from PIAAC. *Social Indicators Research*, *131*,1211-1249. https://doi.org/10.1007/s11205-016-1292-7

14. Forsythe, E., Kahn, L. B., Lange, F. & Wiczer, D. (2022). Where have all the workers gone? Recalls, retirements, and reallocation in the COVID recovery. *Labour Economics*, *78*, 102251. https://doi.org/10.1016/j.labeco.2022.102251

15. Ghayad, G. & Dickens, W. (2012). *What Can We Learn by Disaggregating the Unemployment-Vacancy Relationship?* (Research Department Public Policy Brief No. 12-3). Boston, MA: Federal Reserve Bank of Boston.

16. Ghayad, R. (2022). *The Labor Market through the Lens of the Beveridge Curve*. Linkedin Economic Graph. https://economicgraph.linkedin.com/content/dam/me/economicgraph/en-us/PDF/eg-analytics-report-aug2022.pdf

17. Gomez-Salvador, R. & Soudan, M. (2022). *The US labour market after the COVID-19 recession* (ECB Occasional Paper Series No. 298). Frankfurt am Main: European Central Bank. https://doi.org/10.2866/964976

18. Hornstein, A., Krusell, P. & Violante, G. L. (2005). Unemployment and Vacancy Fluctuations in the Matching Model: Inspecting the Mechanism. *Federal Reserve Bank of Richmond Economic Quarterly*, *91*(3), 19-51.

19. Kohlbrecher, B., Merkl, C. & Nordmeier, D. (2014). *Revisiting the Matching Function* (Kiel Working Paper No. 1909). Kiel: Kiel Institute for the World Economy. https://doi.org/10.2139/ssrn.2505377

20. Kuhn, M., Manovskii, I. & Qiu, X. (2021). *The Geography of Job Creation and Job Destruction* (NBER Working Paper No. 29399). Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w29399

21. Lange, F. & Papageorgiou, T. (2020). *Beyond Cobb-Douglas: Flexibility Estimating Matching Functions with Unobserved Matching Efficiency* (NBER Working Paper No. 26972). Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w26972

22. Obadić, A. (2005). *Makroekonomika tranzicijskog tržišta rada – S posebnim osvrtom na Hrvatsku*. Politička kultura.

23. Obadić, A. (2016). Tržište rada. In Obadić, A. & Tica, J. (Eds.), *Gospodarstvo Hrvatske* (pp. 203-246). Faculty of Economics and Business, University of Zagreb.

24. Obadić, A. (2020). Influence of technological change and digital technology on job polarization and occupational change. In Družić, G. & Gelo, T. (Eds.). *Conference Proceedings of the International Conference on Economics of Decoupling (ICED)* (pp. 433-450). Zagreb: Croatian Academy of Sciences and Arts (Economic Research Division) and Faculty of Economics and Business, University of Zagreb.

25. Petrongolo, B. & Pissarides, C. A. (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature*, *39*(2), 390-431, https://doi.org/10.1257/jel.39.2.390

26. Pizzinelli, C. & Shibata, I. (2023). Has COVID-19 induced labor market mismatch? Evidence from the US and the UK. *Labor Economics*, *81*, 102329. https://doi.org/10.1016/j.labeco.2023.102329

27. Pissarides, C. A. (2000). *Equilibrium Unemployment Theory*. MIT Press.

28. Public Employment Service Austria (2022). *Public Employment Service Austria (AMS)*. https://www.ams.at/organisation/public-employment-service-austria

29. Shimer, R. (2005). A Theory of the Matching Function. In *2005 Meeting Papers* (Vol. 673)*.* Society for Economic Dynamics.

30. Spanish Public Employment Service (2022). Statistical bulletins. https://www.sepe.es/HomeSepe/en/que-es-el-sepe/estadisticas/publicaciones-estadisticas.html

Appendix

Table 1 - Matching function estimated coefficients and p-values (shown in brackets), Austria and Croatia

Austria			Croatia		
Occupation	Constant	lnTightness	ISCED occupation	Constant	lnTightness
Agricultural	-0.81 (0.00)	0.77 (0.00)	2 – Professionals	-1.32 (0.00)	0.55 (0.00)
Industry and small trade – first subgroup	-0.81 (0.24)	0.56 (0.00)	3 – Technicians and associate professionals	-1.28 (0.00)	0.49 (0.00)
Industry and small trade – second subgroup	-0.18 (0.53)	0.45 (0.00)	4 – Clerical support workers	-1.27 (0.00)	0.49 (0.00)
Industry and small trade – third subgroup	-0.34 (0.37)	0.89 (0.00)	5 – Service and sales workers	-1.47 (0.00)	0.45 (0.00)
Goods and services, sales personnel, transport	-1.19 (0.00)	0.36 (0.05)	6 - Skilled agricultural, forestry and fishery workers	-2.06 (0.00)	0.29 (0.00)
Services	-0.92 (0.00)	0.57 (0.00)	7 - Craft and related trades workers	-1.97 (0.00)	0.32 (0.00)
Trained technicians	0.73 (0.09)	0.84 (0.00)	8 - Plant and machine operators, and assemblers	-1.82 (0.00)	0.34 (0.00)
Administrative and clerical	-0.40 (0.46)	0.85 (0.00)	9 – Elementary occupations	-1.99 (0.00)	0.45 (0.00)
Health service, teaching and cultural occupations	-1.47 (0.00)	0.23 (0.15)			

Source: Authors' calculations

Table 2 - Matching function estimated coefficients and p-values (shown in brackets), Slovenia and Spain

Slovenia			Spain		
ISCED occupation	Constant	lnTightness	ISCED occupation	Constant	lnTightness
1 – Managers	-2.95 (0.00)	0.02 (0.60)	1 – Managers	-0.75 (0.02)	0.33 (0.00)
2 – Professionals	-2.44 (0.00)	0.12 (0.02)	2 – Professionals	0.39 (0.06)	0.54 (0.00)
3 – Technicians and associate professionals	-2.45 (0.00)	0.21 (0.00)	3 – Technicians and associate professionals	-0.23 (0.33)	0.41 (0.00)
4 – Clerical support workers	-2.51 (0.00)	0.22 (0.00)	4 – Clerical support workers	-0.13 (0.49)	0.43 (0.00)
5 – Service and sales workers	-2.30 (0.00)	0.25 (0.00)	5 – Service and sales workers	0.08 (0.75)	0.44 (0.00)
6 - Skilled agricultural, forestry and fishery workers	-1.83 (0.00)	0.47 (0.00)	6 - Skilled agricultural, forestry and fishery workers	-0.65 (0.00)	0.17 (0.00)
7 - Craft and related trades workers	-2.47 (0.00)	0.17 (0.02)	7 - Craft and related trades workers	-0.01 (0.94)	0.45 (0.00)
8 - Plant and machine operators, and assemblers	-2.50 (0.00)	0.23 (0.00)	8 - Plant and machine operators, and assemblers	-0.65 (0.00)	0.28 (0.00)
9 – Elementary occupations	-2.62 (0.00)	0.20 (0.00)	9 – Elementary occupations	-0.20 (0.27)	0.47 (0.00)

Source: Authors' calculations

Ana Babić
University of Rijeka
Faculty of Economics
and Business
51000 Rijeka, Croatia
ana.babic@efri.uniri.hr

Andrea Arbula Blecich
University of Rijeka
Faculty of Economics
and Business
51000 Rijeka, Croatia
andrea.arbula.blecich@efri.uniri.hr

Nikolina Dukić Samaržija
University of Rijeka
Faculty of Economics
and Business
51000 Rijeka, Croatia
nikolina.dukic.samarzija@efri.uniri.hr

SMART CITY EFFICIENCY ASSESSMENT MODEL: MULTI-CRITERIA ANALYSIS OF 127 CROATIAN CITIES

Abstract

Purpose: The aim of this paper is to present a model for the efficiency assessment of smart cities based on 38 indicators (ISO standard 37120, ISO standard 37122 and additional indicators) in six dimensions of a smart city in order to produce a ranking of 127 cities in Croatia.

Methodology: In this study, the Data Envelopment Analysis (DEA) method was used, which was preceded by the translator invariance method due to the standardization of 38 absolute values. The analysis was performed using the input-oriented BCC model. The input values are previously formed indices for six dimensions of smart cities; the index of the development of smart cities was selected as the output.

Results: According to the results of the ranking, 33 (26%) cities are efficient, while 94 (74%) cities are inefficient. The most efficient cities are Korčula, Split, Pazin, Rijeka and Dubrovnik, while the most inefficient cities are Skradin, Petrinja, Bakar, Komiža, Glina and Kutina.

Conclusion: By identifying the dimensions that have the greatest impact on the efficiency of smart cities, DMUs gain valuable information about the position of an individual city compared to other cities. Providing an overview of existing efficiency levels and suggesting improvement measures enables targeted changes towards efficiency.

Keywords: Smart city, relative efficiency, Data Envelopment Analysis (DEA), ISO standards

1. Introduction

Smart City (SC) is a term that stands for various technologies and concepts that aim to make cities more efficient, sustainable, socially inclusive and technologically advanced. The Sustainable Development Goals (SDGs) introduced by the United Nations in its 2015 Agenda for Sustainable Development (UN, 2015) and the concept of smart city (SC) are closely linked as both aim to address various global challenges and improve the quality of life of people around the world. SC initiatives contribute to SDG 11 (Sustainable Cities and Communities) by improving urban planning, expanding public transportation, promoting sustainable infrastructure and ensuring access to basic services for all residents. Smart grids, energy-efficient buildings

and smart energy management systems contribute to SDG 7 (Affordable and Clean Energy) and SDG 13 (Climate Action) by improving access to affordable, reliable, sustainable and modern energy, thus helping to mitigate climate change. By introducing sustainable consumption and production practices (reducing waste and promoting the circular economy), SC initiatives contribute to SDG 12 (Responsible Consumption and Production). SC initiatives also contribute to SDG 3 (Good Health and Well-Being) by ensuring healthy living and promoting well-being for all ages. By improving water management, reducing water waste and improving access to clean water and sanitation, SC initiatives contribute to SDG 6 (Clean Water and Sanitation). In summary, by integrating sustainability principles and leveraging technology, SCs can play a critical role in realizing the broader sustainable development agenda outlined in the SDGs.

Efficiency is a core principle of sustainable development and a smart use of city resources leads to greater efficiency and directly impacts the creation of greater economic value and the well-being of citizens (OECD, 2019). City efficiency primarily means targeted, coordinated, and integrated management of a smart city's infrastructure resources (e.g., energy, business, transportation, waste management, public facilities, and green spaces) using information-communication technologies (ICT) that maximize the impact and significantly reduce costs, increasing at the same time the city's sustainability. The first step is to define a vision and a strategy that reflects the characteristics of a particular city (e.g., economy, climate, natural capital, social capital, geographic location, specific industries, and infrastructure) and to implement it through all administrative structures. The second step is to integrate and coordinate policies, regulations, administrative frameworks, acts, and institutional hierarchies within which city departments operate. Formulating well-integrated management processes and activities is critical to effective urban planning and shaping a vision for the future.

In general, a greater number of quality standards met, reflecting citizens' quality of life, should improve the efficiency of the city. According to the ISO 37120 standard, the certificate is awarded to cities based on the number of indicators collected, which gives them an additional incentive, i.e. the more indicators prescribed in the standard a city meets, the higher the award. The same is true for

ISO 37122, as cities that are successfully certified to ISO 37120 and are part of the World Council on City Data (WCCD) are eligible for ISO 37122 certification and able to lead the global development of smart cities with ISO-standardized, comparable city data (WCCD, 2021).

The goal of the study is to create a model for assessing the efficiency of smart cities and determine the relative efficiency of Croatian cities using DEA. This enables the identification and quantification of the efficiency factors of Croatian cities according to the Smart City Index, as all Croatian cities were included in the study. The model takes into account a large number of indicators, i.e. 38 indicators that each city must fulfil are distributed among six different smart city areas/dimensions. The inputs of the DEA analysis are the six dimensions of smart cities mentioned above, and the output is defined by the city development index.

After the introduction, this paper provides a literature review in the field of smart city efficiency assessment. After opting for the use of DEA analysis, the methodology itself is presented. It is then followed by a comprehensive description data gathering process and a variable determination process. The paper continues with a presentation of the results and finishes with a discussion and a presentation of the main findings of the paper.

2. Literature review

The operation of cities encompasses a rather diverse and extensive range of activities, from economic activities and governance to quality of social capital, environmental protection and sustainability, and connectivity through broadband and transportation infrastructure. Therefore, most authors group all these areas of a city into the dimensions of a smart city in order to define the most important activities and implement the strategy of a smart city. Smart cities are not a goal in themselves, but rather serve as tools to achieve a higher quality of life and sustainability goals. Achieving efficiency in various dimensions of smart cities is critical to their success.

The efficiency of smart cities can be viewed from many different angles. For example, the influence of smart city performance (SCP) on the technical efficiency and sustainability of cities was investigated by Auci and Mundula (2012) and Yigitcanlar and Kamruzzaman (2018). Auci and Mundula (2012) used a stochastic frontier analysis and found that people

and environment have a positive impact on technical efficiency. Conversely, other smart indicators such as governance, economy, mobility and housing have a negative impact. Kutty et al. (2022) used the Double-Frontier Slack Based Measure DEA (DFSBM-DEA) model, which accounts for undesirable factors in long-term technological sustainability performance, to comprehend the methods cities use to tackle the problem of sustainability. The most efficient smart city was determined based on six different dimensions of sustainable development, including economic dynamics, governance and institutions, energy and environmental resources, social cohesion and solidarity, safety and security, and climate change. This approach allows policymakers to compare cities and identify areas of improvement.

The DEA allows benchmarking of smart cities based on specific dimensions such as transportation, energy consumption, waste management, and healthcare. Yu and Zhang (2019) study the influence of smart city policy (SCPol) implementation on energy efficiency (EE) in China. The findings indicate notable differences between cities in different regions, with SCPol showing a significant positive impact on EE, implying that SCPol can effectively enhance EE. Fan et al. (2021) investigated the impact of SCPol implementation on the promotion of a low-carbon economy (LCE) in China. Their findings reveal a significant improvement in the LCE of cities through the implementation of SCPol. Smart energy management strategies are critical to reducing costs and environmental footprint. In their paper, Nguyen et al. (2020) discuss the benefits of implementing the Internet of Things (IoT) in the smart city transportation system. They address the use of a hierarchical approach for overall traffic management, i.e. they define the concept, methodology, and required developments of submodels describing the optimization problems of the overall system. Transportation optimization systems are important in smart cities to alleviate traffic congestion. Yi and Ma (2019) used DEA's C2R model to evaluate the solid waste logistics system. Based on the evaluation findings, waste management and transportation were rationalized. The results indicate that the system manages logistics costs efficiently, transportation efficiency is high, the risk of waste pollution is minimized, and reliability is ensured. Smart health systems play a critical role in urban environments. Research by Pacheco Rocha et al. (2019) shows that smart cities can impact pub-

lic health in a number of ways, including through disease prevention and the promotion of health initiatives. The use of smart technologies in healthcare plays an important role in patient-centred care based on patient preferences (Dukić Samaržija, Arbula Blecich & Samaržija, 2018).

Research on SC efficiency in Croatia covers various aspects related to the implementation and impact of smart city initiatives in urban areas across the country. While not measuring SC efficiency, Kelman et al. (2017) used SWAT analysis to identify the strategically important factors of a set of indicators used to govern, measure performance and improve the quality of life, using ISO 37120 Sustainable development of communities - Indicators for city services and quality of life. Jurlina Alibegović et al. (2018) assessed the largest cities in Croatia using the smart urban development index and six different dimensions (economy, people, governance, mobility, environment and living) of smart cities. According to the smart urban development index, only eleven major Croatian cities are above average, while it shows the heterogeneity of cities in all six SE dimensions. Čukušić et al. (2019) examined smart city initiatives and applications based on the simulation experiment and the opinions of 60 smart city experts. The total time of the implementation process, expressed in days, was estimated for 11 smart city applications in the city of Split. The time represents the average number of days until the start of implementation of a particular application. The results of the simulation experiment are in line with the expert assessments of the priorities and potential of smart city applications.

Buntak et al. (2021) analyzed the state of smart city development in Croatia by collecting and analyzing key dimensions of the smart city. They concluded that the implementation of smart cities is not satisfactory, which is mainly due to the lack of strategic thinking. The lack of smart city research is also evident in the systematically peer-reviewed publications dealing with research on SC in Southeast Europe (SEE). Ninčević Pašalić et al. (2021) examined, analyzed and classified seventy-four papers based on their focus on SC topics and common subtopics. While smart governance has been extensively researched in the SEE region, topics related to smart economy and smart people have received little attention from researchers. SC research in SEE is still in the conceptualization and planning stage, and

there is very little evidence of the actual implementation and follow-up activities.

Overall, studies on smart cities in Croatia contribute to understanding and evaluating the efficiency of smart city initiatives in Croatia, considering citizen perspectives, financial capacities, and governance strategies. The main contribution of this paper is a holistic approach (based on 38 indicators over six SC dimensions) to efficiency assessment in order to provide a complete and credible ranking of all 127 Croatian cities, with the aim of identifying the sources of inefficiency and providing suggestions for improving the dimensions of SC.

3. Methodology

Data Envelopment Analysis (DEA) is a non-parametric mathematical method based on linear programming. It is often used to evaluate the relative efficiency of homogeneous decision-making units (DMUs) in different domains, with the main objective of optimizing resource allocation, improving performance and benchmarking. It was originally introduced by Charnes et al. (1978) and it is the estimation of the production function. The ratio of the weighted inputs to the weighted outputs is calculated for each DMU. The value q can range between 0 and 1, whereby $q = 1$ means the relative efficiency of 100%, and a value below 1 means that the DMUs operate relatively inefficiently. Relatively efficient DMUs are compared and serve as a benchmark for relatively inefficient DMUs. Relatively efficient DMUs are not able to increase any output without increasing inputs or reducing the remaining outputs, and conversely, they cannot decrease any input without simultaneously decreasing any of their outputs or increasing any of the remaining inputs.

There are two basic DEA models, the CCR model (Charnes et al., 1978), the basic DEA model, which assumes constant returns to scale (CRS), and the BCC model (Banker et al., 1984), which assumes variable returns to scale (VRS). The CRS assumes that an increase in each input for each DMU leads to a proportional increase in each output, while the VRS assumes that this relationship need not be proportional. When choosing the appropriate model, the orientation (input or output) should also be selected based on the strategy pursued by the DMUs, analysts, managers, etc. If the goal is to minimize inputs while achieving (at least) a previously reached output level, an input-oriented model is used, while an output-oriented model is chosen if the goal of DMUs is to maximize outputs with (at most) a given number of inputs. According to Mundula and Auci (2016), an efficient city has the ability to maximize its own impact (well-being) with respect to a set of inputs, i.e. to minimize the use of its own resources (inputs) to achieve a given outcome.

Since an input-oriented BCC model (VRS) is used in this paper, it can be expressed as follows (Aminuddin & Ismail, 2016):

$$Max\theta_0 = \sum_{j=1}^{m} u_j y_{j0} + u_0)$$

subject to

$$\sum_{i=1}^{s} v_i x_{i0} = 1$$

$$\sum_{i=1}^{m} u_j y_{jk} - \sum_{i=1}^{s} v_i x_{ik} + u_0 \leq 0$$

$$v_i \geq 0, \quad u_j \geq 0, \quad u_0 \text{ free in sign}$$

(1)

Although DEA was originally developed to assess the relative efficiency of public sector DMUs, its application was later extended to the non-public sector as it can take into account multiple inputs and multiple outputs expressed in different units. DEA is often used in the area of public policy and services to evaluate the efficiency of health services (Dukić Samaržija, Arbula Blecich and Najdek, 2018; Mourad et al., 2021), educational institutions (Arbula Blecich, 2021; 2020; Arbula, 2012; Marto et al., 2022), transportation systems (Rivero Gutiérrez et al., 2022), and municipal service providers (Cerović et al., 2017). The application of DEA also extends to environmental analysis, where organizations and industries seek to assess their environmental efficiency. In this context, DEA can evaluate how efficiently resources are used to minimize environmental impacts such as pollution and waste (Elhami et al., 2016). This is especially important in today's world where sustainability and environmental responsibility are of utmost importance to investors and society. Moreover, in urban planning, DEA can help evaluate the efficiency of various urban services and infrastructure components in smart cities. DEA can evaluate how efficiently resources are used to deliver services such as transportation, energy, and waste management in terms of quality and quantity of service delivery (Nguyen et al., 2020; Yi & Ma, 2019). It can also help find ways to improve or expand urban services with-

out significantly increasing resource consumption, contributing to the development of sustainable and efficient smart cities.

DEA is a useful analytical tool for increasing the efficiency of smart cities in a variety of areas. From assessing a city's overall performance and benchmarking specific dimensions to promoting sustainability, smart energy management, and optimizing transportation and health services, DEA plays an important role in promoting the success of smart city initiatives. Benchmarking helps cities identify best practices and areas for innovation. Researchers such as Giffinger et al. (2007) have introduced a multidimensional framework for smart cities that includes dimensions such as smart economy, smart governance, smart citizens, smart environment, smart living, and smart mobility. These dimensions are used in this paper to evaluate the efficiency of Croatian cities in their operation as smart cities and to help policy makers identify areas for improvement and allocate resources efficiently.

4. Determination of variables and data collection

The model variables were determined in several phases. In the first phase, the following dimensions of smart cities were determined: Smart Economy (SE), Smart Governance (SG), Smart Citizens (SC), Smart Living (SL), Smart Environment (SEN), and Smart Mobility (SM).

An economy is considered smart if it is able to develop innovative and modern solutions that meet the needs of the market. This means that the concept of a smart city encompasses several urban dimensions associated with SE, such as innovation, entrepreneurship and the labor market. In addition, factors such as productivity, flexibility and integration into national and international markets are also important components that need to be considered (Griffinger et al., 2007; Babić et al., 2022). Certainly, the development of a country's ICT sector plays an important role (Babić, 2021a) in achieving sustainable economic, environmental and social development.

When planning smart city solutions, particular attention is paid to the economic, social and environmental dimensions, as well as to the accountability and transparency of the governance. Extensive databases, spatial decision support systems and corresponding geodata technologies are used to facilitate city administration. In addition, smart cities

are continuously innovating in the field of e-government to provide more efficient public services to their citizens (Vinod Kumar & Dahiya, 2017).

Today, city governance focuses mainly on physical features, including roads, the built environment, sewerage systems, and green spaces. Spatial planning, housing, transportation and waste management are at the forefront, while little attention is paid to e-government, i.e. the collection and processing of the vast amounts of data that can be collected as part of urban governance.

The main reason for the existence of a city and the formulation of its policies are its citizens. A crucial aspect in the development of smart cities is therefore active participation of smart citizens in urban activities in various roles (Madakam, 2014). The concept of Smart Citizens implies the educational level of the citizens of a given city and their willingness to engage in lifelong learning and progress.

Smart Living means the application of ICT with information systems for urban services in the areas of water, gas and electricity supply, telecommunications, banking, etc. in order to increase the well-being of citizens. In addition, it implies modern homes equipped with smart appliances and automation systems (Bawa et al., 2016). Although healthcare is often located outside of the city limits, the availability and accessibility of primary healthcare is an important feature of smart cities.

Smart Environment refers to the effective use of ICT to protect natural resources and cultural heritage at city level (Staffans & Horelli, 2014). It also refers to the protection of the environment and natural heritage, but especially to sustainability (Al-Nasrawi et al., 2015), as it provides advanced tools and technologies to monitor, detect, measure, and record all changes in the urban environment, as well as tools and technologies to ensure sustainability (waste management, wastewater treatment, intelligent transportation systems, etc.).

Mobility plays a crucial role in modern smart cities, as the transportation of people and goods inside and outside of the city plays a fundamental role in economic development and the improvement of everyday urban life. The difference between mobility and Smart Mobility is that in smart mobility, the public has access to real-time information and monitoring of public transportation, using ICT such as the Internet of Things, GPS, smart cameras and geographic information systems (Albino et al., 2015; Vanolo, 2014), as well as the use of environmentally friendly fuels in public transportation (Manville et al., 2014). Smart mobility is possible

with an appropriate 5G network concept that combines the advantages of fixed and mobile communications (Babić et al., 2019), and this is supported by the fact that all Croatian cities are beneficiaries of the WiFi4EU initiative (European Commission, 2023), i.e. all of them have a free Wi-Fi network (Babić, 2021b).

In order to provide a holistic approach, the model consists of 38 indicators, and some of the indicators have been added so that the model can provide a more complete picture and a more credible ranking. The number of trades was included because trades play a crucial role in the economy of cities; in some cities, the number of trades exceeds the number of companies. The number of employees in the ICT sector is an indicator of the ISO 37122 standard, but it is supplemented by data on the number of ICT companies in a given city, as the application and implementation of ICT in all areas of the economy and society is the basis for creating competitiveness and ensuring the city's continued economic and social progress. In addition, the road connection to the nearest airport, the number of e-charging stations and the share of expenditure on environmental protection in the city's total expenditure, the transparency of the city's budget and the total household expenditure per inhabitant, as well as the share of citizens holding a university degree per 1,000 inhabitants, are recorded.

The data was collected from publicly available or officially requested data from public institutions. Part of the data was collected through a detailed search and analysis of the websites of all 127 Croatian cities, as well as precise measurements of the distance of all cities to the nearest airport on the Google map. The data was collected for the years 2019 and 2020 to facilitate monitoring and comparison with the same indicators in the future period, which ensures verification and comparison of the data.

Standardization was inevitable due to the range of values for the different indicators. This was done using the z-transformation method. This method involves standardization of the indicator values with the mean of 0 and standard deviation of 1 as normal distribution conditions.

$$z = \frac{-x-\mu}{\sigma} \qquad (2)$$

The z-value, i.e. the standard score, quantifies the number of standard deviations by which the individual values of the observed numerical characteristic are below or above its average value. In other words, the z-value provides information about the relative position of a particular value within the overall distribution compared to the average value. Values for which a higher value is less desirable are inverted by multiplying by -1, as is the case for greenhouse gas emissions, PM10 concentration, etc.

Finally, indices are calculated and created for different smart city dimensions. Each index is assigned a z-value based on equation (2) for each indicator. Based on the z-values of the indicators, the average scores of the dimensions and consequently the average scores of each city were calculated.[1]

Each index of smart city dimensions for each individual city (DMU) represents an input for the analysis of the efficiency of the Croatian cities. The output in this model is the index of development (y1) for each city (DMU).

Table 1 Inputs and outputs

CROATIAN CITIES	Inputs - Dimension indices						Output
	Smart Economy	Smart Governance	Smart Citizens	Smart Living	Smart Environment	Smart Mobility	Index of development
DMU DOMAINS	Economy Finances	Management	Education Recreation Culture	Health Housing Safety	Energy Environment and climate Solid waste Water	Telecommunications Traffic	
DMU	SE	SM	SC	SL	SEN	SM	IoD
1 - 127	x_1	x_2	x_3	x_4	x_5	x_6	y1

Source: Authors

---

1   This work is a result of research conducted as part of the doctoral dissertation: Babić, A. (2021) Učinkovitost gradova Republike Hrvatske prema normama ISO 37120, ISO 37122 i dimenzijama pametnih gradova. Due to the extensive analysis, the reader is referred to the mentioned literature for additional explanations.

The development index was created using the balanced z-score method, a non-linear technique for creating composite indices. In this method, the values of the individual indicators are converted into standardized scores and combined into a composite index using the arithmetic mean and a penalty coefficient (Denona Bogović et al., 2017). This method is also known as the Mazziotta-Pareto index (De Muro et al., 2011). This index was developed with the aim of solving the problem of objectively measuring, evaluating, comparing and ranking units at various territorial administrative levels on the basis of their level of development in a given period. This

was particularly important when individual or multiple units have mismatched sets of indicators, i.e. when they perform above average on some indicators and below average on others.

5. Results and discussion

After determining the DMUs, i.e. Croatian cities, and the input and output values, the efficiency analysis was carried out according to the BCC and CCR models. The following table presents the descriptive statistics for the inputs and outputs included in the model.

Table 2 Descriptive statistics of inputs and outputs

	X1	X2	X3	X4	X5	X6	Y
Max	0.77528	1.09076	1.06757	3.3404	83.9615	0.72155	115.637
Min	0.2556	0.33049	0.29752	0.30977	0.30613	0.25179	91.167
Average	0.51855	0.52438	0.52931	0.56863	1.17456	0.48127	103.613
SD	0.08798	0.11752	0.12872	0.32787	7.37609	0.10881	5.09866

Source: Authors

The comparison of the results of the input-oriented models CCR and BCC is summarized in Table 3 and a detailed presentation of the results

of the efficiency analysis for both models for all 127 DMUs individually can be found in Babić (2021b).

Table 3 Comparison of CCR and BCC models

RESULTS OF THE ANALYSIS	CCR UU model	BCC UU model
Relatively efficient cities (number)	20 (16%)	33 (26%)
Relatively inefficient cities (number)	107 (84%)	94 (74%)
Average relative efficiency	0.8968	0.9382
Standard deviation	0.0777	0.0568
The lowest relative efficiency value	0.662	0.7708
The number (%) of cities whose relative efficiency is lower than the average	56 (44%)	60 (47%)

Source: Authors

The efficiency frontier is determined based on those DMUs that achieve the maximum efficiency ($\theta=1$) of inputs to achieve a given level of output. Looking at relative efficiency according to the CCR model, out of the total population of all cities in the Republic of Croatia, 20 cities are relatively efficient and define the efficiency frontier. These cities are: DMU5

(Sveta Nedelja), DMU9 (Zaprešić), DMU11 (Klanjec), DMU15 (Zabok), DMU33 (Varaždin), DMU34 (Varaždinske Toplice), DMU53 (Opatija), DMU55 (Rijeka), DMU60 (Senj), DMU69 (Nova Gradiška), DMU72 (Biograd na Moru), DMU76 (Zadar), DMU100 (Sinj), DMU102 (Split), DMU111 (Buzet), DMU114 (Pazin), DMU119 (Vodnjan - Dignano),

DMU120 (Dubrovnik), DMU121 (Korčula), and DMU125 (Čakovec). Since 16% of the cities are relatively efficient, their results appear as a reference set for relatively inefficient ones. These cities allow benchmarking for all inefficient cities, i.e. the reference set for inefficient cities consists of relatively efficient cities whose input-output orientation is closest to that of an inefficient city, but which are efficient based on the output achieved with the given resources.

Each inefficient city is assigned at least one reference DMU that represents best practice. According to the BCC model and taking into account the VRS, DMUs have a higher relative efficiency compared to the CCR model. The lowest value of relative efficiency is by 0.11 lower in the CCR model than in the BCC model, but the number of cities where relative efficiency is below average is higher in the

BCC model (60) than in the CCR model (56). For further analysis, the input-oriented BCC model is used because it is translationally invariant with respect to both inputs and outputs, in contrast to the CCR model, which is not translationally invariant because of the shape of the envelope it forms (Knox Lovell & Pastor, 1995). This is important for the analysis presented in this paper.

The reference set allows relatively inefficient DMUs to pursue achievable goals. Since the projection of inefficient cities to the efficiency frontier is expressed by a linear combination of efficient cities from the reference set whose coefficients represent their share of a single city's projection to the efficiency frontier, the city with the highest corresponding coefficient should be chosen as the reference. The reference set for inefficient cities is shown in the following table.

Table 4 Reference set for inefficient cities

DMU	Reference city	Number of cities	%	Reference set of inefficient cities
DMU120	Dubrovnik	28	30%	Cres, Crikvenica, Garešnica, Glina, Ivanić-Grad, Kraljevica, Križevci, Krk, Kutina, Ludbreg, Makarska, Mali Lošinj, Novi Vinodolski, Omiš, Poreč, Rab, Rovinj, Samobor, Šibenik, Komiža, Sisak, Slavonski Brod, Solin, Trogir, Umag, Velika Gorica, Vis, Vrbovec
DMU114	Pazin	15	16%	Belišće, Daruvar, Županja, Ivanec, Hvar, Krapina, Lepoglava, Opuzen, Slatina, Orahovica, Supetar, Virovitica, Vrgorac, Novigrad, Hrvatska Kostajnica
DMU34	Varaždinske Toplice	10	11%	Knin, Trilj, Mursko Središće, Oroslavje, Ozalj, Petrinja, Popovača, Pleternica, Sveti Ivan Zelina, Zlatar
DMU14	Pregrada	6	6%	Beli Manastir, Benkovac, Donja Stubica, Duga Resa, Ilok, Otok
DMU95	Imotski	5	5%	Buje, Donji Miholjac, Kutjevo, Novi Marof, Vukovar
DMU11	Korčula	5	5%	Prelog, Ploče, Pag, Novalja, Grubišno Polje
DMU121	Klanjec	4	4%	Obrovac, Skradin, Pakrac, Slunj
DMU15	Zabok	4	4%	Nin, Požega, Lipik, Gospić
DMU111	Buzet	4	4%	Vrlika, Stari Grad, Delnice, Bakar
DMU80	Đakovo	3	3%	Metković, Drniš, Valpovo
DMU125	Čakovec	3	3%	Vodice, Novska, Jastrebarsko
DMU69	Nova Gradiška	2	2%	Vinkovci, Našice
DMU102	Split	1	1%	Kaštela
DMU76	Zadar	1	1%	Bjelovar

DMU	Reference city	Number of cities	%	Reference set of inefficient cities
DMU56	Vrbovsko	1	1%	Čabar
DMU55	Rijeka	1	1%	Karlovac
DMU5	Sveta Nedelja	1	1%	Kastav
TOTAL		94	100%	

Source: Authors

The cities most frequently mentioned in the reference sets of inefficient cities are DMU120 (Dubrovnik), DMU114 (Pazin) and DMU34 Varaždinske Toplice. By comparing empirical and projected values for all 127 cities, the sources of inefficiency and their magnitude were identified for six inputs and one output. A larger percentage difference between the projected and empirical values of a particular input or output means that the input or output is a greater source of inefficiency. The following table shows, for all inputs and outputs, the changes that inefficient cities must make to achieve relative efficiency. These are average percentage changes per inefficient city and changes for individual cities that require the greatest improvement in individual inputs or outputs.

Table 5 Results on average improvements for relatively inefficient inputs

Inputs/Outputs	Empirical value	Projected value	Difference (%)
(x1) Smart Economy	0.5185	0.4789	7.0622
(x2) Smart Governance	0.5244	0.4674	9.8196
(x3) Smart Citizens	0.5293	0.4618	10.8195
(x4) Smart Living	0.5686	0.5053	9.5471
(x5) Smart Environment	1.1746	0.4761	7.9944
(x6) Smart Mobility	0.4813	0.4219	11.2067
(y1) Index of development	103.613	106.756	3.1506

Source: Authors

Since all input values are inverted, all output values are increased, even though the interpretation of the results speaks of a decrease in input. The projection shows that input x1 (SE) can achieve the same output level with 7.06% less input, input x2 (SG) with 9.82% less input, input x3 (SC) with 10.82% less input, input x4 (SL) with 9.55% less input, input x5 (SEN) with 8% less input, and for input x6 (SM) with 11.21% less input with the same output level, which represents increases in the non-inverted values. This shows that inputs x3 and x6 have the greatest impact on city inefficiency, followed by x2, x4, and x5, while input x1 has the least impact on city inefficiency.

6. Discussion and conclusion

Smart cities use data and technology to operate efficiently, promote economic development, improve sustainability and enhance the quality of life of people living and working in the city. The indicators of ISO standard 37120 - *Sustainable cities and communities*, and ISO standard 37122 - *Indicators for smart cities and additional indicators applicable to all cities* were used as a basis for creating a model for assessing the efficiency of Croatian smart cities within the framework of smart urban units, regardless of their size. The main contribution of this paper is that for the first time, all Croatian cities have been ranked using comparable and verifiable data based on a comprehensive database (38 indicators), which provides a model for assessing the efficiency of Croatian smart cities.

DEA was used for the efficiency analysis. The input values were represented by previously formed indices for six dimensions of smart cities, and the output was the development index, which is a reliable official statistic and one of the most important instruments of regional policy. The ranking resulted in efficient cities that are at the efficiency frontier and whose reference value is equal to one, while inefficient cities have values below one. The analysis was carried out using the input-oriented BCC model. This analysis identified efficient cities, inefficient cities and a reference group consisting of efficient cities that are most similar to the inefficient cities. For cities identified as relatively inefficient, the causes of relative inefficiency were identified in order to make recommendations and predictions on how much at least one dimension needs to be improved in order to reach the efficiency frontier based on their reference group. By identifying the dimensions that have the greatest impact on the efficiency of smart cities, cities gain valuable information about their individual position compared to other cities. Providing an overview of existing efficiency levels and suggesting improvement measures enables targeted changes toward efficiency.

However, the study has certain limitations, which mainly relate to the data set from 2019 and 2020. This is due to the fact that the paper is based on a comprehensive doctoral dissertation investigation, which includes both publicly available data and data officially requested from public institutions, as well as a detailed search and analysis of the websites of all 127 Croatian cities and accurate measurements of the distance of all cities to the nearest airport on Google Maps. Furthermore, the output variable in this model is the index of development of Croatian cities, for which no new data has been published yet. For further research, it is advisable to conduct a detailed analysis of the individual SC dimensions and introduce the second stage statistical analysis to determine the influence of exogenous factors. Since no data over a longer period of time was available, a DEA analysis based on one year was performed, while for further research it would be advisable to perform a dynamic DEA analysis. Finally, the model also has additional potential by extending the analysis of the existing model to new regional entities (municipalities) and applying new methods and measurement approaches. This comes with limitations, mainly because many initiatives in Croatian cities are still in their infancy and there are still many obstacles to overcome, especially when it comes to collecting data for the purpose of additional analysis in cities.

Regarding Smart Mobility in the context of telecommunications and transportation in particular, it is clear that not all cities have the same speed or access to the internet and that not all cities have a developed transportation system equipped with digital parking solutions, payment technologies, online parking or the presence of electric charging stations. Without these smart technologies, inefficient cities use telecommunication and transportation resources in a sub-optimal way, which is also reflected in their level of development.

The Smart Citizen dimension is the second largest source of inefficiency in inefficient cities. It relates to education, sport and culture, and it is one of the most important dimensions when it comes to a city's social capital. The implementation of more effective and targeted population policies as well as the co-financing of student activities and scholarships, investment in sports infrastructure and sporting events are just some of the measures that could improve the use of resources in this dimension.

The impact of the Smart Living dimension relates to the quality of life in terms of smart technologies (smart electricity), security (live cameras) and primary healthcare (the ratio of primary healthcare physicians to inhabitants). These are indicators that need to be improved and are not always the responsibility of the city administration. The specific increase in the share of smart electricity is the responsibility of Hrvatska elektroprivreda (HEP d.d.), and there are numerous ongoing projects related to the implementation of smart living solutions that have a direct impact on the better quality of life of citizens.

This study also suggests that Smart Governance is a very important dimension that needs to be improved, as current governance models should utilize the city's resources (businesses, people and technology) more efficiently. Smart Environment and Smart Economy are two dimensions that have proven to be less problematic, but still require better resource utilization, such as better waste management and water supply systems, as well as better road connections and investment in research and development.

To summarize, this paper provides a comprehensive assessment of Croatian (smart) cities and identifies the main sources of inefficiency. This should help policy makers and urban planners to focus on the areas where improvements are needed in order to promote overall urban development and smart city initiatives.

References

1.   Albino, V., Berardi, U. & Dangelico, R. M. (2015). Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology*, *22*(1), 3-21. https://doi.org/10.1080/10630732.2014.942092

2.   Al-Nasrawi, S., Adams, C. J. & El-Zaart, A. (2015). A conceptual multidimensional model for assessing smart sustainable cities. *Journal of Information Systems and Technology Management*, *12*(3), 541-558. https://doi.org/10.4301/S1807-17752015000300003

3.   Aminuddin, W. M. & Ismail, W. R. (2016). Integrated simulation and data envelopment analysis models in emergency department. *AIP Conference Proceedings*, *1784*, 050003. https://doi.org/10.1063/1.4966822

4.   Arbula Blecich, A. (2020). Factors affecting relative efficiency of higher education institutions of economic orientation. *Management, 25*(1), 45-67. https://doi.org/10.30924/mjcmi.25.1.3

5.   Arbula Blecich, A. (2021). Relative efficiency of R&D in European countries. *Journal of the Polytechnic of Rijeka*, *9*(1), 169-185. https://doi.org/10.31784/zvr.9.1.11

6.   Arbula, A. (2012). Using Data Envelopment Analysis to Evaluate the Efficiency of Higher Education Expenditure in Europe. *Journal of International Scientific Publications: Economy &* Business, *6*, 59-69.

7.   Auci, S. & Mundula, L. (2012). *Smart Cities and a Stochastic Frontier Analysis: A Comparison among European Cities*. Available at SSRN: https://ssrn.com/abstract=2150839 or http://dx.doi.org/10.2139/ssrn.2150839

8.   Babić, A. (2021a). Digital skills as a perspective of economic development and an important factor of digital transformation. *Economic Review*, *72*(1), 59-87. https://doi.org/10.32910/ep.72.1.3

9.   Babić, A. (2021b). *A Study of the Efficiency of Cities in the Republic of Croatia in accordance with ISO 37120 and ISO 37122 standards and the dimensions of smart cities* [Doctoral dissertation, University of Rijeka]. University of Rijeka.

10.   Babić, A., Jardas Antonić, J. & Buljat, B. (2022). Ranking of Croatian Cities according to Hellwig's Information Capacity Method in the Smart Economy Dimension. *Ekonomska misao i praksa*, *31*(2), 347-368. https://doi.org/10.17818/EMIP/2022/2.1

11.   Babić, A., Vlačić, E. & Sokolić, D. (2019). Strategic Positioning of Emerging 5G Technology – Barriers and Perspectives. *The Journal of Corporate Governance, Insurance, and Risk Management (JCGIRM)*, *1*, 10-22. https://doi.org/10.56578/jcgirm060102

12.   Banker, R. D., Charnes, A. & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078-1092. https://doi.org/10.1287/mnsc.30.9.1078

13.   Bawa, M., Caganova, D., Szilva, I. & Spirkova, D. (2016). Importance of Internet of Things and Big Data in Building Smart City and What Would Be Its Challenges. In Leon-Garcia A. et al. (Eds.), *Smart City 360*, *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol. 166* (pp. 605-616). Springer. https://doi.org/10.1007/978-3-319-33681-7_52

14.   Buntak, K., Budimir Šoško, G., Grgurević, D. & Cvitković, I. (2021). Key Factors of Effective Development of Smart Cities in Croatia. Pomorski zbornik, *61*(1), 91-103. https://doi.org/10.18048/2021.61.07.

15.   Cerović, Lj., Šegota, A. & Maradin, D. (2017). Efficiency of municipal service providers in the Republic of Croatia. *Croatian Operational Research Review*, *8*(2), 537-562. https://doi.org/10.17535/crorr.2017.0035

16.   Charnes, A., Cooper, W. W. & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429-444. https://doi.org/10.1016/0377-2217(78)90138-8

17.   Čukušić, M., Jadrić, M. & Mijač, T. (2019). Identifying challenges and priorities for developing smart city initiatives and applications. *Croatian Operational Research Review, 10*(1), 117-129. https://doi.org/10.17535/crorr.2019.0011

18. De Muro, P., Mazziotta, M. & Pareto, A. (2011). Composite Indices of Development and Poverty: An Application to MDGs. *Social Indicators Research*, *104*, 1-18. https://doi.org/10.1007/s11205-010-9727-z

19. Denona Bogović, N., Drezgić, S. & Čegar, S. (2017). *Studija - Evaluacija postojećeg i prijedlog novog modela za izračun indeksa te izračun novog indeksa razvijenosti jedinica lokalne i područne samouprave u Republici Hrvatskoj*. Center for Local Economic Development, University of Rijeka, Faculty of Economics and Business.

20. Dukić Samaržija, N., Arbula Blecich, A. & Najdek, T. (2018). Investigation of the Reimbursement Scheme in Croatian Public Hospitals: A Data Envelopment Analysis Approach. In Omazić, M. A. et al. (Eds.). *Proceedings of the 28th International Scientific Conference on Economic and Social Development* (pp. 161-165). Varaždin: Varaždin Development and Entrepreneurship Agency; Warsaw: Faculty of Management University of Warsaw; Koprivnica: University North.

21. Dukić Samaržija, N., Arbula Blecich, A. & Samaržija, L. (2018). The Paradigm of Patient-Centered Care in the Public Health Decision-Making. *Ekonomska misao i praksa*, *27*(2), 503-516.

22. Elhami, B., Akram, A. & Khanali, M. (2016). Optimization of energy consumption and environmental impacts of chickpea production using data envelopment analysis (DEA) and multi objective genetic algorithm (MOGA) approaches. *Information Processing in Agriculture, 3*(3), 190-205. https://doi.org/10.1016/j.inpa.2016.07.002

23. European Commission (2023). *Shaping Europe's digital future*. https://digital-strategy.ec.europa.eu/hr/activities/wifi4eu-citizens

24. Fan, S., Peng, S. & Liu, X. (2021). Can Smart City Policy Facilitate the Low-Carbon Economy in China? A Quasi-Natural Experiment Based on Pilot City. *Complexity in Economics and Busines*. https://doi.org/10.1155/2021/9963404

25. Griffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovic, N. & Meijers, E. (2007). *Smart cities - Ranking of European medium-sized cities*. http://www.smart-cities.eu/download/smart_cities_final_report.pdf

26. International Organization for Standardization (2018). *ISO 37120 2018 - Sustainable cities and communities - Indicators for city services and quality of life*. https://www.iso.org/obp/ui/#iso:std:iso:37120:ed-2:v1:en

27. International Organization for Standardization (2019). *ISO 37122 2019 - Sustainable development in communities - Indicators for Smart Cities*. https://www.iso.org/obp/ui/#iso:std:iso:37122:ed-1:v1:en

28. Jurlina Alibegović, D., Kordej-De Villa, Ž. & Šagovac, M. (2018). Smart City Indicators: Can They Improve Governance in Croatian Large Cities?. *EIZ Working Papers*, *5*, 5-48.

29. Kelemen, R., Zdjelar, R., Jakšić, M. & Hleb, M. (2017). Modeling the smart city performance, Croatian Case. *Smart Cities and Regional Development Journal*, *1*(1), 21-42. https://doi.org/10.25019/scrd.v1i1.7

30. Knox Lovell, C. A. & Pastor J. T. (1995). Units invariant and translation invariant DEA models. *Operations Research Letters*, *18*(3), 147-151. https://doi.org/10.1016/0167-6377(95)00044-5

31. Kutty, A. A., Kucukvar, M., Abdella, G., Bulak, M. E. & Onat, N. (2022). Sustainability Performance of European Smart Cities: A Novel DEA Approach with Double Frontiers. *Sustainable Cities and Society*, *81*. https://doi.org/10.1016/j.scs.2022.103777

32. Madakam, S. (2014). Smart Cities - Six Dimensions (A Scholarstical Articles Review). In *Proceedings of the International Conference on Advances in Computing and Information Technology*. http://dx.doi.org/10.3850/978-981-07-8859-9_09

33. Manville, C., Cochrane, G., Cave, J., Millard, J., Pederson, J. K., Tharuup, R. K., Liebe, A., Wissner, M., Massink, R. & Kotterink, B. (2014). *Mapping Smart Cities in the EU*. European Parliament. https://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/507480/IPOL-ITRE_ET(2014)507480_EN.pdf

34. Marto, M., Lourenço Marques, J. & Madaleno, M. (2022). An Evaluation of the Efficiency of Tertiary Education in the Explanation of the Performance of GDP per Capita Applying Data Envelopment Analysis (DEA). *Sustainability*, *14*(23), 15524, https://doi.org/10.3390/su142315524

35. Mourad, N., Habib, A. & Tharwat, A. (2021). Appraising healthcare systems' efficiency in facing COVID-19 through data envelopment analysis. *Decision Science Letters*, *10*(3), 301-310. https://doi.org/10.5267/j.dsl.2021.2.007

36. Mundula, L. & Auci, S. (2016). Smart cities: a policy tool for city efficiency? REAL CORP 2016 – SMART ME UP!. In *Proceedings of the 21st International Conference on Urban Planning, Regional Development and Information Society* (pp. 589-597).

37. Nguyen, D. D., Rohács, J, Rohács, D. & Boros, A. (2020). Intelligent Total Transportation Management System for Future Smart Cities. *Applied Sciences*, *10*(24), 8933. https://doi.org/10.3390/app10248933

38. Ninčević Pašalić, I., Ćukušić, M. & Jadrić, M. (2021). Smart city research advances in Southeast Europe. *International Journal of Information Management*, *58*, 102-127. https://doi.org/10.1016/j.ijinfomgt.2020.102127

39. OECD (2019). *The Economy of Well-being: Creating opportunities for people's well-being and economic growth* (SDD Working Paper No. 102). Paris: OECD.

40. Pacheco Rocha, N., Dias, A., Santinha, G., Rodrigues, M., Queirós. A. & Rodrigues, C. (2019). Smart Cities and Healthcare: A Systematic Review. *Technologies*, *7*(3), 58. https://doi.org/10.3390/technologies7030058

41. Rivero Gutiérrez, L., De Vicente Oliva, M. A. & Romero-Ania, A. (2022). Economic, ecological and social analysis based on DEA and MCDA for the management of the Madrid Urban Public Transportation System. *Mathematics*, *10*(2), 172. https://doi.org/10.3390/math10020172

42. Staffans, A. & Horelli, L. (2014). Expanded Urban Planning as a Vehicle for Understanding and Shaping Smart, Liveable Cities. *The Journal of Community Informatics*, *10*(3). https://doi.org/10.15353/joci.v10i3.3439

43. UN (2015). *Transforming our world: the 2030 Agenda for Sustainable Development.* https://documents.un.org/doc/undoc/gen/n15/291/89/pdf/n1529189.pdf?token=pOnhvXEp4zQeHg 6Yqq&fe=true

44. Vanolo, A. (2014). Smartmentality: The Smart City as Disciplinary Strategy. *Urban Studies*, *51*(5), 883-898. https://doi.org/10.1177/0042098013494427

45. Vinod Kumar, T. M. & Dahiya, B. (2017). *Smart Economy in Smart Cities: Advances in 21st Century Human Settlements*. Spinger. https://doi.org/10.1007/978-981-10-1610-3

46. World Council on City Data (WCCD) (2021). *ISO 37120.* https://www.dataforcities.org/wccd

47. Yi, H. & Ma, H. (2019). Optimization of Municipal Solid Waste Logistics System based on Data Envelopment Analysis. *Ekoloji Dergisi*, *107*, 2643-2654.

48. Yigitcanlar, T. & Kamruzzaman, M. (2018). Does smart city policy lead to sustainability of cities?. *Land Use Policy*, *73*, 49-58. https://doi.org/10.1016/j.landusepol.2018.01.034

49. Yu, Y. & Zhang, N. (2019). Does smart city policy improve energy efficiency? Evidence from a quasi-natural experiment in China. *Journal of Cleaner Production*, *229*, 501-512. https://doi.org/10.1016/j.jclepro.2019.04.316

Amira Pobrić
University of East Sarajevo
Faculty of Economics Brčko
76100 Brčko, Bosnia
and Herzegovina
amira.pobric@gmail.com

GOING CONCERN AUDIT REPORTING IN BOSNIA AND HERZEGOVINA

Abstract

Purpose: This paper deals with going concern audit reporting in Bosnia and Herzegovina. The research objectives are to determine whether, in conditions of increased economic uncertainty, auditors issue this type of audit report more frequently, whether the bankruptcy of a company can be predicted based on a going concern audit report, and whether companies that receive this type of audit report engage in audit opinion shopping.

Methodology: The research was conducted on a sample of audit reports of 187 companies referring to the period from 2017 to 2021. Content analysis method was used.

Results: The average rate of going concern audit reports was 19.2%. Observed by year, the rate of going concern audit reports ranges from 18.1% to 19.9%. All companies that received a going concern audit report in the considered period one or more times are still operating. In 17.5% of cases, companies replaced the auditor after receiving a going concern audit opinion, while in 16.4% of cases, companies replaced the auditor even though they did not receive a going concern audit opinion.

Conclusion: The increase in economic uncertainty during the coronavirus pandemic led to only a slight increase in the rate of going concern audit reports. This type of audit report cannot serve as a predictor of the company's bankruptcy. Companies that receive a going concern audit report do not engage in audit opinion shopping in order to avoid receiving the same type of audit report in the following year.

Keywords: Going concern audit report, going concern assumption, material uncertainty related to going concern

1. Introduction

Financial statements are the basic means by which management communicates information about the financial condition, profitability and cash flows of the company to owners and other external stakeholders. Financial statements should contain complete and reliable information about the financial effects of all company activities undertaken in one accounting period.

According to International Accounting Standard 1, financial statements are prepared using the going concern assumption. This assumption implies that the company will continue its operations for the foreseeable future and that it has the operational and financial capabilities to maintain business

continuity (Simamora & Hendarjatno, 2019). This means that the company will be able to use its assets and meet its obligations through normal business operations. The going concern assumption cannot be applied if the management intends to liquidate or cease operations or has no alternative but to do so. "When the company's management is aware that there are significant uncertainties related to the events or the conditions that may cast doubt on the company's ability to continue operating indefinitely, the company is required to disclose any such uncertainty in its financial statements. When a company does not prepare financial statements on a going concern basis, it must disclose this fact, together with the basis on which the financial statements have been prepared, as well as the reason why the going concern assumption has not been met" (IASB, 2003).

The main purpose of auditing financial statements is to determine whether the information contained in financial statements is true and objective, i.e., whether financial statements are prepared in accordance with the applicable financial reporting framework. During the audit, the auditor is required, among other things, to assess the company's ability to operate for an unlimited period. The auditor should collect a sufficient volume of adequate audit evidence on the basis of which they will conclude whether the management's use of the going concern assumption in the preparation of financial statements is appropriate. Furthermore, based on the collected evidence, the auditor should conclude whether there is uncertainty regarding the company's ability to continue as a going concern.

If the financial statements are prepared using the going concern assumption and the auditor considers that the use of this assumption is not appropriate, they will express an adverse opinion. If the auditor concludes that there is significant doubt about the company's ability to continue as a going concern, but that the application of the going concern assumption is appropriate, the auditor will assess whether the financial statements disclose any material uncertainty related to events or circumstances that give rise to significant doubts about the company's ability to continue as a going concern, and whether these events or circumstances are adequately described. If the financial statements include the required disclosures, the auditor will express an unqualified opinion and include in the auditor's report a separate paragraph

under the subtitle "Material Uncertainty Related to Going Concern" (MURGC), in which they will state that there are events or circumstances that indicate the existence of significant uncertainty that may cause significant doubt about the company's ability to continue as a going concern, and also draw attention to the notes in the financial statements describing these matters and state that the auditor's opinion has not been modified in relation to that matter. If the financial statements do not include appropriate disclosures about material uncertainty, the auditor will express a qualified opinion or an adverse opinion, as appropriate (IAASB, 2015).

An audit opinion given by an auditor in conditions where there is significant doubt about the company's ability to continue as a going concern one year after the date of the financial statement is called a going concern audit opinion. A going concern audit opinion means an unqualified audit opinion with an emphasis of matter for the existence of material uncertainty related to going concern, i.e., with the MURGC paragraph, and a modified audit opinion where the existence of the specified uncertainty is a reason for modifying the audit opinion. An audit report that contains this type of audit opinion is called a going concern audit report.

The subject of this paper is going concern audit reporting in Bosnia and Herzegovina. The research objectives are to determine the frequency of issuing a going concern audit report, whether, in conditions of increased economic uncertainty, auditors issue this type of audit report more frequently, whether the bankruptcy of a company can be predicted based on a going concern audit report, and whether companies that receive this type of audit report engage in audit opinion shopping. The paper is organized as follows. The second section provides an overview of previous research on a going concern audit opinion. Research design is presented in the third section. The research results are presented in the fourth section. Concluding considerations are given in the fifth section.

2. Theoretical and conceptual background

In the audit literature, the going concern audit opinion has been investigated from different aspects. The identified research areas relate to determining the frequency of issuing a going concern audit opinion, identifying the factors that influence the issuance of this type of audit opinion, develop-

ing a model that will help auditors to give the most appropriate audit opinion in circumstances when there is doubt about a company's ability to continue as a going concern, and examining the impact of the going concern audit opinion on the behavior of participants in the financial market.

Cheffers et al. (2010) investigated the frequency of issuing a going concern audit opinion on a population of audit reports submitted to the US Securities and Exchange Commission between 2000 and 2009. They determined that the rate of issuing going concern audit opinions among US companies ranged from 14.4% in 2003 to 20.3% in 2008. This research showed that the frequency of issuing going concern audit opinions increased significantly during and after the global financial crisis in 2007 and 2008 compared to the period before this crisis. This means that auditors express a going concern audit opinion more frequently in a period of increased economic uncertainty. Xu et al. (2011) reached the same conclusion. They investigated the frequency of issuing different types of audit opinions in Australia in the period from 2005 to 2009. They conducted the research on a sample of 8,382 Australian listed companies, which is approximately 95% of the total number of companies listed on the Australian stock exchange during that period. The results of their research showed that the frequency of issuing going concern audit opinions increased significantly during and after the global financial crisis, from 12% in the period from 2005 to 2007, to 18% and 22% in 2008 and in 2009, respectively. The lack of credit liquidity and a decrease in economic activity during the global financial crisis significantly increased business risk, which led to an increase in uncertainty regarding the ability of companies to continue as a going concern, which was reflected in audit reports. Circumstances that led the auditors to express doubts about the ability of companies to continue as a going concern relate to the fact that the companies were making a loss or a negative operating net cash flow, that they were not able to settle their obligations on the due date, and that they had limited opportunities to obtain additional sources of financing.

Carson et al. (2016) also analyzed trends in audit reporting in Australian listed companies, covering the period from 2005 to 2013. The research was conducted on a sample of 15,855 audit reports. Their research showed that in the period from 2005 to 2007, the percentage of audit reports containing

a going concern audit opinion was stable at a level of around 12%. Most of these reports contain an unqualified audit opinion with an emphasis of matter for the existence of material uncertainty related to going concern. In the period from 2008 to 2010, the percentage of audit reports relating to going concern increased to around 22%, which is a consequence of the impact of the global financial crisis on the operations of Australian companies. It is interesting to note the trend of increasing the percentage of going concern audit opinions in the period from 2011 to 2013, with a peak of 33.3% in 2013. The authors hypothesize that the increase in the frequency of going concern audit opinions in the period from 2011 to 2013 is a consequence of the impact of the slowdown in Chinese economic growth on the Australian economy and the increased scrutiny of auditors by regulatory authorities regarding the applicability of the going concern assumption. Carson et al. (2016) also found that a company that received a going concern audit opinion in one year is more likely to receive the same opinion in the following year. They also determined that the rate of companies that receive a going concern audit opinion and continue with business in the following year is at a level of 92% to 94%.

Carson et al. (2016) also found that Australian companies that receive a going concern audit opinion are not inclined to audit opinion shopping, that is, they do not replace the auditor in search of an auditor that will give them a more favorable audit opinion. Namely, the percentage of companies that received a going concern audit opinion and replaced the auditor in the following year is lower compared to companies that did not receive a going concern audit opinion, but replaced the auditor. This is in contrast to the results obtained by Chung et al. (2019). Based on a sample of 11,628 US financially distressed companies, covering the period from 2004 to 2012, they found that these companies successfully engage in audit opinion shopping in order to avoid a going concern audit opinion. The results of other research studies on this topic are not harmonized. For example, Carcello and Neal (2003) found that companies tend to replace auditors after receiving an unfavorable audit opinion. However, Krishnan (1994) and Krishnan and Stephens (1995) found that companies that replace the auditor after receiving an unfavorable audit opinion do not receive a more favorable audit opinion by the next auditor. This could mean that the replacement

of the auditor is not related to audit opinion shopping or that audit opinion shopping has not been successfully realized because the next auditor does not agree to compromise their independence. In contrast to the research mentioned earlier, Lennox (2000) found evidence that companies successfully use audit opinion shopping.

Successful audit opinion shopping indicates a lack of independence of the auditor, which can affect audit quality. However, financially distressed companies may have a strong motive to engage in audit opinion shopping in order to avoid a going concern audit opinion because this type of audit opinion can have negative consequences for the company's operations. Menon and Williams (2010) found that institutional investors react negatively to the issuance of a going concern audit report. Namely, institutional investors react to the going concern auditor opinion by reducing their shares in the capital of the companies that have received this type of audit opinion, thus causing a drop in the stock price of these companies. This research has shown that the reaction of institutional investors is more negative if the inability of the company to settle its obligations on the due date and problems in obtaining additional sources of financing are cited as reasons for expressing doubts about the company's ability to continue as a going concern. Geiger and Kumas (2018) also determined that after the issuance of a going concern audit report, institutional investors increase the sale of stocks of the companies that received this type of audit report. Chen et al. (2016) determined that the going concern audit opinion also affects lending conditions. They compared the lending conditions in the year after the companies received a clean audit opinion and the lending conditions in the year after the companies received a going concern audit opinion. They found that in the year after the going concern audit report, creditors approve a smaller credit limit, with higher interest rates and most often with the collateral requirement.

And finally, we will mention the results of the research conducted by Vučković-Milutinović (2019). All aforementioned research studies were conducted on a sample of companies from developed countries. Vučković-Milutinović (2019) investigated the frequency of modified audit opinions, including going concern audit opinions, among listed companies in Serbia, which is a developing country, just like Bosnia and Herzegovina. The re-

search was conducted on a sample of audit reports of 112 listed companies that were issued in the period from 2015 to 2017. This research showed that in the period from 2015 to 2017 the average rate of going concern audit reports was 21.4%. The highest rate was recorded in 2016, when 27% of companies received a going concern audit report. About 2% of audit reports contained an unqualified audit opinion with an emphasis of matter for the existence of material uncertainty related to going concern, approximately 10% of audit reports contained a modified audit opinion with an emphasis of matter for the existence of material uncertainty related to going concern, while in approximately 9% of audit reports, the existence of uncertainty related to going concern was the reason for modifying the audit opinion. Auditors most often expressed doubts about a company's ability to continue as a going concern due to factors such as operating at a loss, having current liabilities that exceeded current assets, carrying substantial debt, and violating credit obligations.

3. Research design

The research was conducted on a sample of companies from Bosnia and Herzegovina listed on the Banja Luka Stock Exchange. The initial sample consisted of all companies listed on the Banja Luka Stock Exchange. Securities on the Banja Luka Stock Exchange are included in the official stock market and the free market. At the time of data collection, there were 32 companies on the official stock market, while the securities of 422 companies were listed on the free market. This means that the initial sample consisted of 454 companies.

The research covered a five-year period, from 2017 to 2021. For 152 companies, audit reports were available for all five years under consideration. In the case of 302 companies, audit reports for one or more years were not available. It was decided to include the companies for which audit reports were available for three or more years of the period considered in the sample. The final sample included 187 listed companies, or 41.2% of the initial sample. A total of 886 audit reports were collected. The number of audit reports analyzed by year is given in Table 1.

Table 1 Distribution of audit reports analyzed in the period from 2017 to 2021

	2017	2018	2019	2020	2021
Number of companies in the sample	187	187	187	187	187
Number of missing audit reports	8	10	13	7	11
Number of available audit reports	179	177	174	180	176

Source: Author's calculation

Table 2 shows the structure of the sample according to the main industry sectors. As can be seen in Table 2, the sample includes companies from 17 different industry sectors. The most represented companies are those involved in water supply, sewerage, waste management and environmental remediation activities, followed by those that deal with the processing industry and those engaged in financial and insurance activities.

Table 2 The structure of the sample by main industry sectors

	Main industry section	No. of companies	% of companies
A	Agriculture, forestry and fishing	8	4.3
B	Ore and stone extraction	4	2.1
C	Processing industry	28	15.0
D	Production and supply of electricity, gas, steam and air conditioning	14	7.5
E	Water supply, sewerage, waste management and environmental remediation activities	43	23.0
F	Construction	8	4.3
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	14	7.5
H	Transportation and storage	11	5.9
I	Accommodation and food preparation and service activities	4	2.1
J	Information and communication	3	1.6
K	Financial and insurance activities	24	12.8
L	Real estate activities	11	5.9
M	Professional, scientific and technical activities	10	5.3
N	Administrative and support service activities	1	0.5
P	Education	1	0.5
Q	Health care and social work activities	1	0.5
R	Arts, entertainment and recreation	2	1.1
	Total	187	100.0

Source: Author's calculation

Table 3 shows the structure of the sample according to ownership. The sample includes both private and public enterprises, with a greater representation of private enterprises. Out of the total number of companies that make up the sample, 65.2% are privately owned, and 34.8% are publicly owned.

Table 3 The structure of the sample by ownership of companies

	No. of companies	% of companies
Private enterprises	122	65.2
Public enterprises	65	34.8
Total	187	100.0

Source: Author's calculation

All collected audit reports underwent content analysis. Content analysis was used according to Smith (2003). It is an approach to document and text analysis to quantify their content. This research method is commonly used in accounting research and has been employed by other authors investigating audit reporting (e.g. Carson et al., 2016; Vučković-Milutinović, 2019). Quantification is carried out within predefined categories and in a systematic and replicable manner. Content analysis was performed to determine whether the audit report contains a going concern audit opinion, which events and circumstances were cited as the reason for expressing doubt about the company's ability to continue as a going concern, and which audit firm performed the audit and issued an audit report.

4. Results

Table 4 shows the percentage share of going concern audit reports of listed companies in Bosnia and Herzegovina in the period from 2017 to 2021. Out of a total of 886 analyzed audit reports, 170 audit reports contain a going concern audit opinion. The average rate of going concern audit reports for the period from 2017 to 2021 is 19.2%. Observed by year, the rate of going concern audit reports ranges from 18.1% to 19.9%.

Table 4 Proportion of going concern audit reports in the period from 2017 to 2021

	2017	2018	2019	2020	2021	Total
Going concern audit reports	**19.0%**	**18.1%**	**19.0%**	**19.9%**	**19.9%**	**19.2%**
Unqualified audit opinion with MURGC paragraph	9.5%	10.2%	9.9%	9.4%	11.4%	10.1%
Modified audit opinion with MURGC paragraph	8.4%	6.8%	8.0%	9.4%	8.5%	8.2%
Modified audit opinion with modification on going concern	1.1%	1.1%	1.1%	1.1%	0.0%	0.9%
Other audit reports	**81.0%**	**81.9%**	**81.0%**	**80.1%**	**80.1%**	**80.8%**

Source: Author's calculation

When it comes to the structure of going concern audit reports, it can be noticed that in the majority of cases, auditors only emphasize the matter of material uncertainty related to the going concern. In less than 1% of cases, the existence of material uncertainty related to the going concern is a reason for modifying the audit opinion. On average, approximately 10% of audit reports contain an unqualified audit opinion with the MURGC paragraph, while approximately 8% of audit reports contain a modified audit opinion with the aforementioned paragraph. In the case of a modified audit opinion with the MURGC paragraph, the auditor draws the user's attention to the existence of significant doubt about the company's ability to continue as a going concern, while the reason for the modification of the audit opinion is something other than the existence of the aforementioned doubt.

Table 5 shows how often companies received a going concern audit opinion in the considered period. As can be seen in Table 5, 70% of companies never received a going concern audit opinion, while in almost 30% of companies, doubts about their ability to continue as a going concern were identified at least in one year. As many as 8% of companies received a going concern audit opinion four times in the period of five years, while in 5.3% of companies, material uncertainty related to going concern was identified in all five years.

Table 5 Repetition of the audit opinion on going concern in the period from 2017 to 2021

No. of received going concern audit reports	No. of companies	% of companies
Five	10	5.3
Four	15	8.0
Three	11	5,9
Two	7	3.7
One	13	7.0
Zero	131	70.1
Total	187	100.0

Source: Author's calculation

Out of the total number of companies that received a going concern audit opinion in 2017, 67.6% received the same opinion in 2018. In the following two years, this rate increased to 75%, while 80.6% of the total number of companies that received a going concern audit opinion in 2020 received the same opinion in 2021.

The question arises as to whether companies with identified events and circumstances that cast doubt on their ability to continue as a going concern manage to maintain business. This especially applies to companies that received a going concern audit opinion for several consecutive years. Based on the information contained in the Register of Business Entities, it was determined that all companies that received a going concern audit opinion once or more than once during the considered period are still operating.

Table 6 shows the circumstances that created significant doubt about the company's ability to continue as a going concern and led the auditors to issue a going concern audit opinion. Doubt about the company's ability to continue as a going concern can be generated by one or more events or circumstances.

Table 6 Circumstances which, individually or collectively, created significant doubts about the company's ability to continue as a going concern

Circumstances	No. of going concern auditor's reports	% of going concern auditor's reports
Net liability or net current liability position	135	79.4
The company made a loss	85	50.0
Loss is greater than equity	32	18.8
Adverse key financial ratios	22	12.9
The company does not perform its core activities or has significantly reduced the scope of its activities	18	10.6
Inability to pay creditors on due dates	16	9.4
The company achieved a negative operating cash flow	10	5.9

Source: Author's calculation

As can be seen in Table 6, the circumstances that led to doubts about the ability of companies to continue as a going concern are predominantly of financial nature. In the largest number of cases, almost 80%, it was about short-term liabilities being greater than current assets. This indicates existing or potential problems with the company's liquidity. The next most frequently mentioned circumstance is that the company made a loss. This is about the fact that the company made a loss in several consecutive periods and that the accumulated loss exceeds the amount of the variable part of own capital. In

almost 19% of cases, the fact that the accumulated loss exceeds the amount of own capital is cited as the reason for doubting the company's ability to continue as a going concern. This means that the company's liabilities are greater than its assets and that the company makes a loss at the expense of creditors. In the case of bankruptcy, the company will probably not be able to pay off all the obligations to the creditors. Other circumstances refer to negative key financial indicators, most often indicators related to liquidity, then partial or complete cessation of the main activity, inability to settle obligations on the due date, and the realization of a negative net operating cash flow.

Some believe that after receiving a going concern audit opinion, companies replace the auditor in search of an auditor who will give them a more favorable audit opinion. Table 7 shows the percentage of companies that replaced the auditor after receiving the going concern audit opinion, as well as the percentage of companies that did not receive the going concern audit opinion and replaced the auditor. In 17.5% of cases, companies replaced the auditor after receiving a going concern audit opinion, while in 16.4% of cases, companies replaced the auditor even though they did not receive a going concern audit opinion.

Table 7 The tendency of companies to replace the auditor after receiving a certain type of audit opinion

	% of companies that replaced the auditor	% of companies that did not replace the auditor
Companies that received a going concern audit opinion	17.5	82.5
Companies that did not receive a going concern audit opinion	16.4	83.6

Source: Author's calculation

If companies in Bosnia and Herzegovina were inclined to audit opinion shopping, the percentage of companies that replace the auditor after receiving a going concern audit report should be much higher than the percentage of companies that replace the auditor even though they did not receive this type of audit opinion. It can be seen in Table 7 that this percentage is only slightly higher. This is insufficient to claim that there is a tendency of companies toward audit opinion shopping.

5. Discussion and conclusion

This research shows that the average rate of going concern audit reports in Bosnia and Herzegovina for the period from 2017 to 2021 was 19.2%, and that this rate, observed by year, was stable even though the considered period included two pandemic years. In 2020 and 2021, a significant number of companies were affected by the coronavirus pandemic. Due to the measures implemented to protect against the coronavirus, some companies had to temporarily suspend operations or significantly reduce the scope of their activities. Furthermore, in some industries, there was a significant drop in demand for products and services. All this led to an increase in business risk and business uncertainty.

Previous research has shown that auditors are more likely to express a going concern audit opinion in a period of increased economic uncertainty (Cheffers et al., 2010; Xu et al., 2011), while this research shows that an increase in economic uncertainty leads to only a slight increase in the rate of going concern audit reports.

This research also showed that the majority of going concern audit reports contain an unqualified or modified audit opinion emphasizing the matter of material uncertainty regarding going concern, while in an extremely small number of audit reports the existence of material uncertainty regarding going concern was the reason to modify the audit opinion. Based on this, it can be concluded that in the majority of cases, companies have correctly reported on the existence of doubts about the company's ability to continue as a going concern and the events and circumstances that generate this doubt. However, based on the information available in the audit reports and the company's financial statements, it cannot be determined whether the companies included the required disclosures in their financial reports on their own initiative or did so at the auditor's request to avoid the auditor modifying the audit opinion. Circumstances that most often created significant doubts about the company's

ability to continue as a going concern refer to the fact that the company's short-term liabilities are greater than its current assets, that the company makes a loss, and that the loss exceeds the amount of own capital.

Carson et al. (2016) determined that a large number of companies that receive a going concern audit opinion in one year receive the same opinion in the following year. This research reached the same conclusion. The rate of repetition of the going concern audit report in the considered period ranged from 67.6% to 80.6%. Out of the total number of companies included in the sample, in 13.3% of the companies, material uncertainty related to the going concern was identified in four or all five years. Despite this, all companies continue to operate, which means that the going concern audit opinion cannot serve as a predictor of company bankruptcy.

Although Chung et al. (2019) found that companies that receive a going concern audit opinion in one year replace their auditor in order to avoid receiving the same audit opinion in the following year and thus successfully implement the so-called audit opinion shopping, no evidence for this was found in this research. The reasons for replacing the auditor can be different. Companies that receive an unqualified audit opinion also replace the auditor, for objective or subjective reasons. The percentage of companies that received a going concern audit opinion and replaced the auditor is only slightly higher than the percentage of companies that did not receive this type of audit opinion, but still chose to replace the auditor. This is not enough to claim that companies replace the auditor to obtain a more favorable audit opinion. There is a possibility that companies in Bosnia and Herzegovina will not bear the negative consequences of obtaining a going concern audit opinion and therefore do not have to engage in audit opinion shopping.

When interpreting the results of this research, the limitations that existed in the research should be taken into account. In order to increase the sample, companies for which audit reports were not available for all years covered by the considered period were included in the sample. It is possible that companies did not submit audit reports containing an unfavorable audit opinion, including an audit opinion on going concern, to the Banja Luka Stock Exchange. The problem of missing data was particularly pronounced when determining whether the auditor had been replaced. The lack of an audit report for one year leads to the impossibility of determining whether there was a replacement of auditors in two consecutive years. Furthermore, when interpreting the research results, it should be borne in mind that issuing a going concern audit opinion does not depend only on the existence of circumstances that create significant doubt about the company's ability to continue as a going concern, but also on the quality of the audit. Namely, the issuance of a going concern audit opinion is influenced by the auditor's ability to identify circumstances that create significant doubt about the company's ability to continue as a going concern, as well as their willingness to request the disclosure of information about the existence of these circumstances or to modify the auditor's opinion if the audit client's management refuses to include the said disclosures in the financial statements. On the one hand, managers do not like to include unfavorable information in financial statements, and on the other hand, auditors may abandon their disclosure requirements or their intention to modify audit opinions in order not to lose a client.

Future research could focus on examining how users of financial statements react to the issuance of a going concern audit opinion and whether this type of audit opinion affects the decisions they make.

References

1. Carcello, J. V. & Neal, T. L. (2003). Audit committee characteristics and auditor dismissals following "new" going-concern reports. *The Accounting Review*, *78*(1), 95-117. https://doi.org/10.2308/accr.2003.78.1.95

2. Carson, E., Fargher, N. & Zhang, Y. (2016). Trends in auditor reporting in Australia: A synthesis and opportunities for research. *Australian Accounting Review*, *26*(3), 226-242. https://doi.org/10.1111/auar.12124

3. Cheffers, M., Whalen, D. & Thrun, M. (2010). *Going Concern Review*. Audit Analytics. https://www.auditanalytics.com/doc/AuditAnalytics_GoingConcernRpt_04-10.pdf

4. Chen, P. F., He, S., Ma, Z. & Stice, D. (2016). The information role of audit opinions in debt contracting. *Journal of Accounting and Economics*, *61*(1), 121-144. https://doi.org/10.1016/j.jacceco.2015.04.002

5. Chung, H., Sonu, C. H., Zang, Y. & Choi, J. H. (2019). Opinion shopping to avoid a going concern audit opinion and subsequent audit quality. *Auditing: A Journal of Practice & Theory*, *38*(2), 101-123. https://doi.org/10.2308/ajpt-52154

6. Geiger, M. A. & Kumas, A. (2018). Anticipation and reaction to going-concern modified audit opinions by sophisticated investors. *International Journal of Auditing*, *22*(3), 522-535. https://doi.org/10.1111/ijau.12135

7. IAASB (2015). *International Standards on Auditing 570 (revised): Going Concern*, International Auditing and Assurance Standards Board, London.

8. IASB (2003). *International Accounting Standard 1: Presentation of Financial Statements*, International Accounting Standard Board, London.

9. Krishnan, J. (1994). Auditor switching and conservatism. *Accounting Review*, *69*(1), 200-215.

10. Krishnan, J. & Stephens, R. G. (1995). Evidence on opinion shopping from audit opinion conservatism. *Journal of Accounting and Public Policy*, *14*(3), 179-201. https://doi.org/10.1016/0278-4254(95)00020-F

11. Lennox, C. (2000). Do companies successfully engage in opinion-shopping? Evidence from the UK. *Journal of Accounting and Economics*, *29*(3), 321-337. https://doi.org/10.1016/S0165-4101(00)00025-2

12. Menon, K. & Williams, D. D. (2010). Investor reaction to going concern audit reports. *The Accounting Review*, *85*(6), 2075-2105. https://doi.org/10.2308/accr.2010.85.6.2075

13. Simamora, R. A. & Hendarjatno, H. (2019). The effects of audit client tenure, audit lag, opinion shopping, liquidity ratio, and leverage to the going concern audit opinion. *Asian Journal of Accounting Research*, *4*(1), 145-156. https://doi.org/10.1108/AJAR-05-2019-0038

14. Smith, M. (2003). *Research Methods in Accounting*. Sage Publications. https://doi.org/10.4135/9781849209809

15. Vučković-Milutinović, S. (2019). Analysis of modifications to auditor's opinion on financial statements of listed companies in Serbia. *Ekonomika preduzeća*, *67*(3-4), 212-223. https://doi.org/10.5937/EKOPRE1904212V

16. Xu, Y., Jiang, A. L., Fargher, N. & Carson, E. (2011). Audit reports in Australia during the global financial crisis. *Australian Accounting Review*, *21*(1), 22-31. https://doi.org/10.1111/j.1835-2561.2010.00118.x

Dunja Meštrović
University of Rijeka
Centre for Quality Assurance and
Institutional Research
51000 Rijeka, Croatia
dmestrovic@uniri.hr

Lidija Bagarić
University of Rijeka
Faculty of Tourism and Hospitality
Management in Opatija
51410 Opatija, Croatia
lidijab@fthm.hr

Marija Ham
Josip Juraj Strossmayer
University of Osijek
Faculty of Economics
and Business in Osijek
31000 Osijek, Croatia
marija.ham@efos.hr

EMBRACING SUSTAINABLE MARKETING IN HIGHER EDUCATION VIEWED THROUGH SERVICE QUALITY: STUDENTS' PERSPECTIVE

Abstract

Purpose: This paper explores the application of sustainable marketing practices in higher education institutions, focusing in particular on the perspectives and experiences of students as direct users of the service. Given that sustainability has become an important issue globally, integration of sustainable principles into the marketing strategies of educational institutions becomes imperative. It was found that the existing academic literature contains only a very limited number of studies in the area of applying sustainable marketing in public higher education institutions, despite its recognized importance. This study aims to help close this gap.

Methodology: Using a quantitative research methodology, including questionnaires, this study examines student perceptions and attitudes towards the application of sustainable marketing in higher education, focusing on the quality of services. A structural model was tested, the reliability and validity of which had been confirmed previously. Partial least squares structural equation modeling (PLS-SEM) was used to examine the relationships between the constructs.

Results: All direct relationships in the structural model were found to be statistically significant and positive, confirming the hypothesis that sustainable marketing of a higher education institution has a positive impact on the performance of the higher education institution when considering the quality of services.

Conclusion: By highlighting the importance of sustainable marketing in public higher education institutions, this research contributes to the growing sustainability discourse in the academic community and provides actionable insights for educational institutions seeking to improve their marketing strategies in line with environmental and social responsibility goals.

Keywords: Sustainability, sustainable marketing, higher education institutions, students, performance

1. Introduction

In recent years, the sustainability imperative has permeated various sectors, including higher education. The importance of sustainable marketing in higher education is undisputed. As educational institutions that have an impact on society, universities are responsible for promoting sustainability in all areas of their activities and in society in general. Despite the increasing importance placed on

sustainability in various sectors, there is a notable gap in the academic literature regarding the application of sustainable marketing practices in public higher education institutions. Previous research has been recognized mainly in the field of sustainability in education (Casarejos et al., 2017; Aleixo et al., 2018; Fuchs et al., 2020; Zhou et al., 2020; Budihardjo et al., 2021) and the efficiency and quality of higher education institutions (Barilović et al., 2013; Štimac, 2013; Degtjarjova et al., 2018; Arbula Blecich, 2024). Although the importance of sustainable marketing in higher education is recognized in theory, there are few empirical studies examining its implementation and impact on institutional performance.

Therefore, the research problem addressed in this paper focuses on investigating student perceptions and attitudes towards sustainable marketing initiatives in public higher education institutions, with a focus on service quality. By investigating this research problem, the study aims to fill a gap in the literature and provide empirical evidence of the positive impact of sustainable marketing on the performance of higher education institutions from a service quality perspective. A central focus of this study is to investigate the relationship between sustainable marketing and the performance of HEIs through service quality from the perspective of students as direct users of the service. Through the use of a structural model tested for reliability and validity, this research aims to provide empirical evidence of the positive impact of sustainable marketing practices on HEI performance metrics, particularly in relation to service quality. Students represent an important stakeholder group whose perceptions and attitudes towards sustainable marketing initiatives can provide valuable insights into their effectiveness and impact on the overall success of higher education institutions.

This paper contributes to the academic literature by expanding the understanding of sustainable marketing in higher education and providing practical implications for higher education institutions seeking to improve their marketing strategies in line with sustainability principles. The study highlights the links between sustainable marketing, student perceptions and the performance of higher education institutions, and aims to promote positive change towards a more sustainable future in public higher education institutions. The paper continues with a theoretical background of the research topic, the research methodology and data analysis, the results and discussion, and finally the conclusions and limitations.

2. Theoretical background

Sustainable marketing in higher education encompasses a wide range of activities to promote sustainability principles, practices and initiatives in academic institutions. This approach recognizes the unique position of higher education institutions as catalysts for social change and innovation that can influence not only their own activities, but also broader societal attitudes and behaviors towards sustainability. Sustainable marketing in higher education recognizes the central role of students as key stakeholders in driving sustainability initiatives and fostering a culture of sustainability on campus.

2.1 Sustainable marketing in higher education

Sustainable marketing in higher education is a strategic approach to promoting sustainability principles and practices in academic institutions with the aim of fostering a more environmentally conscious, socially responsible and economically sustainable future for all stakeholders (Fuchs et al., 2020; Budihardjo et al., 2021; Meštrović et al., 2021). The concept of sustainability education is based on the idea that educational institutions play a crucial role in promoting sustainable development and providing students with the knowledge and skills needed to tackle social, environmental and economic challenges (Hübscher et al., 2022). Higher education institutions use various strategies to engage stakeholders and demonstrate their commitment to sustainability (Shawe et al., 2019). These include integrating sustainability into academic curricula across all disciplines, fostering interdisciplinary research collaborations on sustainability-related topics, and implementing sustainability-oriented initiatives in campus operations and facilities management (Annan-Diab & Molinari, 2017).

The adoption of sustainable marketing practices in higher education reflects a broader trend towards corporate social responsibility (Arena et al., 2018). Higher education institutions are increasingly integrating sustainability considerations into their strategic plans, governance structures and operational activities to address environmental challenges and meet stakeholder expectations (Aleixo et al., 2018). Sustainable marketing in higher education has gained importance in response to global environmental challenges and changing societal expectations. As institutions of learning and societal impact, higher education institutions have a unique opportunity and responsibility to promote the principles of sustainability (Casarejos et al., 2017; Zhou et al., 2020).

2.2 Students as stakeholders

The theoretical background of students as primary stakeholders and direct service users in higher education institutions is based on stakeholder theory, the concept of students as customers and the recognition of students as direct service users (Sandmaung & Ba Khang, 2013). Students are viewed as primary stakeholders and direct service users in higher education institutions, and their needs and preferences should be considered in the design and implementation of sustainable initiatives and programs (Pedro et al., 2020). Higher education institutions are expected to put student interests first and design their programs, services and policies to meet their diverse needs and expectations (Jongbloed et al., 2008).

Students play a diverse role in shaping institutional policy, practice and culture. As consumers of educational services, students influence enrollment decisions, tuition revenue, and institutional reputation (Degtjarjova et al., 2018). Consequently, institutions are expected to adopt a customer-centric approach and focus on providing high-quality services that meet or exceed student expectations (Širola & Mihaljević, 2016). This theoretical view emphasizes the importance of understanding students' needs, preferences and satisfaction and using this information to continuously improve the quality of educational services (Degtjarjova et al., 2018; Meštrović et al., 2021). Research shows that students are increasingly concerned with sustainability issues and expect their universities to take a leadership role in this area (Lozano et al., 2013). They place importance on environmental responsibility, social justice and ethical leadership when evaluating higher education institutions. Therefore, students' perceptions, attitudes and behaviors regarding sustainability have a significant impact on institutional decision-making and strategic planning (Degtjarjova et al., 2018; Pedro et al., 2020).

3. Methodology

This section outlines the methodological framework of the empirical study, which is grounded in an extensive literature review and prior research on the relationship between sustainable marketing—operationalized as sustainable marketing orientation—and performance, measured by service quality in the context of higher education.

The academic marketing literature lacks consensus on a widely accepted approach to measuring higher education (HE) performance. However, Pearce et al. (1987) suggest that subjective measures of public organization performance, based on respondents' assessments, can be interpreted as reliable and equivalent to objective indicators. In line with this, students' subjective assessments were chosen as a non-financial performance indicator, following the approach of Teeroovengadum et al. (2016) and Ruiz de Sabando et al. (2018) for evaluating a university's service quality.

To achieve empirical objectives of the study, exploratory primary research was conducted, building on the literature review. This approach is well-suited for exploring the attitudes and perceptions of key stakeholders regarding sustainable marketing in higher education institutions. Based on this, the authors proposed a conceptual model illustrated in Figure 1, that consists of two constructs: (1) sustainable marketing in higher education (SMHE), defined as a multidimensional construct, and (2) performance, assessed as service quality (SQ) of a higher education institution (HEI), treated as a unidimensional construct.

Figure 1 Conceptual model
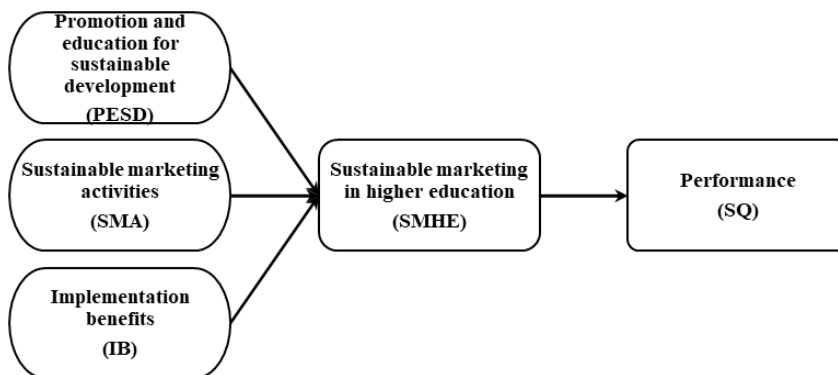
Source: Authors

The primary quantitative empirical research was conducted using a measurement instrument developed by Meštrović et al. (2021) to capture the attitudes and perceptions of key stakeholders in higher education—specifically students as direct service users—towards sustainable marketing. The instrument for measuring service quality in higher education was adapted from the SERVQUAL scale, previously tested by Leko Šimić and Štimac (2013) and Barilović et al. (2013). Both instruments utilized a 7-point Likert scale. The research took place between 10 and 30 May 2019, using an anonymous online questionnaire via Google Forms, with data collected from a convenience sample, supplemented by the snowballing technique.

Univariate and multivariate statistical analyses were performed using the Statistical Package for Social Sciences (SPSS 24.0) to analyze and organize the collected data. Descriptive statistical analysis was employed to characterize the research sample. To examine the relationships between the constructs, partial least squares structural equation modeling (PLS-SEM) was applied using SmartPLS 3.3.3. PLS-SEM, a multivariate method widely used in the social sciences (Purwanto & Sudargini, 2021), which integrates factor and regression analysis and does not require the assumption of data normality (Hair et al., 2013).

4. Participants and procedure

A total of 1,663 students, as direct users of college services, completed the questionnaire across 15 components (i.e., faculties) of the University of Rijeka (UNIRI). The sample comprised 1,089 women (66.03%) and 565 men (33.97%), with an average age of 25 years. Table 1 presents the distribution of respondents based on the faculty and degree program they are enrolled in.

Table 1 Sample - Respondents' faculty and study program level (n = 1,663)

UNIRI Component	Graduate University Study and Specialist Graduate Professional Study - ISCED 7	Integrated Undergraduate and Graduate University Study - ISCED 7	Undergraduate University or Professional Study (min. 180 ECTS) - ISCED 6	Undergraduate Professional Study (less than 180 ECTS) - ISCED 5	Total Sample Share
AARTS	7 (14.90%)	0 (0%)	40 (85.10%)	0 (0%)	47 (2.80%)
BIOTECH	17 (36.20%)	0 (0%)	30 (63.80%)	0 (0%)	47 (2.80%)
CIVILEN	47 (43.90%)	0 (0%)	54 (50.50%)	6 (5.60%)	107 (6.40%)
ECON	1 (33.30%)	0 (0%)	2 (66.70%)	0 (0%)	3 (0.20%)
ENGINEER	111 (48.10%)	0 (0%)	120 (51.90%)	0 (0%)	231 (13.90%)
HEALTH	39 (19.50%)	0 (0%)	161 (80.50%)	0 (0%)	200 (12.00%)
HUMAN	114 (57.60%)	0 (0%)	84 (42.40%)	0 (0%)	198 (11.90%)
INFDT	31 (39.70%)	0 (0%)	47 (60.30%)	0 (0%)	78 (4.70%)
LAW	11 (8.10%)	90 (66.20%)	24 (17.60%)	11 (8.10%)	136 (8.20%)
MARIT	76 (39.60%)	0 (0%)	116 (60.40%)	0 (0%)	192 (11.50%)
MATH	8 (24.20%)	0 (0%)	25 (75.80%)	0 (0%)	33 (2.00%)
MEDIC	5 (1.80%)	211 (76.20%)	61 (22.00%)	0 (0%)	277 (16.70%)
PHYS	12 (35.30%)	1 (2.90%)	21 (61.80%)	0 (0%)	34 (2.00%)
POLY	0 (0%)	1 (3.80%)	25 (96.20%)	0 (0%)	26 (1.60%)
TEACH	14 (25.90%)	13 (24.10%)	27 (50.00%)	0 (0%)	54 (3.20%)
Total	493 (29.60%)	316 (19.00%)	837 (50.30%)	17 (1.00%)	1,663 (100%)

Note: AARTS = Academy of Applied Arts, BIOTECH = Faculty of Biotechnology and Drug Development, CIVILEN = Faculty of Civil Engineering, ECON = Faculty of Economics and Business, ENGINEER = Faculty of Engineering, HEALTH = Faculty of Health Studies, HUMAN = Faculty of Humanities and Social Sciences, INFDT = Faculty of Informatics and Digital Technologies, LAW = Faculty of Law, MARIT = Faculty of Maritime Studies, MATH = Faculty of Mathematics, MEDIC = Faculty of Medicine, PHYS = Faculty of Physics, POLY = University of Rijeka, Department of Polytechnics, TEACH = Faculty of Teacher Education.
Source: Authors

Considering the characteristics of the proposed measurement model, the minimum required sample size was calculated using the G*Power 3.1.9.7 application (Faul et al., 2009), based on the guidelines of Hair et al. (2013) and Memon et al. (2020). With the desired statistical power ($1-\beta = 0.95$), effect size ($f^2 = 0.15$), and statistical significance ($\alpha = 0.05$) typical in social science research, the minimum required sample size was determined to be 119 respondents. Kline (2016) notes that for structural equation modeling, a sample of 100 is small but acceptable for simple models, a sample between 100 and 200 is considered medium, and a sample of 200 or more is regarded as large. Therefore, the sample of 1,663 respondents was deemed more than adequate for PLS-SEM analysis.

Table 2 presents the results of descriptive statistical analysis of the Sustainable Marketing in Higher Education (SMHE) construct, which is measured by three dimensions: (1) Sustainable Marketing Activities (SMA), (2) Promotion and Education for Sustainable Development (PESD), and (3) Implementation Benefits (IB). The IB dimension received the high-est mean score ($\bar{x} = 5.64$, $\sigma = 1.109$), while the SMA dimension had the lowest mean score ($\bar{x} = 5.15$, $\sigma = 1.001$). The PESD dimension was rated with a mean score of 5.48 ($\sigma = 1.211$).

Within the IB dimension, the highest rated item was IB7, "Increasing the visibility of the higher education institution" ($\bar{x} = 5.84$, $\sigma = 1.319$), while the lowest rated item was IB3, "Creating added value for users while considering the long-term interests of society and the environment" ($\bar{x} = 5.42$, $\sigma = 1.419$). For the PESD dimension, the highest rated item was PESD1, "Improving the overall quality of the higher education system" ($\bar{x} = 5.82$, $\sigma = 1.282$), while the lowest rated was PESD3, "Implementation of study programs in sustainable development" ($\bar{x} = 5.18$, $\sigma = 1.603$).

In the SMA dimension, the highest rated item was SMA5, "Partnership with the local community" ($\bar{x} = 5.51$, $\sigma = 1.213$), while the lowest rated was SMA9, "Regularly considering the impact of business decisions on various stakeholders (e.g., employers, students, parents, employees, higher education institutions), natural and financial resources, and society at large" ($\bar{x} = 4.52$, $\sigma = 1.441$).

Table 2 Descriptive statistical analysis of the SMHE construct

Code	Item	Mean	SD
Sustainable Marketing Activities (SMA)			
SMA1	Adjusting business processes to laws and legal regulations while striving to achieve own business goals	4.64	1.327
SMA2	Concern about environmental and societal long-term benefits while striving to achieve own business goals	5.18	1.264
SMA3	Dialogue with key stakeholders (e.g., employers, students, prospective students, parents, employees, institutions, ministries, society at large)	5.20	1.303
SMA4	Partnership with regional and local government bodies	5.22	1.239
SMA5	Partnership with the local community	5.51	1.213
SMA6	Partnership with economic entities	5.17	1.244
SMA7	Partnership with competitors	5.10	1.461
SMA8	Concern about all employees while striving to achieve own business goals	5.27	1.281
SMA9	Considering impacts of business decisions on stakeholders (e.g., students, employees, institutions, society at large)	4.52	1.441
SMA10	Increasing the application of modern ICT in business processes and teaching methods	5.41	1.286
SMA11	Increasing availability of formal, informal, and non-formal education to all stakeholders	4.94	1.436

Code	Item	Mean	SD
SMA12	Anticipating and respecting the needs of the broader community and future generations	5.32	1.263
SMA13	Transparency and availability of data on contributions to society and efforts to reduce environmental impact	5.11	1.371
SMA14	Acceptance and application of sustainable development principles as part of business culture	5.27	1.319
SMA15	Promoting new ideas that contribute to sustainability as a lifestyle and business philosophy	5.35	1.327
Implementation Benefits (IB)			
IB1	Creating societal change	5.60	1.289
IB2	Rationalizing resource usage	5.68	1.312
IB3	Creating added value for users while considering long-term societal and environmental interests	5.42	1.419
IB4	Adapting and/or creating new study programs	5.70	1.253
IB5	Increasing loyalty and satisfaction of users and stakeholders	5.51	1.414
IB6	Increasing study success	5.57	1.352
IB7	Increasing the visibility of the higher education institution	5.84	1.319
IB8	Intensifying internal and external mobility of students and employees	5.79	1.282
IB9	Increasing ethics, morality, and transparency in procurement and donations	5.80	1.248
IB10	Education for sustainable development	5.54	1.531
IB11	Creating benefits for all stakeholders and wider society	5.68	1.403
IB12	Achieving competitive advantage	5.56	1.467
IB13	Achieving environmental, societal, and economic goals simultaneously	5.68	1.342
Promotion and Education for Sustainable Development (PESD)			
PESD1	Improving the overall quality of the higher education system	5.82	1.282
PESD2	Enhancing continuous professional development and training of higher education employees	5.71	1.361
PESD3	Implementing study programs in sustainable development	5.18	1.603
PESD4	Offering elective courses in sustainable development	5.37	1.412
PESD5	Reporting efforts and achievements related to sustainable development	5.29	1.405
PESD6	Promoting sustainable development principles through business practices	5.53	1.424

Note: SD = standard deviation.
Source: Authors

The descriptive statistical analysis results of the Service Quality (SQ) construct are shown in Table 3.

Table 3 Descriptive statistical analysis of the SQ construct

Code	Item	Mean	SD
Service Quality (SQ)			
SQ1	Organization of studies	3.71	1.717
SQ2	Acquired knowledge meets expectations	3.90	1.751
SQ3	Usefulness of the acquired knowledge	4.52	1.774
SQ4	Social necessity of the acquired knowledge	4.84	1.649
SQ5	Quality of the study program	3.86	1.898
SQ6	The knowledge and skills acquired during the study program will help me find a new job or get promotion in my current position	4.86	1.399
SQ7	The skills acquired during the study program will significantly impact the efficiency and success of my future work	5.28	1.724
SQ8	The study program enables personal development	4.19	1.815
SQ9	The available resources (space, facilities, literature, technical equipment) meet student needs	4.57	1.935
SQ10	Performance/efficiency of the teaching staff	4.53	1.960
SQ11	Performance/efficiency of the administrative staff	4.53	2.096
SQ12	Teaching staff knowledge, experience, and skills	3.72	2.000
SQ13	Administrative staff knowledge, experience, and skills	4.24	1.768
SQ14	Courtesy of the teaching staff	4.78	1.583
SQ15	Courtesy of the administrative staff	4.44	1.551
SQ16	Reputation/credibility of the teaching staff	4.52	1.972
SQ17	Reasonable working hours	5.36	1.563
SQ18	Convenient class schedule	3.37	1.983
SQ19	Reputation of the study program in the local community	4.90	1.946

Note: SD = standard deviation.
Source: Authors

As presented in Table 3, the highest rated item within the SQ construct was SQ17, "Reasonable working hours" ($\bar{x}$ = 5.36, σ = 1.563), followed by SQ7, "The skills acquired during the study program will significantly impact the efficiency and success of my future work" ($\bar{x}$ = 5.28, σ = 1.724). On the other hand, the items with the lowest mean scores were SQ1, "Organization of studies" ($\bar{x}$ = 3.71, σ = 1.717), and SQ18, "Convenient class schedule" ($\bar{x}$ = 3.37, σ = 1.983).

The calculated Cronbach's alpha values for the measurement scales were as follows: 0.938 for SMA, 0.958 for IB, 0.926 for PESD, and 0.939 for the SQ construct. Additionally, all calculated correlations with the associated constructs exceeded the threshold of 0.30 (Hair et al., 2013), confirming the reliability of the measurement scales and the internal consistency of the constructs.
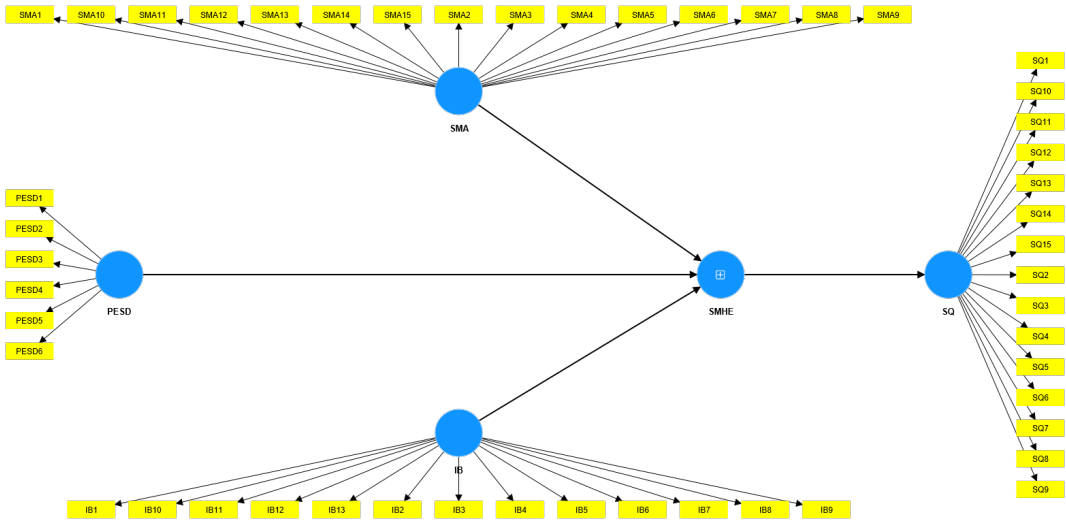
5. Results

Following the recommendations of Hair et al. (2013), Confirmatory Factor Analysis (CFA) was conducted before developing the structural model SMHE-SQ, as shown in Figure 2. This model outlines the relationship between SMHE and SQ, with SMHE defined as an exogenous higher-order construct composed of three lower-order constructs: (1) SMA, (2) PESD, and (3) IB, and SQ as the endogenous latent construct. The SMHE higher-order construct was modeled as a reflective-formative Type II construct (Jarvis et al., 2003; Sarstedt et al., 2019).

In line with the recommendations of Sarstedt et al. (2019) and Hair et al. (2020), the evaluation of the higher-order constructs was performed using the PLS-SEM method with a two-stage approach. This method allows for different numbers of indicators across lower-order constructs (Becker et al., 2012). The first stage involved assessing the reflective measurement model of the lower-order constructs, followed by the evaluation of the structural model.

Figure 2 Structural and measurement model SMHE-SQ

Source: Authors

This section of the paper presents the results of both the outer (measurement) model evaluation and the inner (structural) model evaluation. The higher-order exogenous construct, SMHE (reflective-formative Type II), comprising three lower-order reflective constructs (SMA, PESD, IB), and the endogenous construct, SQ, determined by reflective variables, were analyzed using the PLS-SEM method. Following Hair et al. (2013), which recommends this approach for models with higher-order constructs, the path weighting scheme was applied with the "A" setting for formative construct indicator weighting. This approach allows for a maximum of 300 iterations with a stop criterion of 10^{-7} (Becker et al., 2012).

The evaluation of the reflective measurement model included assessing indicator reliability, internal consistency, as well as discriminant and convergent validity, following the guidelines of Sarstedt et al. (2019) and Hair et al. (2020). All indicators in the SMHE-SQ model had standardized factor loadings above the 0.70 threshold (Hair et al., 2019), except for SMA1, SMA15, IB10, IB12, and PESD1, which were removed. However, the removal of SMA3, SMA10, SQ6, and SQ7, which also had loadings below 0.70, did not result in the expected improvement in composite reliability (CR) and average variance extracted (AVE) values, so they were retained for further analysis.

Table 4 Reflective measurement model reliability, internal consistency and validity analysis

Construct	Item	Standardized factor loading	Cronbach α	Rho_A	CR	AVE
SMA	SMA10	0.673	0.940	0.943	0.947	0.581
	SMA11	0.729				
	SMA12	0.832				
	SMA13	0.808				
	SMA14	0.768				
	SMA2	0.799				
	SMA3	0.669				
	SMA4	0.809				
	SMA5	0.772				
	SMA6	0.806				
	SMA7	0.719				
	SMA8	0.776				
	SMA9	0.725				
IB	IB1	0.760	0.921	0.923	0.939	0.718
	IB11	0.823				
	IB13	0.886				
	IB2	0.782				
	IB3	0.805				
	IB4	0.820				
	IB5	0.849				
	IB6	0.800				
	IB7	0.860				
	IB8	0.812				
	IB9	0.835				
PESD	PESD2	0.856	0.918	0.927	0.939	0.755
	PESD3	0.774				
	PESD4	0.886				
	PESD5	0.920				
	PESD6	0.901				
SQ	SQ1	0.752	0.946	0.950	0.952	0.573
	SQ10	0.807				
	SQ11	0.791				
	SQ12	0.653				
	SQ13	0.695				
	SQ14	0.782				
	SQ15	0.757				
	SQ2	0.830				
	SQ3	0.803				
	SQ4	0.689				
	SQ5	0.815				
	SQ6	0.641				
	SQ7	0.640				
	SQ8	0.814				
	SQ9	0.797				

Note: Removed items SMA1, SMA15, IB10, IB12, PESD1.
Source: Authors

As shown in Table 4, the standardized factor loadings for all variables in the reflective SMHE-SQ model ranged from 0.640 to 0.920, while the Cronbach's alpha coefficients, ranging from 0.918 to 0.946, indicated a high level of reliability for the latent construct measures. The Rho_A coefficients, ranging from 0.923 to 0.950, further confirmed the stability and consistency of the model. Additionally, Composite Reliability (CR) values between 0.939 and 0.952 verified the reliability of the indicators, while Average Variance Extracted (AVE) values, all exceeding the 0.50 threshold (ranging from 0.573 to 0.755), confirmed the internal consistency and convergent validity of the SMHE-SQ model. These findings satisfied the necessary criteria for both validity and reliability across all constructs in the outer measurement model (Hair et al., 2019).

Furthermore, discriminant validity of the measurement model was assessed following Henseler et al. (2015) and Hair et al. (2020), using three methods: (1) standardized cross-loadings, (2) the Fornell-Larcker criterion, and (3) the Heterotrait-Monotrait ratio (HTMT). The first criterion for discriminant validity was satisfied as all outer factor cross-loadings were greater than their loadings on other indicators, as highlighted in bold in Table 5.

Table 5 Standardized cross-loadings of items of the reflective measurement model

Item/Construct	SMA	IB	PESD	SQ
SMA10	0.673	0.435	0.276	0.250
SMA11	0.729	0.336	0.264	0.265
SMA12	0.832	0.543	0.389	0.224
SMA13	0.808	0.597	0.456	0.272
SMA14	0.768	0.470	0.277	0.127
SMA2	0.799	0.624	0.568	0.276
SMA3	0.669	0.522	0.393	0.204
SMA4	0.809	0.627	0.613	0.345
SMA5	0.772	0.565	0.496	0.209
SMA6	0.806	0.582	0.620	0.379
SMA7	0.719	0.364	0.239	0.158
SMA8	0.776	0.437	0.355	0.276
SMA9	0.725	0.412	0.372	0.260
IB1	0.671	0.760	0.563	0.246
IB11	0.492	0.823	0.777	0.426
IB13	0.603	0.886	0.812	0.340
IB2	0.490	0.782	0.593	0.219
IB3	0.512	0.805	0.619	0.254
IB4	0.580	0.820	0.614	0.230
IB5	0.577	0.849	0.708	0.337
IB6	0.508	0.800	0.614	0.146
IB7	0.550	0.860	0.730	0.377
IB8	0.520	0.812	0.623	0.133

Item/Construct	SMA	IB	PESD	SQ
IB9	0.547	0.835	0.659	0.216
PESD2	0.528	0.812	0.856	0.431
PESD3	0.377	0.546	0.774	0.190
PESD4	0.449	0.609	0.886	0.380
PESD5	0.513	0.743	0.920	0.327
PESD6	0.525	0.781	0.901	0.458
SQ1	0.273	0.129	0.264	0.752
SQ10	0.207	0.202	0.336	0.807
SQ11	0.151	0.246	0.421	0.791
SQ12	0.147	0.022	0.142	0.653
SQ13	0.262	0.171	0.265	0.695
SQ14	0.309	0.334	0.359	0.782
SQ15	0.348	0.332	0.339	0.757
SQ2	0.256	0.147	0.290	0.830
SQ3	0.174	0.273	0.408	0.803
SQ4	0.308	0.353	0.278	0.689
SQ5	0.204	0.103	0.215	0.815
SQ6	0.258	0.212	0.167	0.641
SQ7	0.258	0.348	0.379	0.640
SQ8	0.210	0.194	0.277	0.814
SQ9	0.199	0.221	0.374	0.797

Source: Authors

The second discriminant validity criterion – Fornell-Larcker criterion, was also confirmed, as shown in Table 6, since all the dimensions' calculated square root AVE values are greater than its correlations with other model's dimensions.

Table 6 Reflective measurement model discriminant validity - Fornell-Larcker criterion

	SMA	IB	SQ	PESD
SMA	0.762			
IB	0.632	0.847		
SQ	0.328	0.300	0.757	
PESD	0.553	0.766	0.417	0.869

Source: Authors

As shown in Table 7, all calculated HTMT values were lower than the threshold of 0.90, thus confirming the third discriminant validity criterion according to Henseler et al. (2015).

Table 7 Reflective measurement model discriminant validity - HTMT ratio of correlations

	SMA	IB	SQ	PESD
SMA				
IB	0.670			
SQ	0.330	0.305		
PESD	0.574	0.820	0.420	

Source: Authors

Finally, all calculated reflective model's path coefficients were found to be statistically significant and positive, while SMA had the highest calculated statistically positive relationship with SMHE (β = 0.565), as shown in Table 8, while the effects of lower construct weight values on higher value construct SMHE are shown in Table 9.

Table 8 Reflective measurement model path coefficients

	Path coefficient	SD	t	p
SMA → SMHE	0.565	0.008	72.162	0.000
IB → SMHE	0.307	0.005	61.441	0.000
PESD → SMHE	0.269	0.006	41.801	0.000

Source: Authors

Table 9 Reflective measurement model lower construct weight value effects on the higher value construct

Higher-order construct	Lower-order construct	Weight value	t-value	p-value
SMHE	SMA	0.352	31.214	0.000
	IB	0.329	30.115	0.000
	PESD	0.459	33.932	0.000

Source: Authors

The calculated weight values for all lower-order constructs were statistically significant at the 1% level (p < 0.01), thereby confirming the third criterion for discriminant validity. This indicates that SMHE in the proposed conceptual model can be composed of three dimensions (SMA, IB, and PESD), with PESD being the most significant, with a weight value of 0.459. Discriminant validity of the higher-order formative construct was assessed through the correlation between the constructs, with the calculated correlation value of 0.406, well below the threshold of 0.7, confirming its discriminant validity.

The evaluation of the inner (formative) structural model involved assessing the relationships between the constructs and the model predictive capability (Hair et al., 2013). This included the following: (1) estimating indicator collinearity by checking the variance inflation factors (VIF) of the latent constructs, (2) testing the significance and relevance of structural model relationships by calculating path coefficients, which represent the hypothesized relationships, using a nonparametric resampling method, (3) assessing predictive significance through the coefficient of determination (R^2), evaluating effect size using the coefficient of influence (f^2), and calculating the blindfolding-based cross-validated redundancy measure (Q^2), and (4) evaluating the structural model quality using the standardized root mean square residual (SRMR).

For the first evaluation criterion, collinearity was assessed by examining the VIF values of the latent constructs in the SMHE-SQ model (Figure 3), which resulted in a VIF value of 1.000, confirming no collinearity between the constructs (Hair et al., 2013).

The significance and relevance of the structural model relationship between SMHE and SQ were evaluated by calculating the path coefficient representing the hypothesized relationship (H1). As shown in Table 10, the results indicated that at a 5% significance level, the calculated t-value (22.335) exceeded the threshold of 1.96. This confirmed that

the relationship between SMHE and SQ in the SM-HE-SQ structural model is both statistically signifi-

cant and positive, thereby supporting the proposed hypothesis.

*Table 10 Hypothesis testing - examining the direct relationship in structural model SMHE-SQ*

Hypothesis		Original sample β	t - value	p - value	95% confidence interval	confirmation
H1	SMHE → SQ	0.406	22.335	0.000	0.372 – 0.443	+

Source: Authors

As shown in Table 10, the proposed hypothesis H1—"Sustainable marketing statistically signifi-cantly, directly, and positively impacts perfor-mance, assessed as higher education institution service quality" (ß = 0.406, t-value = 22.335, p-value = 0.000)—was confirmed. The empirical relation-ship between the constructs in the proposed model was statistically significant at the 5% level (p < 0.05) and strong, with a 95% confidence interval for the population estimate ranging from 0.372 to 0.443.

Predictive significance, the second criterion for evaluating the structural model, was assessed by calculating the coefficient of determination (R^2), the effect size (f^2), and the cross-validated redundancy measure (Q^2). The R^2 values for the dependent con-structs were derived using the corresponding path coefficients and standardized factor loadings for the reflective construct SQ, along with the weight-ing values for the higher-order formative construct SMHE.

Figure 3 Structural model SMHE-SQ
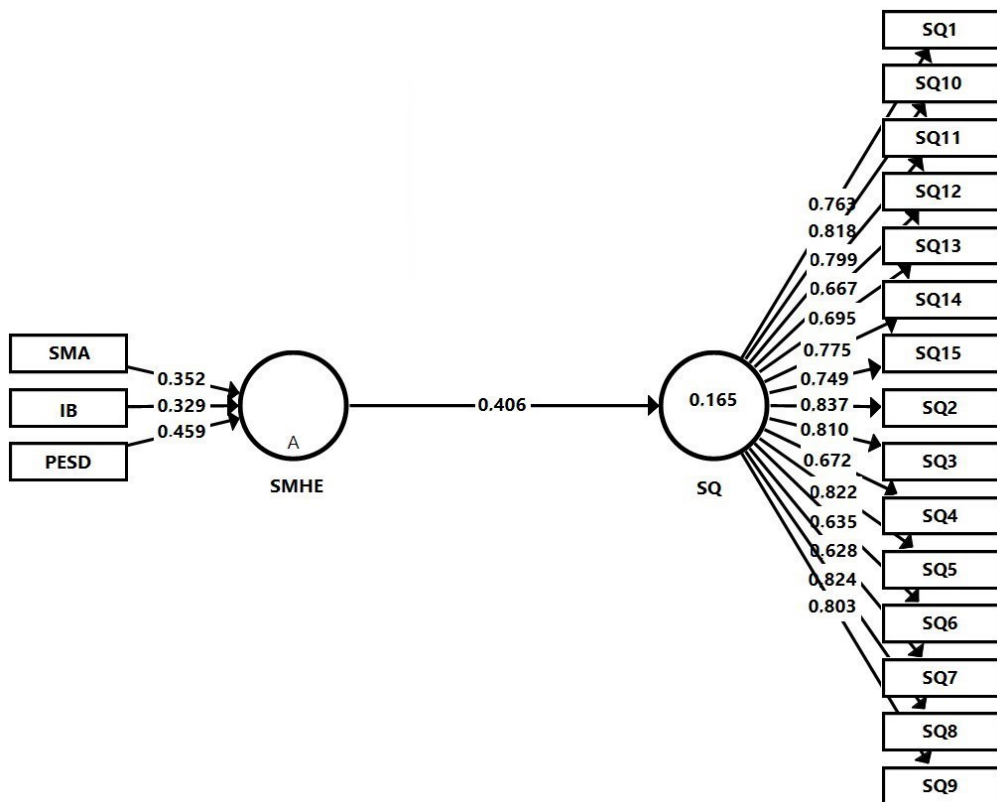
Source: Authors

Figure 3 shows that, according to Ritchey (2008, cited in Figueroa-García et al., 2018), the calculated R² value for the dependent variable SQ (R² = 0.165), though relatively low, falls within an acceptable range. This indicates that the SMHE-SQ model explains 16.5% of the variance in the dependent construct SQ. The calculated effect size (f²) for the SMHE-SQ structural model was 0.198, which represents a moderate influence of the exogenous latent construct SMHE on the endogenous variable SQ, based on Hair et al. (2013).

The Stone-Geisser coefficient (Q²) for the endogenous construct SQ was 0.495, which is greater than zero, confirming the predictive relevance of the SMHE-SQ model, as suggested by Hair et al. (2019).

Finally, the overall quality of the model was evaluated using the standardized root mean square residual (SRMR). Since the recommended threshold for SRMR is below 0.10 (Garson, 2016), the obtained value of 0.085 indicates an acceptable quality level for the proposed SMHE-SQ structural model.

6. Conclusion

This paper highlights the importance of sustainable marketing in higher education, focusing specifically on students as key stakeholders. Sustainable marketing is a function of sustainable economic development, which is basically a management tool. As a business and management function, marketing should help to improve the management processes of organizations by aligning products and services with the needs of target markets, future generations and society as a whole.

The main findings of this research are reflected in the answers to the research questions obtained based on the results of the research conducted on a sample of 1,663 students from the University of Rijeka. By applying the PLS-SEM method, which is suitable for testing new research theories, with the aim of further developing the existing theoretical framework using structural models that include one or more formative constructs and their relationships (Hair et al., 2019). A structural model of the relationship between sustainable marketing and business success was developed and tested. After analysis, all direct relationships in the structural model were found to be statistically significant and

positive, confirming the hypothesis: Sustainable marketing of a university has a positive and statistically significant effect on university performance, seen as service quality (ß = 0.406, t-value = 22.335, p = 0.000), with the empirical relationship between the two observed constructs being statistically significant and strong at the p < 0.05 level, with a population estimation interval of 0.372 to 0.443 at the 95% confidence level. A model of sustainable marketing of HEIs was developed and tested, which can serve as a stimulus or starting point for HEI managers to evaluate and improve existing measures or identify specific groups of key activities and appropriate ways to implement sustainable marketing. By understanding student perceptions, attitudes and behaviors regarding sustainable marketing, HEIs can develop more effective strategies that align with student values and contribute to the promotion of sustainability goals.

This research has certain limitations that do not diminish the significance of the findings and conclusions if interpreted appropriately, while providing opportunities for improvement in future research. Empirical research has shortcomings and limitations related to the research approach, the research process and the processing and analysis of data collected with regard to the application of the quantitative approach. The conducted research has limitations related to its implementation at only one public university and the inability to compare it with related previous research, as it primarily focuses on the profit sector and mostly covers only one of the three dimensions of sustainable development. Since the research results based on the proposed model are not considered in a temporal perspective, but exclusively in the context of the given situation, and sustainable marketing necessarily involves a long-term perspective, it is necessary to consider it appropriately in the future and examine the impact of sustainable marketing on the performance of HEIs over certain periods of time (longitudinal research).

Acknowledgement

References

1. Aleixo, A. M., Leal, S. & Azeiteiro, U. M. (2018). Conceptualization of sustainable higher education institutions, roles, barriers, and challenges for sustainability: An exploratory study in Portugal. *Journal of Cleaner Production*, *172*, 1664-1673. https://doi.org/10.1016/j.jclepro.2016.11.010

2. Annan-Diab, F. & Molinari, C. (2017). Interdisciplinarity: Practical approach to advancing education for sustainability and for the Sustainable Development Goals. *The International Journal of Management Education*, *15*(2), 73-83. https://doi.org/10.1016/j.ijme.2017.03.006

3. Arbula Blecich, A. (2024). Efficiency evaluation of higher education sector in Europe—window DEA based approach. *Journal of the Polytechnic of Rijeka*, *12*(1), 71-90. https://doi.org/10.31784/zvr.12.1.9

4. Arena, M., Azzone, G. & Mapelli, F. (2018). What drives the evolution of Corporate Social Responsibility strategies? An institutional logics perspective. *Journal of Cleaner Production*, *171*, 345-355. https://doi.org/10.1016/j.jclepro.2017.09.245

5. Barilović, Z., Leko Šimić, M. & Štimac, H. (2013). Istraživanje kvalitete obrazovne usluge s područja projektnoga menadžmenta. *Ekonomski vjesnik*, *26* (2), 371-380.

6. Becker, J.-M., Klein, K. & Wetzels, M. (2012). Hierarchical Latent Variable Models in PLS-SEM: Guidelines for Using Reflective-Formative Type Models. *Long Range Planning*, *45*(5-6), 359-394. https://doi.org/10.1016/j.lrp.2012.10.001

7. Budihardjo, M. A., Ramadan, B. S., Putri, S. A., Wahyuningrum, I. F. S. & Muhammad, F. I. (2021). Towards sustainability in higher-education institutions: analysis of contributing factors and appropriate strategies. *Sustainability*, *13*(12), 6562. https://doi.org/10.3390/su13126562

8. Casarejos, F., Frota, M. N., & Gustavson, L. M. (2017). Higher education institutions: a strategy towards sustainability. *International Journal of Sustainability in Higher Education*, *18*(7), 995-1017. https://doi.org/10.1108/IJSHE-08-2016-0159

9. Degtjarjova, I., Lapina, I. & Freidenfelds, D. (2018). Student as stakeholder: Voice of customer in higher education quality development. *Marketing and Management of Innovations*, (2), 388-398. https://doi.org/10.21272/mmi.2018.2-30

10. Faul, F., Erdfelder, E., Buchner, A. & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

11. Fuchs, P., Raulino, C., Conceição, D., Neiva, S., de Amorim, W. S., Soares, T. C. & Guerra, A. (2020). Promoting sustainable development in higher education institutions: the use of the balanced scorecard as a strategic management system in support of green marketing. *International Journal of Sustainability in Higher Education*, *21*(7), 1477-1505. https://doi.org/10.1108/IJSHE-02-2020-0079

12. Hair, J. F, Hult, G. T. M., Ringle, C. M. & Sarstedt, M. (2013). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Sage Publications.

13. Hair, J. F., Howard, M. C. & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis, *Journal of Business Research*, *109*, 101-110. https://doi.org/10.1016/j.jbusres.2019.11.069

14. Hair, J. F., Risher, J. J., Sarstedt, M. & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2-24. https://doi.org/10.1108/EBR-11-2018-0203

15. Henseler, J., Ringle, C. M. & Sarstedt, M. (2015). A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling. *Journal of the Academy of Marketing Science*, *43*, 115-135. https://doi.org/10.1007/s11747-014-0403-8

16. Hübscher, C., Hensel-Börner, S. & Henseler, J. (2022). Social marketing and higher education: partnering to achieve sustainable development goals. *Journal of Social Marketing*, *12*(1), 76-104. https://doi.org/10.1108/JSOCM-10-2020-0214

17. Jarvis, C. B., MacKenzie, S. B. & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*(2), 199-218. https://doi.org/10.1086/376806

18. Jongbloed, B., Enders, J. & Salerno, C. (2008). Higher education and its communities: Interconnections, interdependencies and a research agenda. *Higher Education*, *56*, 303-324. https://doi.org/10.1007/s10734-008-9128-2

19. Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling*. The Guilford Press.

20. Leko Šimić, M. & Štimac, H. (2013). Kvaliteta obrazovne usluge - slučaj Ekonomskog fakulteta u Osijeku. In Čavrak, V. & Gelo, T. (Eds.). *Proceedings of scientific conference: Ekonomsko obrazovanje u Republici Hrvatskoj - jučer, danas, sutra* (pp. 249-262). Zagreb: Faculty of Economics and Business.

21. Lozano, R., Lukman, R., Lozano, F. J., Huisingh, D. & Lambrechts, W. (2013). Declarations for sustainability in higher education: becoming better leaders, through addressing the university system. *Journal of Cleaner Production*, *48*, 10-19. https://doi.org/10.1016/j.jclepro.2011.10.006

22. Memon, M. A., Ting, H., Cheah, J. H., Thurasamy, R., Chuah, F. & Cham, T. H. (2020). Sample size for survey research: review and recommendations. *Journal of Applied Structural Equation Modeling*, *4*(2), 1-20. https://doi.org/10.47263/jasem.4(2)01

23. Meštrović, D., Bagarić, L. & Lončarić, D. (2021). Attitudes and perceptions of sustainable marketing in higher education–designing a measurement instrument. *Ekonomski vjesnik/Econviews-Review of Contemporary Business, Entrepreneurship and Economic Issues*, *34*(2), 401-413. https://doi.org/10.51680/ev.34.2.12

24. Pearce, J. A., Robbins, D. K. & Robinson Jr, R. B. (1987). The impact of grand strategy and planning formality on financial performance. *Strategic Management Journal*, *8*(2), 125-134. https://doi.org/10.1002/smj.4250080204

25. Pedro, E. D. M., Leitão, J. & Alves, H. (2020). Stakeholders' perceptions of sustainable development of higher education institutions: An intellectual capital approach. *International Journal of Sustainability in Higher Education*, *21*(5), 911-942. https://doi.org/10.1108/IJSHE-01-2020-0030

26. Purwanto, A. & Sudargini, Y. (2021). Partial least squares structural equation modeling (PLS-SEM) analysis for social and management research: a literature review. *Journal of Industrial Engineering & Management Research*, *2*(4), 114-123.

27. Ruiz de Sabando, A., Forcada, J. & Zorrilla, P. (2018) The marketing orientation as a university management philosophy: a framework to guide its application. *Cuadernos de gestión, 18*(2), 37-57. https://doi.org/10.5295/cdg.150576al

28. Sandmaung, M. & Ba Khang, D. (2013). Quality expectations in Thai higher education institutions: Multiple stakeholder perspectives. *Quality Assurance in Education*, *21*(3), 260-281. https://doi.org/10.1108/QAE-11-2012-0044

29. Sarstedt, M., Hair, J. F, Cheah, J. H., Becker, J. M. & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal*, *27*(3), 197-211. https://doi.org/10.1016/j.ausmj.2019.05.003

30. Shawe, R., Horan, W., Moles, R. & O'Regan, B. (2019). Mapping of sustainability policies and initiatives in higher education institutes. *Environmental Science & Policy*, *99*, 80-88. https://doi.org/10.1016/j.envsci.2019.04.015

31. Širola, D. & Mihaljević, A. (2016). Primjena simulacija u funkciji kvalitetnijeg visokoškolskog poduzetničkog obrazovanja. *Journal of the Polytechnic of Rijeka*, *4*(1), 79-96.

32. Teeroovengadum, V., Kamalanabhan, T. J. & Seebaluck, A. K. (2016). Measuring service quality in higher education: Development of a hierarchical model (HESQUAL). *Quality Assurance in Education*, *24*(2), 244-258. https://doi.org/10.1108/QAE-06-2014-0028

33. Zhou, L., Rudhumbu, N., Shumba, J. & Olumide, A. (2020). Role of higher education institutions in the implementation of sustainable development goals. In Nhamo, G. & Mjimba, V. (Eds.), *Sustainable Development Goals and Institutions of Higher Education* (pp. 87-96). Springer. https://doi.org/10.1007/978-3-030-26157-3_7

Review paper

Predrag Trpeski, Biljana Tashevska, Gunter Merdzan
The role of institutions in economic growth: A systematic literature review

Predrag Trpeski
Ss. Cyril and Methodius
University in Skopje
Faculty of Economics
1000 Skopje, Republic
of North Macedonia
predrag.trpeski@eccf.ukim.edu.mk

Gunter Merdzan
Ss. Cyril and Methodius
University in Skopje
Faculty of Economics
1000 Skopje, Republic
of North Macedonia
gjunter.merdzan@eccf.ukim.edu.mk

Biljana Tashevska
Ss. Cyril and Methodius
University in Skopje
Faculty of Economics
1000 Skopje, Republic
of North Macedonia
biljana.tashevska@eccf.ukim.edu.mk

THE ROLE OF INSTITUTIONS IN ECONOMIC GROWTH: A SYSTEMATIC LITERATURE REVIEW

Abstract

Purpose: This paper gives a systematic literature review of the research field of institutions and economic growth. The goal of the paper is to provide information on predominant trends in publishing studies, prevailing sentiments on the role of institutions, and emerging themes in the field, to identify gaps in the existing literature and possible avenues for further research.

Methodology: The study systematically reviews papers on institutions and economic growth from the Scopus database following the PRISMA protocol. The research uses descriptive analysis (annual distribution of articles, frequently referenced papers, central and emerging topics), sentiment analysis, and keyword co-occurrence network analysis.

Results: A total number of 78 papers published between 2006 and 2023 were analysed. The results suggest a growing interest in the subject, encompassing topics such as governance, human capital, trade openness, and the resource curse. Sentiment analysis suggests that most of the literature is optimistic about the impact of institutions on fostering economic growth. Keyword analysis indicates that institutions, governance, and economic growth remain key areas of interest, with increasing emphasis on region-specific research and empirical approaches.

Conclusion: The prevailing research indicates that quality institutions play a crucial role in economic growth, enhancing the impact of other factors like financial development, trade openness, and human capital. The review underscores the pivotal role of institutions in sustaining long-term economic progress and suggests further exploration of less researched areas, such as regional development and entrepreneurship, and the utilisation of additional scientific databases to deepen our comprehension of institutional dynamics.

Keywords: Economic growth, institutions, systematic literature review

1. Introduction

Three main theories explain how economies grow: the neo-Keynesian Harrod-Domar theory, the neo-classical Solow theory, and the new growth theories created by Paul Romer (Romer, 1986; 1990) and Robert Lucas. These theories and models suggest that savings, investment, overall productivity, technology, knowledge, and innovation can help explain development differences among countries.

One theory, the geography hypothesis, proposes that geography, culture, and beliefs influence a country's development. Another hypothesis, overlooked for a long time, is the impact of institutions on economic growth. The emergence of New Institutional Economics and the work of Douglass North (North, 1987; 1989; 1990) highlights the crucial role of institutions in economic growth. North defines institutions as the "rules of the game in a society that govern or regulate interpersonal relationships". Institutions determine economic incentives and resource distribution in a society. Well-designed institutions reduce uncertainty and increase incentives for investment and economic activity, enhancing economic growth and development.

Dani Rodrik (Rodrik, 2000; 2007) emphasises the interconnectedness of economic and political institutions, highlighting the crucial role of governance quality in upholding economic growth. He asserts that while economic institutions, such as property rights and regulatory frameworks, are vital, they must be underpinned by political institutions that foster transparency and accountability, mitigating the risks of political instability or corruption that can hinder progress. For example, countries with democratic institutions generally experience more substantial economic growth over time, as political competition and oversight contribute to transparent and accountable governance (Rodrik et al., 2004).

Drawing on the pivotal contributions of North and Rodrik, contemporary research has increasingly emphasised the important role of institutions. Institutions, whether formal or informal, not only shape the legal and regulatory framework for economic activities but also influence the broader economic landscape, by impacting social trust, political stability and governance. Recent empirical studies in New Institutional Economics indicate that nations with stronger institutions tend to experience higher levels of foreign investment, more extensive technological diffusion, and enhanced rates of eco-nomic growth (Acemoglu et al., 2001; La Porta et al., 1999).

Researchers often consider various institutional dimensions, including property rights, regulatory efficiency, the rule of law, and control of corruption. Strong property rights enforcement, for example, creates a secure investment environment, encouraging domestic and foreign capital inflows. Similarly, efficient regulation can foster entrepreneurial activity by reducing business costs or hinder it through excessive bureaucracy and red tape (Knack & Keefer, 1995). A crucial aspect of the institutional framework is the role of governance in mediating the relationship between institutions and economic growth. Research indicates that better governance leads to more efficient public service delivery, reduced levels of corruption, and increases investor confidence, all of which enhance economic outcomes (Kaufmann et al., 2009). Another aspect is the influence of historical factors on institutional development. North (1990) stresses that institutional change tends to be gradual, as societies are bound by their past decisions. Acemoglu and Robinson (2012) further contend that inclusive institutions, fostering broad participation in economic and political activities, are more conducive to long-term economic growth than extractive institutions that centralise power and wealth in the hands of a few.

Given the complex nature of institutions and their impact on economic growth, this paper takes a comprehensive approach to examining the literature on institutions and economic growth. Bibliometric analysis through a systematic literature review aims to describe how the research field has evolved, identify significant trends, and uncover new themes in the field. More specifically, this study aims to address the following research questions:

RQ1. What are the bibliometric trends in the research of institutions and economic growth, including the volume of published papers over time and the top 10 most cited papers?

RQ2. How are the sentiments in the literature on institutions and economic growth distributed, and what does this indicate about the tone and impact of institutions on economic growth?

RQ3. What are the primary themes and emerging research focal points in institutions and economic growth?

In order to address these research questions, we utilised the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol to

identify, screen, and assess the eligibility of 78 articles sourced from the Scopus database. Initially, we conducted an extensive search using pertinent keywords related to institutions and economic growth, ensuring comprehensive coverage of articles on this topic. Following the compilation of the initial pool of studies, we applied the PRISMA screening process, which encompassed the phases of identification, screening, eligibility, and inclusion. These steps were essential for refining the pool of studies directly pertinent to our research focus. Subsequently, we conducted descriptive analysis, sentiment analysis, and keyword analysis.

Regarding the first research question, a descriptive analysis of the dynamics of published articles seeks to uncover significant trends in the publication of research on institutions and economic growth and determine whether there has been a rise or fall in academic interest in this area. Moreover, by identifying the top 10 most-cited papers, the study will shed light on the research that has had the most profound impact on academic discussions.

Addressing the second research question, we use sentiment analysis to examine the distribution of sentiments (positive, neutral or negative) in the abstracts of the reviewed literature. The study anticipates that the tone of the literature will reflect the prevailing perspectives on whether institutions contribute to or impede economic growth. Reflecting on existing literature, we expect a larger inclination towards a positive sentiment.

The third research question addresses the structure of the focus of studies in the field. By analysing keyword co-occurrence and utilising network, overlay, and density visualisation techniques, the study aims to identify clusters of related topics, revealing the main research themes and emerging trends within the field that have gained prominence in recent years.

The findings will contribute to a deeper comprehension of the existing research on the role of institutions on economic growth. The outcomes will offer scholars and researchers clear insight into the evolution and advancement of the field, pinpointing where significant contributions have been made and identifying the interconnectedness of concepts and the prevailing sentiment towards the role of institutions in economic growth. This knowledge will aid in determining potential avenues for future research. Understanding the persistent and emerging trends in the literature can inform strategic decisions and enhance evidence-based decision-making regarding institutional reforms, investment policies, and governance frameworks for governments and businesses, ensuring that their efforts align with global best practices and the latest research. For policymakers and government officials, understanding whether the academic literature predominantly reflects a positive or negative sentiment towards institutions could guide future reforms. For instance, a predominantly positive tone may indicate that institutional reforms have been largely effective and should be continued or expanded. Conversely, a negative tone could indicate the need for more fundamental changes or interventions to address institutional shortcomings.

This paper is organised into several sections. The Introduction delves into the theoretical background and research questions. Section 2 reviews the relevant empirical literature on the relationship between institutions and economic growth. Section 3 explains the methodology employed in the systematic literature review. Section 4 encompasses descriptive analysis, sentiment analysis and keyword analysis results and discussion. Finally, Section 5 provides concluding remarks and directions for further research.

2. Empirical background

The impact of institutions on economic growth has been extensively studied in empirical research. Many studies focus on the direct influence of quality institutions on economic efficiency through reducing transaction costs, protecting property rights, and maintaining law and order. These are vital factors for stimulating investment and innovation for driving economic growth. However, other research shows that high-quality institutions also affect economic growth indirectly, for example by amplifying the benefits of infrastructure, technology, and foreign investment.

Despite this research, most analyses only examine the macroeconomic level, neglecting the microeconomic aspects crucial for understanding how institutions impact individual decision-making, entrepreneurship, and business operations. Furthermore, the relationship between institutional quality and economic outcomes is complex and context-dependent, making it difficult to draw definitive conclusions. While institutions generally have a positive long-term impact on economic growth, their short-term effects are still subject to debate, suggesting the need for further research. Interestingly, resource-rich countries benefit

more from quality institutions, while those with fewer resources and weaker institutions face significant economic challenges.

Institutions play a significant role in shaping a country's long-term and short-term economic growth. Institutions' impact can be positive and negative, depending on how they are structured and managed. In general, and in most of the papers we reviewed, the institutions' impacts on economic growth are positive. Countries with stable, transparent, and accountable institutions tend to have better long-term economic growth. These institutions create a favourable environment for businesses and attract investment. Education and human capital development also contribute to sustained economic growth by providing individuals with the skills and knowledge needed for a modern workforce.

It has been observed that favourable institutions play a crucial role in promoting economic growth. The indices associated with institutional and policy rents are more significant than those focusing on curbing political rents and reducing transaction risks, particularly in developing countries. Additionally, there are positive complementarities between the indices related to political rents and those associated with risk-reducing technologies. These findings highlight the importance of creating and maintaining favourable institutions and policies to foster economic growth and development (Siddiqui & Ahmed, 2013).

The type of growth regime an economy belongs to is largely determined by its political institutions. Economic institutions, on the other hand, have a direct effect on growth rates within each regime. It can be said that political institutions act as a foundation for growth, providing a framework in which economic institutions and other standard factors can operate. This clarifies that effective political institutions are necessary for sustained economic growth. Furthermore, economic institutions are essential to nurturing and maintaining the positive effects of political institutions on growth (Flachaire et al., 2014).

Studies have shown that high institutional quality is vital for economic growth and development. Institutional indicators contribute over 0.3 percentage points to the estimated average annual growth rate of real output (Afonso, 2020). Improving institutional quality benefits countries at different stages of devel-

opment, leading to more investment and growth. This improved investment climate can also lead to more job creation, higher wages, and a general increase in living standards. This creates a more stable and prosperous economic environment (Gwartney et al., 2006; Valeriani & Peluso, 2011).

In developing countries, growth is influenced by institutions, particularly those related to property rights. It has been observed that institutions in one country can indirectly affect the economic growth of neighbouring countries through positive spillover effects. This supports an augmented convergence process faster than the conventional beta-convergence in standard growth models (Ahmad & Hall, 2017).

There are interdependent relationships in the short term between institutional quality, government expenditure, tax revenue, and economic growth. Tax revenue is a significant factor in driving long-term economic growth. To ensure sustained long-term economic growth of countries, it is crucial to develop stronger institutions and more effective fiscal policies simultaneously (Arvin et al., 2021).

It has been found that foreign aid can positively impact the economic growth of developing countries, but this is only the case when the institutional quality is high. This means that the quality of the country's institutions plays a crucial role in determining how effective foreign aid will promote economic growth. Therefore, it is recommended that foreign aid should be directed towards countries with higher institutional quality to achieve maximum effectiveness (Boateng et al., 2021; Maruta et al., 2020).

It has also been observed that institutional quality, CO_2 emissions, and economic growth in developing countries are closely linked. Therefore, adopting holistic co-development policy frameworks that strengthen governance institutions and promote clean-energy industrial strategies can positively impact economic growth in these countries. Policymakers in developing countries are encouraged to implement such policies to reduce their carbon footprint and promote economic growth (Ahmed et al., 2022; Jiang et al., 2023; Nair et al., 2021).

Countries can benefit or suffer from their natural resources, depending on the quality of their institu-

tions and human capital. An analysis was conducted using an autoregressive distributed lag model to examine how natural resources, institutional quality, and human capital interact and affect economic growth. The findings indicate that countries must invest in human capital and anti-corruption measures to effectively utilise natural resources (Abdulahi et al., 2019; Antonakakis et al., 2017; Qiang & Jian, 2020; Zallé, 2019).

The influence of informal institutions on opportunity entrepreneurship is greater than that of formal institutions. In countries across the sample, including Latin American countries, factors like control of corruption, confidence in one's abilities, and access to private credit positively impact opportunity entrepreneurship and contribute to economic growth. By promoting effective institutions, we can foster opportunity entrepreneurship and create conditions for economic growth (Aparicio et al., 2016).

Among the papers we reviewed, only several found negative or neutral relations between institutions and economic growth. These studies suggest that institutions are less influential in economic growth than is often assumed. Furthermore, it has been suggested that other factors, such as technological progress and human capital, largely determine economic growth (Acheampong et al., 2021; Doré & Teixeira, 2023).

In summary, institutions can positively or negatively affect long-term and short-term economic growth. Institutions that function well and encourage stability, transparency, and accountability are crucial for maintaining economic development. Conversely, institutions that function poorly and are characterised by corruption, inefficiency, and a lack of transparency can impede economic growth. Therefore, policymakers must prioritise creating and maintaining effective institutions to promote sustainable and long-term economic growth.
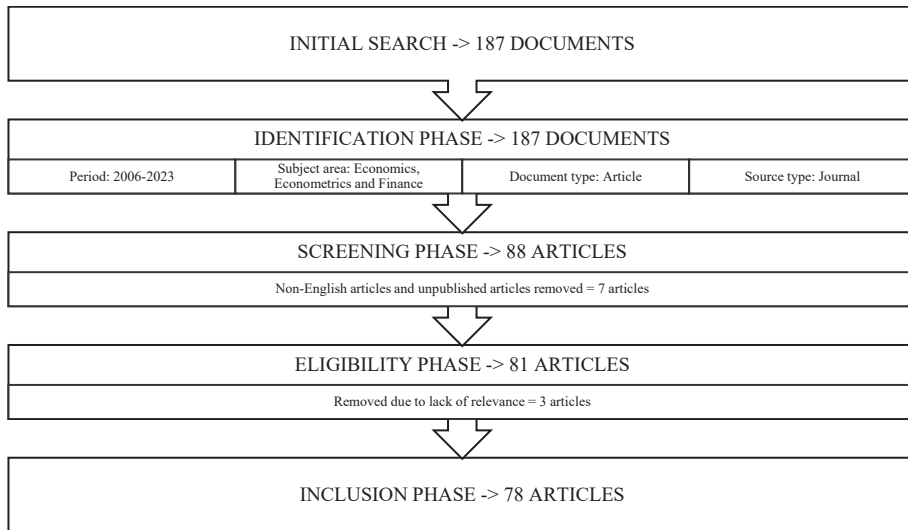
3. Methodology and data

We utilised the PRISMA protocol for systematic reviews to ensure a comprehensive and transparent analysis of the relationship between institutions and economic growth. This protocol comprises four stages: identification, screening, eligibility, and inclusion (Moher et al., 2009). As part of our search strategy,

we combed through the SCOPUS database using the keywords "Institutions AND Economic Growth" OR "Institutional Quality AND Economic GROWTH" from 2006-2023, which yielded 187 papers in the titles, abstracts, and keywords.

Figure 1 demonstrates the implementation of the PRISMA protocol for a systematic literature review of the 187 papers obtained from our search on the Scopus database. The process began with an initial search that retrieved 187 documents from the Scopus database using the keywords "Institutions AND Economic Growth" OR "Institutional Quality AND Economic Growth". The methodology is divided into four phases: identification, screening, eligibility, and inclusion.

1. All 187 documents found during the search are included during the identification phase. These documents are filtered based on their publication year (2006-2023) and limited to the subject areas of Economics, Econometrics, and Finance. Furthermore, only journal articles were included in the further review, while other document types, such as conference papers and book chapters, were excluded.

2. In the screening phase, the initial filters reduced the number of articles to 88. At this stage, non-English articles and those that were not at the final publication stage were excluded, resulting in the removal of 7 articles. This left a total of 81 articles for the next phase of assessment.

3. During the eligibility phase, each article underwent a thorough evaluation to determine its relevance to the research topic. Articles that did not directly relate to the role of institutions in economic growth were excluded, resulting in the removal of three more articles. As a result, 78 articles remained eligible for inclusion.

4. The inclusion phase resulted in 78 articles to be included in the final analysis. These articles serve as the basis for the bibliometric analysis, which explores trends, sentiments, and thematic developments in the literature on institutions and economic growth.

Figure 1 Visualisation of the PRISMA protocol

INITIAL SEARCH -> 187 DOCUMENTS

IDENTIFICATION PHASE -> 187 DOCUMENTS

| Period: 2006-2023 | Subject area: Economics, Econometrics and Finance | Document type: Article | Source type: Journal |

SCREENING PHASE -> 88 ARTICLES

Non-English articles and unpublished articles removed = 7 articles

ELIGIBILITY PHASE -> 81 ARTICLES

Removed due to lack of relevance = 3 articles

INCLUSION PHASE -> 78 ARTICLES

Source: Authors' presentation

We conducted a thorough descriptive analysis of the selected papers to determine whether interest in exploring the relationship between institutions and economic growth has grown over time. This involved examining the annual number of publications to identify potential trends in research output on this topic. Additionally, we aimed to identify the individual papers with the highest citation counts, offering insight into which works have made the greatest impact on the academic discourse.

Next, we conducted a sentiment analysis of the abstracts to gain a deeper understanding of the tone and perspectives expressed. This analysis categorised the abstracts based on whether they conveyed a positive, neutral, or negative outlook on the role of institutions in economic growth. This enabled us to evaluate whether the prevailing sentiment in the literature reflects optimism about institutional reforms or underscores concerns and challenges regarding institutional efficiency and governance. A positive tone in the literature indicates optimism about the role of institutions, encompassing discussions about successful institutional reforms, improvements in governance, or instances where institutions have played a key role in fostering economic growth. A neutral tone may reflect a balanced or indifferent perspective, often found in descriptive or empirical studies without strongly expressing opinions about institutional efficacy. On the other hand, a negative tone would signal concerns or challenges, including institutional failures, corruption, or inefficiency, which could hinder economic growth.

Utilising the VOSviewer software for bibliometric analysis, we constructed a keyword co-occurrence network to visually represent the relationships between key concepts and terms found in the reviewed abstracts. This network enables us to observe the frequency of association between keywords, offering a visual map of the intellectual structure within the field. In addition, we generated an overlay visualisation to track the prominence of specific keywords over time, helping us identify topics that have gained or lost relevance in recent years. The use of density visualisation methods allowed us to pinpoint areas of research concentration within the literature. By visualising the most densely interconnected keywords and concepts, we were able to identify "hotspots" where certain themes are especially well-developed or popular among researchers. This analysis also enabled us to explore which empirical models and methodologies are most frequently used in the studies, providing insights into the prevailing analytical approaches in the field.

It is important to note that the search query used in this analysis, specifically "Institutions AND Economic Growth" OR "Institutional Quality AND Economic Growth", has some limitations. The use of

quotation marks in the search query means that the Scopus database returned papers where these exact phrases appeared together in the title, abstract, or keywords. However, this approach may have limited the scope of our review by excluding papers that discuss institutions and economic growth in broader or more varied contexts. Also, we only focus on articles written in English and published in journals indexed in the SCOPUS database. Consequently, we might have omitted articles published in other languages and indexed in other databases such as Web of Science, Google Scholar, etc., which might provide additional insight.

4. Results and discussion

4.1 Descriptive analytics

In this subsection, we analyse the frequency of papers focused on institutions and economic growth

from 2006 to 2023. Figure 2 depicts the fluctuating trend in the number of published articles during this period, with several distinct peaks:

- – The number of papers peaked in 2017, with 10 publications.

- – Another peak occurred in 2020, with the highest number of papers (11) recorded in this period.

- – The year 2021 experienced a slight decline to 8 papers, followed by a substantial drop in 2022 to just 1 paper. However, 2023 witnessed a recovery, with 7 papers published.

The dotted trendline indicates an upward trajectory, reflecting a growing interest among researchers in studying the role of institutions in economic growth. This upward trend, notwithstanding yearly fluctuations, indicates an increasing focus on this research topic in academic circles.

Figure 2 Number of papers published by year
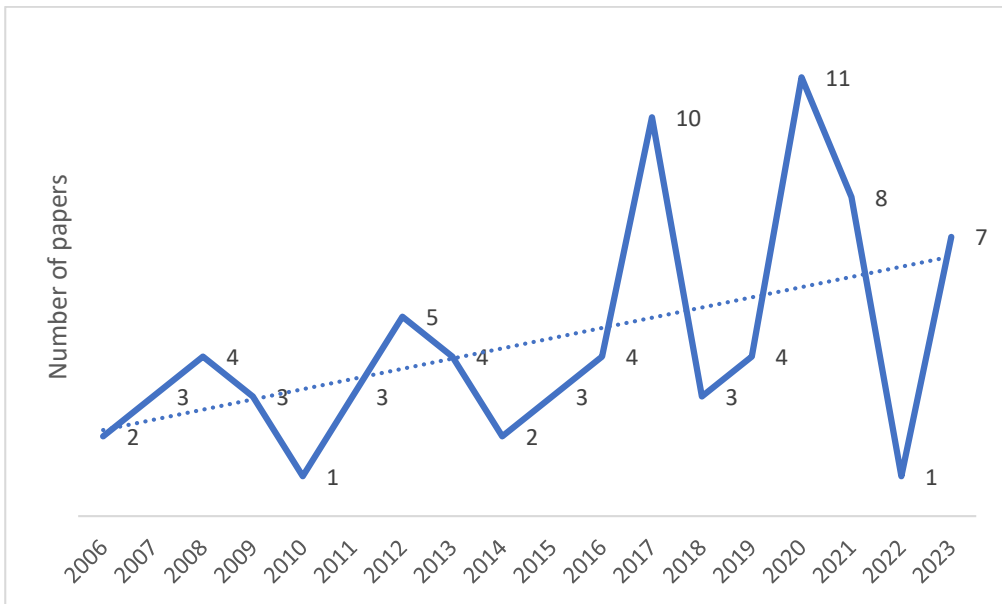
Source: Authors' calculations

The rising volume of published articles indicates a growing interest in the correlation between institutions and economic growth. This pattern suggests that scholars are acknowledging the pivotal role institutions play in shaping economic outcomes. This upward trajectory indicates a field of study with

numerous opportunities for further exploration, especially in areas that may still be underdeveloped. Spikes in publication numbers may also signify breakthroughs or emerging trends in institutional research, motivating scholars to expand on these foundations.

For policymakers, the mounting interest underscores the significance of institutions in devising effective economic policies. Governments, particularly in developing or transitional economies, can interpret this as a call to prioritise institutional reforms, recognising their essential role in long-term economic success. The expanding body of literature also emphasises the importance of robust institutions in providing a stable and predictable environment for business. Additionally, international organisations involved in development efforts can leverage this trend as empirical support for advocating institutional reforms to promote sustainable development.

Table 1 presents the top 10 papers with the highest number of citations in the field of institutions and economic growth, reflecting the most influential studies on the topic. These papers cover a wide array of themes, including the relationship between natural resources, institutions, and economic growth, the effects of remittances and institutional quality on growth, and the role of institutions in innovation, policy volatility, and financial globalisation.

At the top of the list is Brunnschweiler's (2008) paper "Cursing the Blessings? Natural Resource Abundance, Institutions, and Economic Growth", published in World Development, with 423 citations. This study delves into the widely debated resource curse hypothesis and investigates the role of institutions in mitigating the adverse effects of natural resource abundance on economic growth. Next, with 286 citations, there follows "Remittances, Institutions, and Economic Growth" by Catrinescu et al. (2009), also published in World Development, which examines how remittances, when combined with institutional quality, can impact economic growth, emphasising the significance of the institutional context in realising the potential benefits of remittances. Other papers on the list, such as those by Gwartney et al. (2006) and Tebaldi and Elmslie (2013), focus on how institutions influence investment and innovation. Meanwhile, studies by Fatás & Mihov (2013) and Butkiewicz and Yanikkaya (2006) examine policy volatility and the maintenance of the rule of law or democratic institutions as crucial factors for growth.

Table 1 Top 10 papers with the most citations

Title	Authors	Year	Journal	Citations
Cursing the Blessings? Natural Resource Abundance, Institutions, and Economic Growth	Brunnschweiler C.N.	2008	World Development	423
Remittances, Institutions, and Economic Growth	Catrinescu N.; Leon-Ledesma M.; Piracha M.; Quillin B.	2009	World Development	286
Institutions and the impact of investment on growth	Gwartney J.D.; Holcombe R.G.; Lawson R.A.	2006	Kyklos	203
Does institutional quality impact innovation? Evidence from cross-country patent grant data	Tebaldi E.; Elmslie B.	2013	Applied Economics	155
Policy volatility, institutions, and economic growth	Fatás A.; Mihov I.	2013	Review of Economics and Statistics	149
Institutional quality and economic growth: Maintenance of the rule of law or democratic institutions, or both?	Butkiewicz J.L.; Yanikkaya H.	2006	Economic Modelling	136

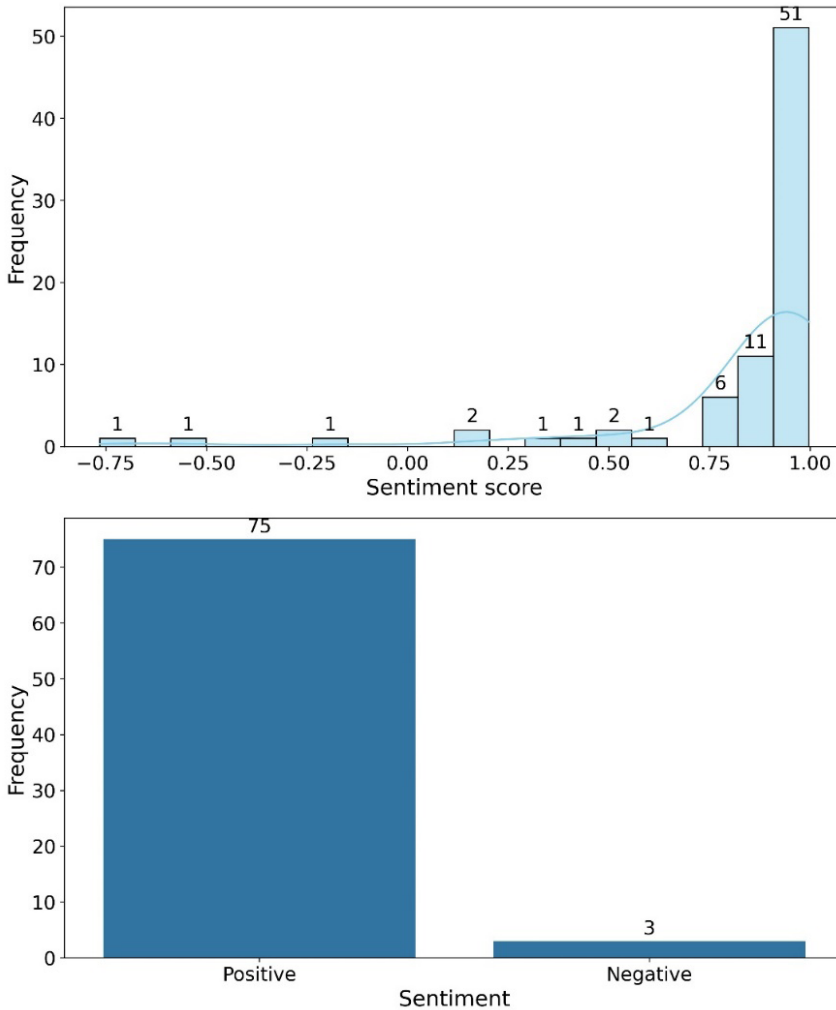| Title	Authors	Year	Journal	Citations
Renewable energy, CO2 emissions and economic growth in sub-Saharan Africa: Does institutional quality matter?	Acheampong A.O.; Dzator J.; Savage D.A.	2021	Journal of Policy Modeling	103
Institutions, human capital, and growth: The institutional mechanism	Dias J.; Tebaldi E.	2012	Structural Change and Economic Dynamics	92
How do financial globalization, institutions and economic growth impact financial sector development in European countries?	Nasreen S.; Mahalik M.K.; Shahbaz M.; Abbas Q.	2020	Research in International Business and Finance	77
Natural Resources, Institutional Quality, and Economic Growth in China	Ji K.; Magnus J.R.; Wang W.	2014	Environmental and Resource Economics	76

Source: Authors' calculations

The diverse range of topics and the substantial number of citations of these papers underscore the importance of institutional quality across different areas of economic research. For academics, these works provide foundational theories and evidence on the influence of institutions on economic growth, covering topics from resource management to innovation. Researchers can use these studies as a basis for exploring new areas or expanding their findings to different environments. For policymakers, the findings stress the significance of strong institutional frameworks in promoting sustainable development. For example, Brunnschweiler's (2008) work on the resource curse and institutional mitigation offers valuable guidance for resource-rich nations, while Catrinescu et al. (2009) underscore the importance of institutional quality in maximising the impact of remittances on growth. For businesses and investors, the focus on institutional stability in these papers offers valuable insights for risk management and long-term planning, especially in regions undergoing institutional changes. Similarly, international organisations like the World Bank and the IMF can leverage these findings to support their advocacy for institutional reforms in developing nations. In summary, these works contribute not only to theoretical debates but also offer practical insights for policymakers, underscoring the significant role that well-functioning institutions play in fostering economic growth and sustainable economic development.

4.2 Sentiment analysis

Sentiment analysis of the abstracts highlights a distinctly positive inclination in the literature concerning institutions and economic growth. Among the 78 abstracts reviewed, 75 portray a positive sentiment, while only 3 convey a negative sentiment (the lower section of Figure 3). The scale of sentiment scores, spanning from -0.765 (negative) to 1.00 (positive), suggests that most studies express optimism or a favourable outlook regarding the influence of institutions on economic growth. The histogram in the upper section of Figure 3 illustrates that the majority of sentiment scores cluster near 1.00, confirming the dominance of positive sentiment.

The prevailing positive sentiment in the literature indicates that institutions are widely recognised as crucial catalysts for economic advancement. This widespread optimism may stem from empirical evidence showing that well-functioning institutions, encompassing strong legal frameworks, property rights, and regulatory bodies, as well as government efficiency and low levels of corruption, can bolster economic performance, mitigate uncertainty, and stimulate both local and foreign investments.

**Figure 3 Abstracts sentiment distribution and categories**

The prevalence of positive sentiment emphasises the necessity of continuing to implement and enhance institutional reforms. These reforms are widely viewed as impacting economic outcomes directly and positively, making them crucial for businesses and investors seeking stable environments for long-term investments. This trend also affirms the support for institutional reforms in developing regions by international organisations. Moreover, the optimistic outlook for governments in transitional economies highlights the broad recognition of institutional strengthening as an effective tool for driving economic growth, urging them to prioritise such efforts in their development agendas.

4.3 Keyword analysis

In this subsection, we analyse the interconnectivity of themes within the literature on institutions and economic growth, as visualised in the keyword co-occurrence network produced by VOSviewer, which is shown in Figure 4. The figure is segmented into clusters, each denoted by a distinct colour, signifying sets of closely associated keywords that often appear together in the same studies. These clusters illuminate pivotal research themes in the field.

Figure 4 Network visualisation for keyword co-occurrence
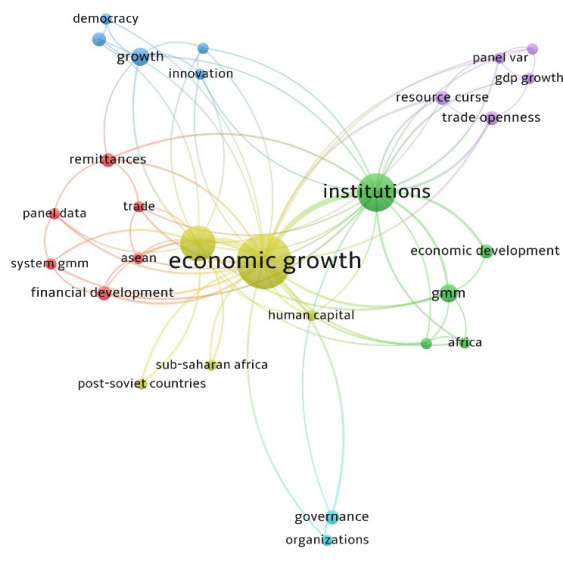
Source: Authors' calculations

Cluster 1 (red cluster) emphasises the significance of financial development, remittances, and trade, particularly focusing on regions such as ASEAN. The mention of "system GMM" and "panel data" suggests that much of the research employs advanced econometric techniques to investigate how financial systems, and trade policies impact economic growth. Remittances are also highlighted, indicating their substantial role in development outcomes. Ultimately, this cluster underscores the necessity of strong financial institutions and effective trade policies, especially in regions that heavily rely on financial inflows.

Cluster 2 (green cluster) is focused on the role of institutions and governance, with a specific emphasis on Africa. The frequent use of "GMM" and "panel estimation" suggests commonly employed empirical methods to assess the impact of governance reforms on growth. This cluster emphasises the crucial importance of institutional quality and governance reforms in driving economic progress, particularly in regions where institutional weaknesses persist.

Cluster 3 (blue cluster) explores the interplay between democracy, innovation, and economic growth. The prominence of the keyword "democ-

racy" indicates that the studies often compare democratic and non-democratic systems, exploring the ways in which political institutions influence economic growth. The research within this cluster delves into how democratic governance creates an environment that is conducive to innovation and technological advancement, both of which play key roles in long-term economic prosperity.

Cluster 4 (yellow cluster) is positioned at the heart of the network and focuses on economic growth, institutional quality, and human capital. By examining specific regional contexts, such as post-Soviet countries and sub-Saharan Africa, the literature delves into how variations in institutional quality across different regions impact economic outcomes. This cluster underscores the crucial importance of investments in education and human capital development in conjunction with institutional reforms to sustain growth.

Cluster 5 (purple cluster) focuses on examining the relationship between trade openness, resource management, and economic growth. The research sheds light on the difficulties that resource-rich nations face in utilising their natural wealth for economic progress (the concept of the "resource curse") and examines trade openness as a potential remedy

to these challenges. This cluster underscores the importance of robust institutional frameworks to ensure that trade and resource wealth have a positive impact on economic performance.

Cluster 6 (light blue cluster) explores the role of governance and organisational structures in supporting institutional quality. This cluster encompasses research highlighting the foundational importance of efficient and transparent governance structures for the effective functioning of institutions, which is a key driver of economic growth.

In summary, the clusters represent various but interconnected themes. Financial development, governance, innovation, human capital and trade openness all play crucial roles in shaping economic performance. Each cluster reinforces the overarching importance of strong institutions in ensuring sustained and inclusive economic growth.

The overlay visualisation in Figure 5 adds a chronological layer to the keyword co-occurrence network. The colours indicate the average publication years for each keyword. The colour gradient, ranging from blue (indicating earlier years) to yellow (indicating recent years), allows us to track the evolution of research focus over time.

Figure 5 Overlay visualisation for keyword co-occurrence
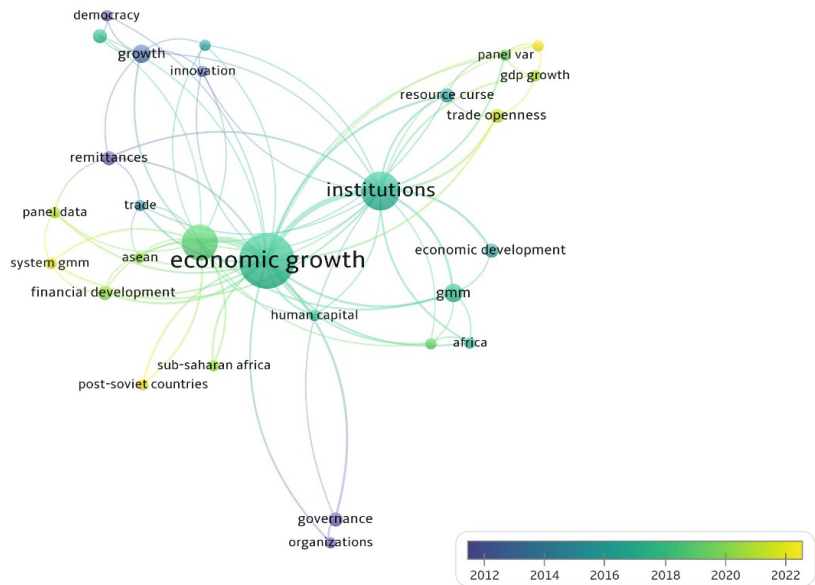
Source: Authors' calculations

Keywords such as "organisations", "governance", "democracy", and "remittances" are highlighted in blue, indicating their significance in earlier years (2008-2013). The research during this time predominantly centred on the fundamental role of governance and institutions in economic growth. Additionally, democracy and remittances garnered early attention, showcasing a curiosity about how political structures and financial inflows from emigrants impact economic performance. The inclusion of "innovation" and "growth" during this period signifies an early examination of the influence of governance on innovation-driven growth.

In the years 2014-2017 (highlighted in green), there was a noticeable increase in the discussion of topics such as "trade", "technological change", "economic development", and "resource curse". During this period, there was a shift in research focus towards understanding the interaction between trade policies, natural resource management, and institutional quality, particularly evident in the emphasis on the "resource curse" keyword. The impact of institu-

tional quality on development across different regions, particularly in Africa, also emerged as a key area of focus during this phase.

The emphasised research topics from 2018 to 2023, indicated in yellow, encompass keywords such as "ASEAN", "financial development", "trade openness", "system GMM", "regional development", and "post-Soviet countries". This trend indicates a growing emphasis on empirical methodologies (e.g., system GMM and panel VAR), region-specific studies, particularly on ASEAN and post-Soviet countries, and on the impact of institutions on global integration and development at both national and regional levels.

The overlay visualisation reveals that the literature on institutions and economic growth has progressed from general debates on governance and democracy to more focused areas such as trade openness, regional development, and on specific regions. This shift over time suggests that recent research is increasingly focused on empirical analysis and the unique institutional challenges in different regions.

In Figure 6, the density visualisation of keyword co-occurrence reflects the intensity of the research focus. Brighter (yellow) areas signify higher studies concentrations, while darker (blue/green) areas indicate less frequent exploration.

As expected, the terms "economic growth" and "institutions" are highlighted in a prominent bright yellow area, indicating that these topics are central to scholarly discourse in the field. A significant amount of literature focuses on how institutions directly impact economic performance, making this the main theme of the overall research landscape.

Figure 6 Density visualisation for keyword co-occurrence
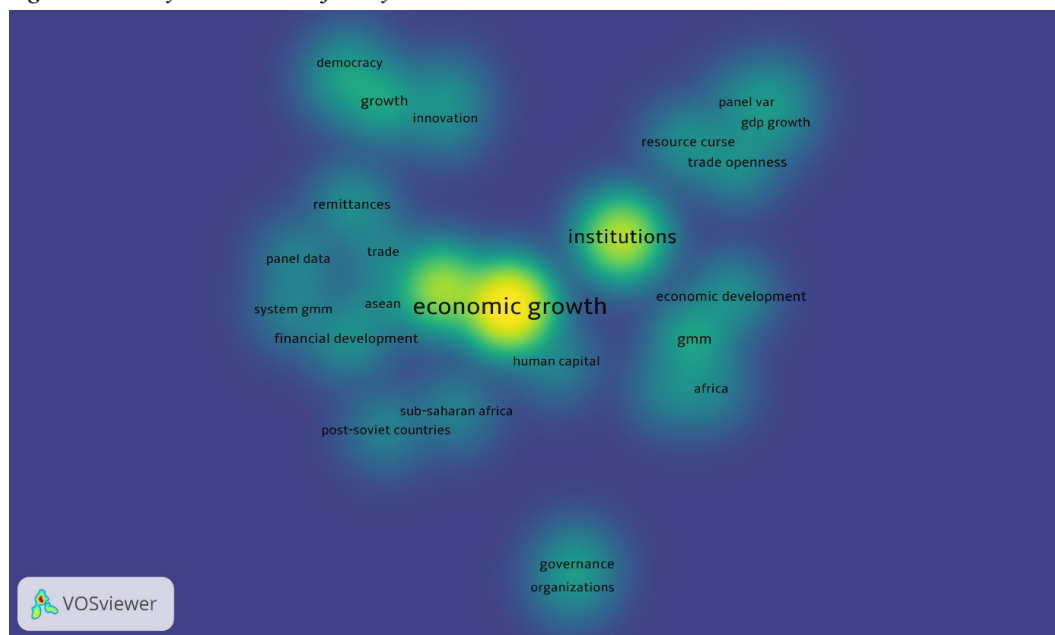
Source: Authors' calculations

Topics like "human capital", "economic development", and "governance" also stand out in yellow, indicating significant attention. These are crucial supporting concepts for understanding how institutional frameworks and the development of human capital contribute to broader economic outcomes.

Topics such as "remittances", "system GMM", "panel data", "innovation", and "democracy" are found in greenish areas in moderate-density areas. These terms denote a moderate research interest, often associated with empirical studies on how financial transactions, governance frameworks, and techno-

logical advancements influence economic development through institutional mechanisms.

The keywords "trade openness", "resource curse", and "GDP growth" appear in blue or light green areas, indicating that they have received less attention compared to the core topics. Nonetheless, their inclusion signifies an interest in understanding how policies, such as trade liberalisation and resource management, interact with institutional quality to impact economic growth.

The use of terms like "ASEAN", "sub-Saharan Africa", and "post-Soviet countries" is less prominent, indicating a scarcer literature on how institutional issues impact economic performance in specific regions. This implies that while regional studies are valuable, they occupy a somewhat peripheral position compared to more expansive discussions on institutional economics.

To sum up, the density visualisation indicates that "economic growth" and "institutions" are the primary focal points in the literature, with considerable emphasis on interconnected ideas such as governance, human capital, and economic development. Conversely, more niche subjects like regional case studies and trade/resource management receive less attention but present opportunities for further research.

5. Conclusion

This paper presents a systematic literature review of articles, that thoroughly examines the scholarly landscape concerning the impact of institutions on economic growth. The analysis is built upon the PRISMA protocol and offers a comprehensive understanding of the topic. The sample for analysis includes 78 articles from the Scopus database, published between 2006 and 2023. Our approach has three main components: a review of publication patterns, a sentiment analysis of abstracts, and a network analysis of keyword co-occurrence. These methods provided insights into important research trends, the overall sentiment of the existing literature, and the interconnections among key concepts in the field.

The analysis indicates a steady increase in research interest over time, with peak interest in certain recent years, reflecting the growing recognition of the significant role institutions play in shaping economic outcomes. Prominent journals and well-cited papers have played a significant role in advancing our un-

derstanding of the impact of institutions on economic growth. The sentiment analysis indicates that most of the literature holds a positive view on the impact of institutions on economic growth. While negative sentiments are relatively infrequent, they do highlight specific challenges in institutional reforms, particularly in developing nations. The keyword co-occurrence analysis reveals that discussions on economic growth consistently revolve around institutions, governance, and human capital. The analysis underscores the field's rigorous and adaptable approach, as novel concepts and tools are integrated to address the intricacies of institutional economics. Emerging areas of focus include the application of econometric techniques such as system GMM, as well as region-specific studies.

The findings provide a basis for some policy recommendations. To enhance economic performance, policymakers should prioritise institutional reforms aimed at improving governance, curbing corruption, refining regulatory frameworks, to establish an environment that encourages economic stability and attracts domestic and foreign investment, ultimately nurturing long-term growth. Policymakers need to leverage the positive sentiment by persisting in the implementation of reforms aimed at enhancing institutional effectiveness, transparency and accountability. The findings also suggest prioritising education and skills development to ensure a well-equipped workforce capable of driving productivity and innovation. Furthermore, they underscore the importance of carefully managing trade policies and natural resources to avoid obstacles that could impede growth. Particularly, countries abundant in natural resources should strengthen their governance structures to mitigate the negative economic impacts of resource dependence. Additionally, the increasing focus on advanced econometric methods indicates the significance of evidence-based policymaking. Governments should invest in developing statistical and analytical capacity to enable more sophisticated, data-driven policy interventions.

Future research should seek to deepen our understanding of institutional impacts by conducting comparative regional studies, particularly across different political and economic contexts. These studies can offer insights into how diverse institutional structures influence growth. Exploring the barriers to successful institutional reforms, especially in developing regions grappling with governance challenges and

corruption, can help fill the gaps in the current literature. In terms of regions, other transition countries (beside post-Soviet countries) could offer interesting insights, as they have undergone significant institutional reforms from the 1990s onward. This is particularly important for the Western Balkan countries that struggle to enhance their institutions and catch up with Central and Eastern European countries in terms of level of economic development and quality of life. With the increasing prevalence of advanced econometric methods like system GMM, there is an opportunity to incorporate innovative empirical approaches, such as machine learning or big data analysis, to analyse complex datasets. Furthermore, future research should prioritise longitudinal studies that track the long-term effects of institutional reforms, enabling policymakers to comprehend how institutional changes unfold over time. Investigating less commonly explored keywords, such as "regional development" and "resource curse", may provide unexplored perspectives on institutional impacts in specific sectors or regions.

Considering the limitations of the study, a future attempt at a systematic literature review can be improved by expanding the search parameters; removing the quotation marks would allow for a more comprehensive literature analysis. This would enable the inclusion of papers that discuss related topics, perhaps with different terminology or in more nuanced ways, thereby providing a complete understanding of how institutions impact economic growth across various fields and regions. Further, exploring other scientific databases, in addition to Scopus, can enhance the quality of research by providing broader and more detailed datasets.

References

1. Abdulahi, M. E., Shu, Y. & Khan, M. A. (2019). Resource rents, economic growth, and the role of institutional quality: A panel threshold analysis. *Resources Policy*, *61*, 293-303. https://doi.org/10.1016/j.resourpol.2019.02.011

2. Acemoglu, D., Johnson, S. & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*(5), 1369-1401. https://doi.org/10.1257/aer.91.5.1369

3. Acemoglu, D. & Robinson, J. A. (2012). *Why Nations Fail.* Crown Business.

4. Acheampong, A. O., Dzator, J. & Savage, D. A. (2021). Renewable energy, CO2 emissions and economic growth in sub-Saharan Africa: Does institutional quality matter? *Journal of Policy Modelling*, *43*(5), 1070-1093. https://doi.org/10.1016/j.jpolmod.2021.03.011

5. Afonso, O. (2020). The impact of institutions on economic growth in OECD countries. *Applied Economics Letters*, *29*(1), 63-67. https://doi.org/10.1080/13504851.2020.1855304

6. Ahmad, M. & Hall, S. G. (2017). Economic growth and convergence: Do institutional proximity and spillovers matter? *Journal of Policy Modelling*, *39*(6), 1065-1085. https://doi.org/10.1016/j.jpolmod.2017.07.001

7. Ahmed, F., Kousar, S., Pervaiz, A. & Shabbir, A. (2022). Do institutional quality and financial development affect sustainable economic growth? Evidence from South Asian countries. *Borsa Istanbul Review*, *22*(1), 189-196. https://doi.org/10.1016/j.bir.2021.03.005

8. Antonakakis, N., Cunado, J., Filis, G. & Gracia, F. P. de. (2017). Oil dependence, quality of political institutions and economic growth: A panel VAR approach. *Resources Policy*, *53*, 147-163. https://doi.org/10.1016/j.resourpol.2017.06.005

9. Aparicio, S., Urbano, D. & Audretsch, D. (2016). Institutional factors, opportunity entrepreneurship and economic growth: Panel data evidence. *Technological Forecasting and Social Change*, *102*, 45-61. https://doi.org/10.1016/j.techfore.2015.04.006

10. Arvin, M. B., Pradhan, R. P. & Nair, M. S. (2021). Are there links between institutional quality, government expenditure, tax revenue and economic growth? Evidence from low-income and lower middle-income countries. *Economic Analysis and Policy*, *70*, 468-489. https://doi.org/10.1016/j.eap.2021.03.011

11. Boateng, E., Agbola, F. W. & Mahmood, A. (2021). Foreign aid volatility and economic growth in Sub-Saharan Africa: Does institutional quality matter?. *Economic Modelling*, *96*, 111-127. https://doi.org/10.1016/j.econmod.2020.12.032

12. Brunnschweiler, C. N. (2008). Cursing the blessings? Natural resource abundance, institutions, and economic growth. *World Development*, *36*(3), 399-419. https://doi.org/10.1016/j.worlddev.2007.03.004

13. Butkiewicz, J. L. & Yanikkaya, H. (2006). Institutional quality and economic growth: Maintenance of the rule of law or democratic institutions, or both?. *Economic Modelling*, *23*(4), 648-661. https://doi.org/10.1016/j.econmod.2006.03.004

14. Catrinescu, N., Leon-Ledesma, M. Piracha, M. & Quillin, B. (2009). Remittances, institutions, and economic growth. *World Development*, *37*(1), 81-92. https://doi.org/10.1016/j.worlddev.2008.02.004

15. Doré, N. I. & Teixeira, A. A. C. (2023). The role of human capital, structural change, and institutional quality on Brazil's economic growth over the last two hundred years (1822-2019). *Structural Change and Economic Dynamics*, *66*, 1-12. https://doi.org/10.1016/j.strueco.2023.04.003

16. Fatás, A. & Mihov, I. (2013). Policy volatility, institutions, and economic growth. *Review of Economics and Statistics*, *95*(2), 362-376.

17. Flachaire, E., García-Peñalosa, C. & Konte, M. (2014). Political versus economic institutions in the growth process. *Journal of Comparative Economics*, *42*(1), 212-229. https://doi.org/10.1016/j.jce.2013.05.001

18. Gwartney, J. D., Holcombe, R. G. & Lawson, R. A. (2006). Institutions and the Impact of Investment on Growth. *Kyklos*, *59*(2), 255-273. https://doi.org/10.1111/j.1467-6435.2006.00327.x

19. Jiang, Y., Sharif, A., Anwar, A., Cong, P. T., Lelchumanan, B., Yen, V. T. & Vinh, N. T. T. (2023). Does green growth in E-7 countries depend on economic policy uncertainty, institutional quality, and renewable energy? Evidence from quantile-based regression. *Geoscience Frontiers*, *14*(6), 101652. https://doi.org/10.1016/j.gsf.2023.101652

20. Kaufmann, D., Kraay, A. & Mastruzzi, M. (2009). *Governance Matters VIII: Aggregate and Individual Governance Indicators 1996-2008* (World Bank Policy Research Working Paper No. 4978). Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-4978

21. Knack, S. & Keefer, P. (1995). Institutions and economic performance: cross-country tests using alternative institutional measures. *Economics & Politics*, *7*(3), 207-227. https://doi.org/10.1111/j.1468-0343.1995.tb00111.x

22. La Porta, R., Lopez-de-Silanes, F., Shleifer, A. & Vishny, R. (1999). The quality of government. *Journal of Law, Economics, and Organization*, *15*(1), 222-279. https://doi.org/10.1093/jleo/15.1.222

23. Maruta, A. A., Banerjee, R. & Cavoli, T. (2020). Foreign aid, institutional quality and economic growth: Evidence from the developing world. *Economic Modelling*, *89*, 444-463. https://doi.org/10.1016/j.econmod.2019.11.008

24. Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, *339*, b2535. https://doi.org/10.1136/bmj.b2535

25. Nair, M., Arvin, M. B., Pradhan, R. P. & Bahmani, S. (2021). Is higher economic growth possible through better institutional quality and a lower carbon footprint? Evidence from developing countries. *Renewable Energy*, *167*, 132-145. https://doi.org/10.1016/j.renene.2020.11.056

26. North, D. C. (1987). Institutions, Transaction Costs and Economic Growth. *Economic Inquiry*, *25*(3), 419-428. https://doi.org/10.1111/j.1465-7295.1987.tb00750.x

27. North, D. C. (1989). Institutions and economic growth: An historical introduction. *World Development*, *17*(9), 1319-1332. https://doi.org/10.1016/0305-750X(89)90075-2

28. North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808678

29. Qiang, Q. & Jian, C. (2020). Natural resource endowment, institutional quality and China's regional economic growth. *Resources Policy*, *66*, 101644. https://doi.org/10.1016/j.resourpol.2020.101644

30. Rodrik, D. (2000). Institutions for high-quality growth: What they are and how to acquire them. *Studies in Comparative International Development*, *35*(3), 3-31. https://doi.org/10.1007/BF02699764

31. Rodrik, D. (2007). *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth*. Princeton University Press. https://doi.org/10.2307/j.ctvcm4jbh

32. Rodrik, D., Subramanian, A. & Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, *9*, 131-165. https://doi.org/10.1023/B:JOEG.0000031425.72248.85

33. Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy*, *94*(5), 1002-1037.

34. Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, *98*(5), S71-S102. https://doi.org/10.1086/261725

35. Siddiqui, D. A. & Ahmed, Q. M. (2013). The effect of institutions on economic growth: A global analysis based on GMM dynamic panel estimation. *Structural Change and Economic Dynamics*, *24*, 18-33. https://doi.org/10.1016/j.strueco.2012.12.001

36. Tebaldi, E. & Elmslie, B. (2013). Does institutional quality impact innovation? Evidence from cross-country patent grant data. *Applied Economics*, *45*(7), 887-900. https://doi.org/10.1080/00036846.2011.613777

37. Valeriani, E. & Peluso, S. (2011). The Impact of Institutional Quality on Economic Growth and Development: An Empirical Study. *Journal of Knowledge Management, Economics, and Information Technology*, *1*(6), 1-25.

38. Zallé, O. (2019). Natural resources and economic growth in Africa: The role of institutional quality and human capital. *Resources Policy*, *62*, 616-624. https://doi.org/10.1016/j.resourpol.2018.11.009

Conference review

Jerko Glavaš, Ivana Unukić
Twentieth conference on Interdisciplinary Management Research – Osijek, IMR XX 2024

Jerko Glavaš
Josip Juraj Strossmayer
University of Osijek
Faculty of Economics
and Business in Osijek
31000 Osijek, Croatia
jerko.glavas@efos.hr

Ivana Unukić
Josip Juraj Strossmayer
University of Osijek
Faculty of Economics
and Business in Osijek
31000 Osijek, Croatia
ivana.unukic@efos.hr

TWENTIETH CONFERENCE ON INTERDISCIPLINARY MANAGEMENT RESEARCH – OSIJEK, IMR XX 2024

IMR 2024: At the heart of the Slavonian Plain, Osijek hosted the twentieth edition of the Interdisciplinary Management Research (IMR) conference that took place on 24 and 25 October 2024. This prestigious gathering, which has been bringing together the academic community, researchers, professionals and doctoral students for two decades, was organized in cooperation with the Faculty of Economics and Business in Osijek, Pforzheim University (Germany), and the Faculty of Economics at the University of Bihać. The IMR XX conference confirmed its status as a leading platform for the exchange of ideas and interdisciplinary research achievements on the international scene, inspiring all attendees with its tradition and influence.

Day 1 of the conference – Artificial Intelligence as a New Business Horizon

The first day of the conference began with a ceremonial opening in the Glagolitic Hall, where the attendees enjoyed the inspiring speeches of the hosts and the organizers. The Dean of the Faculty of Economics and Business in Osijek, Boris Crnković, PhD, together with other distinguished guests, expressed his gratitude and pride in the twenty-year tradition of the IMR conference. In the spirit of the current global challenges, this year's opening was particularly characterized by the Business Forum on the topic of "Artificial Intelligence – a New Era of Business". Robert Kopal, PhD, CEO of Effectus, inspired the audience with his presentation, in which he explained the importance and potential of AI technologies for shaping the future of business.

The introductory lecture was followed by a dynamic panel discussion in which Kopal, PhD, was joined by leading experts in the fields of artificial intelligence, communication and business development. The discussion on artificial intelligence as a disruptive technology, further enriched by insights from Juraj Bilić of CARNET, Marijana Šarolić Robić from the CRO Startup Association, and other experts, opened up new perspectives and emphasized the key issues of adapting business models to modern technological changes.

Academic dialog and a wealth of research topics

The afternoon of the first day saw a series of parallel sessions in which participants presented their research on topics such as public communication and tourism, corporate responsibility and economic models for high-rise construction. With their presentations, the researchers stimulated numerous discussions and raised questions about the impact of new technologies and communication trends on different sectors. The participants also had the

opportunity to explore the historic fortress Tvrđa, which further enriched this event with the local culture and history of Osijek.

Day 2 of the conference – From Tradition to the Future

The second day featured the continuation of inspiring parallel sessions with topics such as corruption in transition countries, teleworking after the pandemic, and the socio-economic impact of digitalization. The richness of this year's research topics clearly demonstrated the importance of an inter-disciplinary approach, and the participants from 8 countries provided fertile ground for further academic discussions and collaborations through the presentation of over 70 research papers.

All papers presented are accompanied by a collection of papers indexed in the world's leading databases, underlining the academic importance and global reach of the IMR conference. With its rich tradition and exceptional organization, IMR XX 2024 confirms that interdisciplinary research plays a key role in modern management and business through the exchange of ideas and collaboration between different professions.